WPS7324

Policy Research Working Paper 7324

Do Improved Biomass Cookstoves Reduce Fuelwood Consumption and Carbon Emissions?

Evidence from Rural Ethiopia Using a Randomized Treatment Trial with Electronic Monitoring

Abebe Beyene
Randy Bluffstone
Zenebe Gebreegzhiaber
Peter Martinsson
Alemu Mekonnen
Ferdinand Vieider

WORLD BANK GROUP
Development Research Group
Environment and Energy Team
June 2015

Abstract

This paper uses a randomized experimental design with real-time electronic stove temperature measurements and controlled cooking tests to estimate the fuelwood and carbon dioxide savings from an improved cookstove program in the process of being implemented in rural Ethiopia. Knowing more about how households interact with improved cookstoves is important, because cooking uses a majority of the fuelwood in the country and therefore is an important determinant of greenhouse gas emissions and indoor air pollution. Creating local networks among stove users generally appears to increase fuelwood savings, and among monetary treatments the most robust positive effects come from free distribution. The paper estimates that on average one improved stove saves approximately 634 kilograms of fuelwood per year or about 0.94 tons of carbon dioxide equivalent per year, which is about half of previous estimates. Using the May 2015 California auction price of $13.39/ton, the carbon sequestration from each stove deployed is worth about $12.59. Such carbon market offset revenues would be sufficient to cover the cost of the stove within one year.

This paper is a product of the Environment and Energy Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at alemu_m2004@yahoo.com.

Do Improved Biomass Cookstoves Reduce Fuelwood Consumption and Carbon Emissions? Evidence from Rural Ethiopia Using a Randomized Treatment Trial with Electronic Monitoring [1]

Abebe Beyene[A]

Randy Bluffstone[B]

Zenebe Gebreegzhiaber[A, C]

Peter Martinsson[D]

Alemu Mekonnen[A, E]
(corresponding author)

Ferdinand Vieider[F]

Keywords: improved biomass cookstoves; controlled cooking test; RCT; REDD+; Ethiopia.

JEL classification: C93, D12, O13, Q41, Q56

---

[1]Authorship is shared and authors are listed in alphabetical order. [A]Environmental Economics Policy Forum for Ethiopia (EEPFE)/ Ethiopian Development Research Institute (EDRI) Addis Ababa, Ethiopia; [B]Department of Economics Portland State University USA; [C]Department of Economics Mekelle; [D]University of Gothenburg, Sweden; [E]Department of Economics, Addis Ababa University (alemu_m2004@yahoo.com); [F]WZB Berlin Social Science Center, Risk & Development Unit.

Acknowledgements: The authors would like to thank the World Bank for funding the research through the Knowledge for Change Program and the Trust Fund for Environmentally and Socially Sustainable Development. We also acknowledge the Berkeley Air Monitoring Group for supplying and advising on the use of the electronic stove use monitors used in this study. Ermias Dessie provided expert data processing that allowed the data to be analyzed.

Do Improved Biomass Cookstoves Reduce Fuelwood Consumption and Carbon Emissions? Evidence from Rural Ethiopia Using a Randomized Treatment Trial with Electronic Monitoring

1. Introduction

With increased incomes and population, energy consumption and associated greenhouse gas emissions in the world are projected to increase and virtually all of the projected increase over the coming quarter century is expected to come from the developing world (Wolfram et al. 2012). This paper focuses on household cooking energy, which comprises about 90% of total primary energy demand in low-income countries like Ethiopia. In Ethiopia about 96% of households primarily use biomass fuels and an estimated 85% rely on fuelwood (Bizzarri, 2010). Within the household sector, cooking comprises virtually 100% of energy demand (Practical Action Ethiopia, 2015). As almost 3 billion people worldwide cook with biomass on a daily basis, biomass makes up about 40% of all household energy demand. Without major policy changes, the number and share of people cooking with biomass fuels are expected to remain virtually constant through 2030 (IEA 2012).

Relying on biomass fuels for cooking has a number of important implications. First, because biomass tends to be burned in homes, cooking with biomass typically generates indoor air pollution, which can affect respiratory health. Approximately 3.5 million people are estimated to die annually due to indoor air pollution, almost all of whom are in low-income countries (Jeuland et al., forthcoming). An additional 0.5 million in premature mortality comes from the outdoor air pollution originating from biomass cooking (Lim et al., 2010). Collecting fuelwood can also be time-consuming, which may divert scarce household labor resources (particularly women and children). If wood is not sustainably harvested, which is the case in most developing countries, biomass combustion releases CO_2 into the atmosphere, contributing to climate change.

Reducing particularly unsustainably harvested biomass energy use in low-income countries is highly desirable from a number of perspectives. Shifting to alternative energy sources is certainly the end goal, but reliance on biomass fuels is likely to continue for the vast majority of people in sub-Saharan Africa and Ethiopia in particular. Commercial energy options, such as natural gas, LPG and electricity, require major public infrastructure investments, supply chain development and purchase of expensive stoves. From the household perspective, these fuels and technologies can be extremely expensive and often unreliable in supply.

Improved cooking stoves (ICS) that use less biomass have therefore received significant attention as important intermediate technologies (Jeuland and Pattanayak, 2012), most of which use fuelwood, which is the most important biomass fuel. ICS have important advantages, because they typically do not involve sophisticated technologies and may require only minor changes in household cooking habits. These important features can make them very attractive if households really are able to cook meals with less wood, ICS are adopted and are regularly used by households. Because of the potential medium-run advantages, Ethiopia has made a particularly important commitment to ICS by including them in official strategy documents. The Climate Resilient Green Economy strategy (FDRE, 2011), for example, commits to promoting ICS that will be used by about 20 million households and reduce greenhouse gas emissions by almost 35 Mt CO_2e by 2030. "Reducing the demand for biomass by increasing fuel efficiency" is indeed currently seen as one of the few strategic priorities in the Ethiopian energy sector" (FDRE 2014).

The Ethiopian government is actively promoting the use of ICS and has announced its intention to distribute 9.4 million improved cookstoves by mid-decade. Most of these are Mirt stoves (which means 'best' in Amharic language). The Mirt stove is used primarily to bake *injera,* which is the main staple bread in Ethiopia and the heart of the national cuisine. Cooking *injera* uses an estimated 50% - 60% of household cooking fuel (Practical Action Ethiopia; Bizzarri, 2010). As about 90% of Ethiopia's total energy demand is from the household sector, baking *injera* consumes roughly half the nation's energy.

Our paper is perhaps the first to use a randomized design with real-time stove monitoring data matched with *in situ* measurement of fuelwood savings per-meal and per-minute of cooking time to identify the fuelwood savings of an ICS intervention. We use controlled cooking tests (CCTs) implemented by cooks in their own homes over three periods to actually measure the per-time-period fuelwood savings from the Mirt stove. To reduce total fuelwood demand, households must not only save wood when the technology is used, but households must actually substitute ICS for less-efficient traditional technologies. We measure the frequency with which Mirt stoves are used with real time electronic stove use monitors that measure the temperatures of our randomly distributed Mirt stoves every 10 minutes during four 37 to 56-day periods over more than one year. In the paper we convert the fuelwood savings to CO_2 emissions reductions using emissions factors and protocols from the Intergovernmental Panel on Climate Change (IPCC) and the Clean Development Mechanism (CDM).

The literature suggests that the payment terms offered and institutional structures within which ICS are distributed can affect stove adoption, use and per-meal/per-minute fuelwood savings (Mobarak et al. 2012). We therefore randomly assign six treatments at the village level to specifically test whether payment terms and institutional structures affect fuelwood savings. We estimate that across the three regional states (making up about 80% of the country's population of over 90 million) from which we randomly sample households, accounting for both per-meal/per-minute savings and in-field monitored stove use over time, the Mirt stove on average reduces fuelwood consumption by 634 kilograms per household per year. Using CDM and IPCC emissions factors, this savings implies 0.94 fewer tons of CO_2 emissions per household per year. We also cautiously conclude that fuelwood savings depend on the treatments applied, with free distribution of Mirt stoves rather than requiring payment and using social networking, generally increasing fuelwood savings.

The plan for the rest of the paper is as follows. Section 2 briefly discusses previous literature on estimating fuelwood savings and greenhouse gas emissions from ICS. Section 3 presents our data, experimental design and identification strategy and Section 4 describes our results. Section 5 discusses implications of the research findings and concludes.

2. Literature on Identifying ICS Fuelwood Savings, the Effects of Behavioral Interventions and Greenhouse Gas Emissions Reductions

Estimating the effect of ICS interventions on total greenhouse gas emissions requires identification of fuelwood savings from the use of improved cookstoves, the emissions reduction from each stove and the non-renewable fuelwood harvest adoption and uses ICS displaces. Each of these factors is discussed below with a primary emphasis on identifying fuelwood savings.

2.1 Fuelwood savings

Estimating the benefits of any policy intervention requires a credible counterfactual, which describes what would have occurred had the intervention not taken place. This critical step in program evaluation requires that as much as possible be held constant except for the policy intervention. Implementing various types of randomized treatment trials, which is the approach used in this paper, are particularly useful, because they explicitly limit potential confounders (Duflo et al., 2008). There is therefore increasing research using randomized designs (e.g. Bensch and Peters, 2013, Pattanayak et al., 2014). Mobarak et al. (2012), for example, analyze the determinants of low demand for improved cookstoves. They find that rural women rely overwhelmingly on free traditional cookstove technologies and are not willing to pay much for new cookstoves. They point to the need to design nontraditional cookstoves with features, such as reduced operating costs, that may be valued highly by users. Very little work has been done that includes credible estimates on whether improved stoves reduce carbon emissions, though Johnson et al. (2009; 2010) is a notable exception.

Possible approaches to measure how much fuelwood is used per unit of time include randomized kitchen performance (KPT), controlled cooking (CCT) and water boiling (WBT) tests. Each has its advantages and drawbacks (Lee et al. 2013). The KPT measures total fuelwood use in a household several times before and after an ICS intervention. The advantage is that it can account for the technology combinations that imply leakage. KPTs, however, have the disadvantage that little is known about the actual cooking use and mechanisms after adoption. It also requires virtually complete on-site measurement of fuelwood use, otherwise measurement errors can be large. Because of the intensive nature of the required measurement, sample sizes tend to be small. This can be a very serious problem if ICS have multiple uses, because in regression models it may not be possible to effectively adjust for potential confounders. Alternatively, and less effectively, experimental subjects can keep fuelwood use logs pre and post intervention. Such an approach has the risk of serious reporting errors that can perhaps only be obviated through direct measurement.

The other two tests focus on differences in fuelwood consumption per unit of time. The WBT measures the fuelwood and time needed to boil a fixed quantity of water. It is the simplest test and assures that cooking tasks are exactly the same for traditional and intervention stoves, but it also has little to do with actual cooking and typically is done in labs. The CCT also focuses on fuelwood consumption per unit of time, but unlike the WBT it examines the cooking of a standardized meal cooked in the home prepared by actual cooks. CCTs provide data on the quantity of fuelwood used per unit of food cooked and per unit of time. Though conducted *in situ*, the CCT only measures wood used and saved to cook typical meals. CCTs are therefore not very appropriate when actual meals deviate substantially from the "standardized" meal. Under such circumstances the measured savings may not be any more related to actual *in situ* cooking than the WBT.

Using a CCT-based method that is similar to the one used in this paper, Burwen and Levine (2012) evaluate a locally made and designed wood burning cookstove in Ghana. They find that the stoves on average reduce fuelwood to cook a standardized meal by 12%. Using electronic stove use monitors they also find that in general the stoves were used very frequently.

Bensch and Peters (2013) conduct an *ex post* evaluation of the Jambar charcoal stove in Senegal. Using OLS and propensity score matching, they find that the Jambar stove reduces charcoal to cook typical meals by 25% compared with traditional stoves. In a randomized control trial also in Senegal, the same authors find that fuel saving per meal cooked on the Jambar stove compared with a control group is even higher at 48%. They also find co-benefits associated with the virtually 100% usage of the stoves, including fewer eye infections and less cooking time.

Thakuri (2009) evaluates a customized improved wood burning stove and vent hood in a sample of 400 Nepali households. Though the data are observational and firewood consumption is based on 24-hour respondent recall, he estimates that improved cookstoves use 42% less wood than traditional stoves. Also based on reported data is Nepal et al. (2010), who use the nationally representative 2003/2004 Nepal Living Standards Survey to estimate the effect of improved stove use on fuelwood consumption. Using OLS, 2SLS and village level fixed effects, they conclude that user-identified ICS do not reduce fuelwood use. Evaluations of other stoves find fuelwood reductions (e.g. Johnson et al, 2009; Smith et al, 2007; Masera et al, 2007), but mass-produced brands are found to use only about 1/3 less than traditional stoves (Adkins et al., 2010).

In much of the above literature ICS fuelwood savings are estimated by measuring fuelwood consumption in traditional and improved stoves over an interval and comparing the two results, with the difference being how much fuelwood is saved. An important issue is leakage, because we know that households often incompletely adopt commercial fuels (e.g. Heltberg, 2004; Brouwer and Falcao, 2004). It therefore makes sense they may not completely substitute ICS for their traditional technologies. Careful and complete measurement of all cooking technologies in the home is therefore ideal.

The Mirt stove is designed primarily for cooking *injera*, which is the main staple in the study area and much of Ethiopia. We therefore see the Mirt stove as an ideal application of the CCT, because we know that on a day-to-day basis households will cook *injera*, which is exactly what is cooked in the CCT. We conduct CCTs applied to both traditional three-stone stoves and Mirt stoves in the same households on the same day three times over a period of a year. Our CCT sample consists of 108 households in 36 sites in rural Ethiopia, which allows us to capture systematic spatial and temporal differences in per-minute fuel savings.

As we analyze the same household's preparation of *injera* holding all things constant except the technology, we believe that we have utilized a particularly effective control in our estimates of fuel savings. Of course, the amount of *injera* batter is standardized (see Gebreegzhiaber et al., 2015 for additional details) and households may typically cook more or less than in the CCT. Total fuelwood savings estimates derived wholly from CCTs are therefore best viewed as indicative.

To adjust for such differences in cooking amounts and therefore get better estimates of total fuelwood savings, we incorporate *duration* and *frequency* of cooking events using electronic stove use monitors. One possibility for deriving stove use frequency information is to interview households as was done, for example, by Dresen et al. (2014) and Bluffstone (1998). However, with such survey data there is a very important tendency to overstate improved stove use frequency (e.g. see Thomas et al. forthcoming) and total fuelwood savings.

We avoid self-reported use and fuelwood savings by measuring the total number of actual cooking events per week and the total time households spend cooking using electronic stove use monitors discussed in detail in the following section. These devices measure the surface temperature of the Mirt stoves every 10 minutes allowing credible measurement of cooking event frequency. As cooking *injera* typically takes about one hour, this logging interval is short enough that we believe we can make reasonable estimates (albeit with some nonsystematic estimation error) of household cooking duration and incorporate that information into our fuelwood savings estimates.

This feature also allows us to explicitly adjust – using measurements – for a variety of unobservable factors that affect cooking time. For example, bringing in measured cooking duration allows us to explicitly incorporate unobservables such as particularly hungry family members and cooks who are less attentive to their cooking fires. Both these factors – largely unobservable – increase baseline fuelwood use and therefore the savings from the Mirt stove.

2.2. Effects of Behavioral Interventions

The literature suggests that the terms under which technologies are provided are important. Miller and Mobarak (2014) investigate the promotion of improved stoves in Bangladesh through social networks. They find a positive effect from learning through social networks, but a negative effect if a local leader promotes the stove (for similar findings see BenYishay and Mobarak, 2014; Munshi, 2004). Interestingly, Cohen and Dupas (2010) find that subjects who pay for treated mosquito bed nets use them more frequently than those who received them for free.

Studies have shown tendencies toward decreased use over time going back to Jones (1988). As Shankar et al. (2014) emphasize, the goal of promoting "improved" cookstoves should be on "correct and consistent use" rather than "acquisition." They and others note that educating users about the advantages of improved biomass cookstoves and promoting them are typically insufficient. Instead, stoves need to be truly appropriate and offer net benefits in the long-run (e.g. Jeuland and Pattanayak, 2012; Mobarak et al., 2012). Outside the ICS literature, using evidence from a large energy efficiency program in the United States, Alcott and Rogers (2014) analyze the long and short run effects of providing social comparison information to energy users. They find that the intervention quickly increases energy efficiency behaviors, but these effects decay once the intervention is removed. They note that this "backsliding" has important effects on the cost-effectiveness of such interventions.

2.3 Estimating Reduced CO_2 Emissions from ICS

CO_2 emissions from the use of cookstoves depend primarily on the technical specifications of the stoves and assumptions about fuels displaced. Under the Framework Convention on Climate Change non-Annex 1 countries, such as Ethiopia, do not have emissions reductions obligations. The relevant emissions reductions displaced are therefore in Annex 1 developed countries using commercial fuels. In 2012 the executive board of the CDM approved protocols for crediting ICS projects. As discussed in UNFCCC (2012), emissions reductions from ICS should be based on "fossil fuel emission factors of substitution fuels likely to be used by similar users." Lee et al. (2013) further suggests "… the use of a weighted average value of 81.6, representing a mix of 50% coal, 25% kerosene, and 25% LPG." In this paper we adopt these norms.

The contribution of fuelwood conservation to reduced forest degradation and therefore fewer carbon emissions depends on the nature of fuelwood harvest (Johnson et al. 2009; Lee et al. 2013). A key parameter is therefore the fraction of woody biomass used "that can be established as non-renewable biomass" (UNFCCC 2012). The degree to which biomass harvests are unsustainable depends primarily on the management regime. If fuelwood is taken under open access regimes, for example, very limited management, including replanting and harvest mitigation, is expected (Ostrom, 1990; Gordon, 1954). Such settings are likely to have close to 100% non-renewable biomass. Estimating non-renewable biomass with precision is difficult and an important area of current research (Bailis et al., 2015; Johnson et al., 2009; Lee et al., 2013). UNFCCC (2012) has therefore provided default estimates of percent non-renewable biomass for low-income countries and many countries in sub-Saharan Africa have default values above 90%. Perhaps because Ethiopia in 2007 passed important forestry legislation that has improved management (Beyene et al., forthcoming; Mekonnen and Bluffstone, 2014), the default value was set at 88%. This is the value used in this paper.

3. Data, Empirical Methods and Identification Strategy

This section discusses our sampling and data collection methods. It also presents our randomized experimental treatments, strategy for identifying fuelwood savings and method for converting fuelwood savings to CO_2 emissions reductions. These are discussed in the following three subsections.

3.1 Sampling and Data Collection

The sample sites are chosen from 110 villages examined during a 2012 study conducted by the Environmental Economics Policy Forum for Ethiopia (EEPFE) based at the Ethiopian Development Research Institute. The project collected critical village information in Amhara, Tigray, SNNP and Oromiya Regional States,[2] which we use to assign treatments and test randomization. Out of the 110 sites, we omit 15 sites that were covered during a pilot survey conducted as part of research for this paper. We also remove 14 sites either because the traditional technology deviates from the three-stone technology, which is the national norm or because *injera* is typically not cooked. The Mirt ICS is designed for baking *injera*, which is the most important use of cooking energy in Ethiopia. Three Borena District sites where *injera* baking is not common are therefore omitted from the sample. We also remove all sites (total of 11) from Tigray Regional State, because three-stone traditional stoves are not typically used.

From the 81 remaining sites we select 36 at random using proportionate random sampling based on regional state forest cover. A total of 20% of observations are sampled from Amhara, 50% from Oromiya and about 30% from SNNP regional states. These percentages also approximate the population distribution across the three regions. Forest cover is used as the sampling proportion, because most fuelwood in the sample areas come from forests and forest area provides a key measure of fuelwood supply. All sites have formal or informal forest user groups. It was at this

[2] The regional state is the most aggregated sub-national jurisdiction.

village level that six behavioral treatments were randomly assigned (6 villages per treatment). These treatments are discussed in subsection 3.2.

Sample households within each site were chosen using systematic random sampling. A total of 504 households were selected (14 households from each site), with 360 households (10 households from each site) receiving the improved Mirt stove and the rest identified as control households who did not receive the Mirt stove. As three CCTs were conducted on both Mirt and three-stone stoves on the same days by each household that received a Mirt stove and we measure Mirt stove use frequency and duration, in this paper, the control households are not relevant. We therefore only examine per-minute fuelwood consumption and measured use duration and frequency for households who received the Mirt stove.

Five supervisors with significant field experience were deployed in each round of field work (each covering 7 – 8 sites). Five enumerators worked under each supervisor. These 30 individuals implemented the research in all four stove use monitoring rounds and three CCT rounds of the study. After identifying sample households, respondents were told by fieldworkers that they were chosen randomly to receive a stove under the same terms as others in their village and they were informed of the terms (i.e. the behavioral treatment). Fieldworkers also gave respondents full information on the stove features. Figures 1 and 2 show the Mirt cookstove.

Figure 1
MIRT Stove with *Injera* Cooking

Source:ethiopiaethos.files.wordpress.com/2010/06/comp5.jpg

Figure 2
Cook Pouring *Injera* Batter on MIRT Stove

Source: energypedia.info/wiki/Baking_with_Improved_Ovens

Stove use frequency is measured using electronic stove use monitors (SUM). The SUM measures stove temperature and as shown in Figure 3 are approximately the size of a watch battery.

They were purchased from Berkeley Air Monitoring Group of Berkeley, California. The recording intervals on the SUM are adjustable and the memory on the DS1922T model used can record stove temperature every ten minutes for approximately 60 days and can tolerate temperatures up to 120 degrees Celsius. Logging software is loaded onto a computer and after completing the monitoring period the SUM is inserted into an interface and the temperature, time and date are downloaded for analysis. We define a cooking event as occurring when temperature exceeds 40 degrees Celsius.

Figure 3 Electronic Stove Use Monitor

Source: https://www.measurementsystems.co.uk/images/categories/main/miniature-data-loggers.jpg

Respondents were shown the SUM device and informed that should they agree to participate, these devices would be placed on the stoves using heat resistant tape and that they would periodically record the surface temperature of the stoves.[3] Fieldworkers also informed participants that stove use is not required to receive a stove, but they are encouraged to use them. Respondents were asked for their formal oral consent and all respondents agreed to participate in the study. There were no refusals and no attrition in the sample.

The SUM was placed exactly at the back of the stove, because it is the coolest location on the Mirt stove, using heat-resistant tape. It was initially believed that the front of the stove was the coolest and SUM devices were ruined, because their temperature tolerances were exceeded. Because of such malfunctions, the panel data used are not balanced, with the second monitoring period having the fewest observations. While over the four rounds a maximum of 1,440 observations (360 per round) are possible, in practice, because of stove use monitor failures, observations in each round average just over 300, with a total of 1,209 observations across all rounds.

[3]Fieldworkers informed respondents that the SUM devices are safe at reasonable temperatures, but they are potentially unsafe if they are put in or very close to fires, because they have flammable components inside them. If SUM fall in the fire, they should be removed immediately and after they are cool replaced on the top of the stove using the heat resistant tape.

Households were typically invited to a centralized location, such as a school or *kebele*[4] office, to receive their stoves and preliminary instructions from field supervisors. Villagers then took the six concrete stove pieces and the clay cooking plate back to their homes. This process is shown in Figure 4. Typically, the following day enumerators came to install the stove in the kitchen area, either in the main home (43%) or in a separate kitchen (53%) by mudding together the six concrete pieces. They also gave cooks training on stove operations and the SUM device was installed on the stove using heat-resistant tape and initiated to begin logging temperature.

Figure 4. Mirt Stove Recipient taking Stove Home

Source: Randall Bluffstone

3.2 *Behavioral Treatments*

The participating households received the stoves under six mutually-exclusive randomized behavioral treatments. There were three aspects to each treatment, with only two levels of each aspect (present or absent) divided equally so one-third of all sites received each treatment aspect. These aspects are: 1) payment for stove use; 2) cost of the stove and 3) networking.

In the first treatment aspect sites were randomly chosen for households to receive a 50 Birr payment if the SUM devices indicated that Mirt stoves were used at least twice per week during the first monitoring period. The 50 Birr payment was made after checking the recorded SUM data at the end of the first round (about 6 weeks after installation of the SUM device). In a manner similar to Charness and Gneezy (2009), this treatment aspect tests the hypothesis that use incentives increase fuelwood savings.

The second treatment aspect is cost. One-third of the sample paid 25 Birr for their Mirt stoves and the remainder received their stoves for free. This is about 13% of the real stove cost. This treatment aspect tests the same type of hypothesis examined by Cohen and Dupas (2010) that

[4] *Kebele* is the lowest administrative unit in Ethiopia and translates to "peasant association."

those who pay for their stoves get better outcomes. Other payment treatments were not used due to budgetary reasons.

The final treatment aspect is the network component. The 1/3 of respondents who received this aspect not only received in-home Mirt stove training, but also were brought together with others in their village for a group meeting with supervisors. The 10 villagers receiving the network treatment in each of the 12 villages were assembled in a common area in the village. The details of stove operation were reiterated and users had the opportunity to ask questions. This treatment aspect tests whether making those who received Mirt stoves aware of each other in a formal setting and allowing users to learn from and potentially network with each other increases fuelwood savings.

Households who were randomly selected received only one of the six treatments below, which remained constant across all four monitoring periods. One treatment is assigned to 6 sites, implying a total of 6 sites receiving the same treatment.

1. Household received 50 Birr stove use payment but no network;
2. Household received 50 Birr stove use payment plus network;
3. Household paid 25 Birr for the stove but received no network;
4. Household paid 25 Birr for the stove plus network;
5. Household received stove for free and no network, and
6. Household received stove for free plus network.

3.3. Strategy for Identifying Fuelwood Savings and Reduced CO_2 Emissions

To identify fuelwood savings we break total savings into household per-meal/per-minute fuelwood savings and frequency of ICS use. We take this approach rather than using a KPT for three reasons. First, using direct measurement allows us to identify the effects of our randomized behavioral treatments discussed in the previous section on actual use of the Mirt stove. Second, this approach offers the opportunity to separate per cooking-minute efficiency from the effect of cooking duration. Finally, in the previous subsection the implementation downsides of KPTs were discussed and we believe direct measurement helps avoid those potential confounders.

Per-minute fuelwood savings from the Mirt stove are identified using our CCTs. As discussed above, the 108 households chosen to participate in CCTs were randomly selected from our total sample of 360 households that received Mirt stoves under our 6 randomized treatments. There were no CCT refusals. CCT rounds were conducted in May/June 2013, October/November 2013 and May-July 2014. The strategy and protocols for identifying the per cooking-minute savings are discussed extensively in Gebreegziabher et al. (2015). Each of the three CCT rounds is conducted on both the Mirt stove and the household's own traditional three-stone stove on the same day cooking the same amount of *injera*. We identify the savings for a household on a particular day as simply the difference between the measured fuelwood used to cook the fixed amount of *injera* batter on the traditional stove compared with the same quantity on the Mirt stove. We believe that with this approach virtually all possible unobservables are adjusted for and confounders eliminated, because the only factor changing is the cooking technology.

During the CCTs we also carefully measure cooking time and use this information to calculate fuel savings per minute. This metric gives us the differential "burn rate" to cook *injera* on the Mirt stove vis-à-vis the three-stone technology. This is the metric from the CCTs that is combined with SUM information on cooking frequency and duration. The equation for computing total fuelwood savings per week is given in equation (1).

$$(1)\ \text{Fuelwood Savings per week} = \underbrace{\text{Fuelwood savings/minute cooking}}_{\text{From CCT}} * \underbrace{\text{Minutes of cooking per week}}_{\text{From SUM}}$$

CCTs measure standardized cooking events that may or may not be fully representative of actual cooking. Though the Mirt stove mainly cooks *injera*, so using the stove for other dishes is very unlikely, and we distributed 4 kilograms of *teff* flour for each CCT based on focus groups, households may cook more or less than this amount during cooking events. We address this concern by combining the measured per cooking-minute savings information with cooking frequency and duration measurements. We define frequency as the measured number of Mirt stove cooking event hours per week, where cooking events are identified if the Mirt stove surface temperature measured by the SUM exceeds 40 degrees Celsius. Using 40 degrees Celsius as the lower-limit for a cooking event obviates the possibility that normal temperature variation is confused with cooking. The maximum ambient temperature in the sample, measured using thermometers, is 35 degrees Celsius and the average measured ambient temperature is 25 degrees.

Using signal processing and analysis (O'Haver, 1997) we identify all cooking events and then calculate their durations. Cooking event duration is calculated as the total time in minutes starting from when the temperature is 40 degrees Celsius at the beginning of an identified cooking event until it comes down again to 40 degrees Celsius at the end of the event. As noted above, because the SUM record temperatures only every 10 minutes, there is likely some measurement error. This error is bounded and approximately 10 minutes in total out of about an hour average CCT cooking event time, assuming 5 minutes error on each of the starting and ending time measurements.

More importantly, because on average households use their Mirt stoves 15 times per measurement round and most households are measured over four rounds, these measurement errors average out. Moreover, we clearly know the direction of the bias. Measurement error can only <u>reduce</u> cooking duration, implying that our per-minute savings from the three rounds of CCTs are applied to <u>fewer</u> measured cooking minutes. The measurement error therefore has the effect of making our estimates conservative.

The SUM devices were initiated and set to record the temperature of the stove every ten minutes. This was done four times over the course of more than one year beginning in June 2013 and ending in July 2014. The four monitoring periods are June-August, 2013; August-October, 2013; March-May 2013 and May-July 2014. During each monitoring period the SUM devices recorded data for between 37 and 56 days. The median recording period was 49 days.

CCT data that match the timing of the monitoring periods were used for the first three monitoring periods. As we have four SUMS monitoring periods and only 3 CCT rounds, in the fourth SUM monitoring period the third round CCT per-minute fuelwood savings were applied to the fourth round SUM data. This approach is likely conservative, because on average total fuelwood savings from the CCTs monotonically increase and cooking time monotonically decreases over time. These together suggest improved performance over time, albeit at a decreasing rate between rounds 2 and 3.

We collect CCT data only for 3 of the 10 households in each of the 36 villages that actually received Mirt stoves and were monitored using SUM. This is because conducting the CCTs is extremely labor-intensive. Fuelwood and *teff* flour must be sourced in villages that may be far from roads. Each household CCT round then takes two to four hours to complete, because it is necessary for households to cook *injera* on both the Mirt and traditional stoves on the same day in the same kitchen. For those households who received Mirt stoves, but did not participate in CCTs, we use the village average per-minute fuelwood savings from the appropriate CCT round as noted above. We calibrate all information to one week time periods, because *injera* tends to not be cooked every day (Dresen et al., 2014). Other data on household characteristics, demographics and cooking environments, including cooking locations, are collected using a household survey.

It is unfortunately not possible to electronically monitor surface temperature of the three-stone technology, because the stones are often moved and disturbed, which can destroy monitors by force or by getting them too close to fires and exceeding temperature tolerances. We therefore implicitly assume that adjusting for differences in fuelwood consumption the time to cook a unit of *injera* is the same regardless of the technology. This exact assumption was tested during the CCTs and as discussed in Gebreegziabher et al. (2015) there is no statistical difference between the time to bake 4kg of *teff* flour into *injera* on the Mirt and traditional three-stone stoves.

We acknowledge that using the measured per-minute CCT savings as an estimate of per-unit time fuelwood savings supposes limited leakage in which households do not fully substitute the ICS for the traditional technology.[5] Indeed, it is even possible that fuel use could increase due to leakage and we acknowledge that our identification strategy does not strictly preclude large leakage effects. That said, three features of the Mirt stove mitigate leakage. First, the Mirt stove is highly specialized for *injera* baking, which makes it very unlikely that it will be shifted to other uses. As shown in Figures 1 and 2, for example, the cooking area is very large and cannot accommodate small pots. Waste gases can be utilized to cook stews and coffee, but the stove is designed for this to be done in conjunction with *injera* baking.

Second, using the SUM we find that the Mirt stove is on average used two to three times per week, which is a very traditional and extremely typical interval for *injera* baking in areas without refrigeration. Indeed it is also almost exactly the mean baking interval reported in Gebreegziabher et

[5] In the energy literature such leakage is often called "rebound" effects. In the results section we discuss the effect of alternative assumptions about leakage.

al. (2015). It is therefore very likely that on average most if not all *injera* baking (but perhaps not all cooking) is done on the Mirt stove and therefore captured by our monitoring.

Finally, it is very unlikely that any increase in real income due to Mirt stove adoption would be realized as significant increased cooking. Our study areas do not have refrigeration, so households that cook *injera* must eat them within one to three days to avoid possible bacterial infection. It is, of course, possible that as real income increases, households over time are able to add extra coffee or stews (typically eaten with *injera*) to their diets. These effects – due to the nature of the cuisine and the Mirt technology - are expected to be limited, however, causing us to not be overly concerned with leakage.

In terms of empirical methods, because our stove distribution is random at the household level and behavioral treatments are also randomized, we begin our empirical analysis with differences in means, rank-sum and differences in median tests. Using OLS with monitoring round and district fixed effects and (in Appendix A) random effects with district fixed effects, we also evaluate the factors affecting estimated fuel savings at the household level. Perhaps not surprisingly given our random sampling, randomized behavioral treatments and CCT experimental methodologies, which explicitly adjust for unobservable differences, few socio-demographic factors affect fuelwood savings. Table 1 presents the variables used in the regression models.

Table 1. Descriptive Statistics of Covariates used in the Pooled OLS Models

Variable	Obs.	Mean	Sd. D.	Min	Max
Age of respondent	1205	42.09	13.07	20	90
Sex of respondent 1 if male, 0 if female	1205	0.88	0.33	0	1
Marital of respondent 1 if married, 0 otherwise	1209	0.90	0.30	0	1
Education-1 if illiterate, 0 if literate	1209	0.39	0.49	0	1
Family Size in adult equivalent	1209	4.93	1.85	0	10.88
Religion -1 if Christian, 0 Muslim	1205	0.73	0.45	0	1
Livestock in tropical livestock units (TLU)	1194	5.00	3.56	0	26.23
Children under 15	1103	3.09	1.51	1	11
Walking Distance from household to nearest road (two-way) in minutes	1205	60.13	97.35	0	840
Participation in controlled cooking test, 1 if yes, 0 if no	1209	0.30	0.46	0	1
Risk coefficient	1197	3.80	1.17	0.08	6.58
Average temperature in ^{0}C	1203	24.75	3.16	15	31
Average number of injera baked at a time	1205	19.79	10.25	0	55
Use the of improved stove for any other purpose other	1209	0.35	0.48	0	1

than baking 1 if yes, 0 if No					
The average quantity of flour in kg used per cooking	1197	4.65	2.31	0.75	15
Type of flour, teff=1, mixed or no teff=0	1209	0.15	0.36	0	1
Place of stove installed - 1 if inside the house, 0 separate kitchen	1203	0.37	0.48	0	1
Main fuel used for baking - 1 if fuel wood, 0 otherwise	1209	0.74	0.44	0	1

A key objective of the analysis is to examine the effect of our six behavioral treatments, applied at the village level, on fuelwood savings. As the treatments are randomized, it is possible to directly compare the households receiving particular treatments with those receiving any and all other treatments. Kruskal-Wallis tests are employed based on 5 key community characteristics and we confirm that the treatments are from the same population; randomization therefore appears to have been largely successful.

Formally, we would therefore like to estimate the average effect of each treatment vis-à-vis all other treatments for the same participating households. That is, ideally, we would like to see the effect of all our 6 treatments on each randomly chosen household. This goal is given in equation 2, where i indicates the treatment, j is the counterfactual, T is the treatment indicator, Y_i is the treated outcome and Y_j is the untreated outcome.

$$(2) \quad ATT_{ij} = E(Y_i \mid T_i = 1) - E(Y_j \mid T_i = 1) \ \forall ij$$

The above criterion is, of course, unobservable, because $E(Y_j|T_i=1)$, which is the counterfactual, did not occur. Though our data are experimental, using observational outcomes purely from the untreated group has the potential to generate selection bias (Andersson et al., 2011). One possible solution is to use $E(Y_j|T_j=0)$, which is observable, as a proxy for $E(Y_j|T_i=1)$ if households in all six treatment groups are comparable on observable and unobservable features.

Though our village level randomization appears to have been successful, it is also true that our analysis is at the household level, but for reasons of fairness as discussed above, treatments had to be applied at the village level. It is therefore possible that some non-random elements were inserted. As a robustness check, we utilize household-level propensity score matching to create matched pairs that construct counterfactuals. Propensity score matching utilizes observables to estimate probit models of the probability that households receive a particular treatment.[6] The predicted values of these probit models are then used as propensity scores that match treatment and control households.

Virtually all the analytical effort in this paper focuses on estimating fuelwood savings. To convert estimated fuelwood savings to reduced carbon emissions we directly apply the standard CDM methodology discussed above that applies a CO_2 emissions factor to each kilogram of fuelwood saved and adjusts that value for fossil fuels displaced and fraction of fuelwood that should be viewed as non-renewable. The equation to calculate carbon emissions reduction (Lee et al. 2013) is given below:

$$ER_y = B_y * f_{NRB,y} * NCV_{biomass} * EF_{projected\_fossilfuel}$$

Where:

ER_y = Emissions reductions during year y in tCO_2e

B_y = Quantity of woody biomass saved in tons

$f_{NRB,y}$ = Fraction of woody biomass saved by the cookstove used in year y defined as non-renewable biomass

$NCV_{biomass}$ = Net calorific value of the non-renewable woody biomass that is substituted (IPCC default for wood fuel, 0.015 TJ/ton)

$EF_{projected\_fossilfuel}$= Emission factor for the substitution of non-renewable woody biomass by similar consumers

4. Results

We now present the results of our empirical analysis. We begin with a discussion of the estimated fuelwood savings per week attributable to use of the Mirt stoves rather than the three-stone tripod. These results are followed by analyses of fuelwood savings by treatment. We use non-parametric tests to evaluate whether fuelwood savings differ by treatment and also present the results of nearest neighbor propensity score matching followed by pooled OLS regressions.

[6] We emphasize that treatments are given at the village level and are assigned randomly. Households within sites were also chosen randomly. There is therefore no reason to believe treatments were non-random.

4.1 Fuelwood Savings of the Mirt Stove Compared to the Three-Stone Tripod

As shown in Table 2, we find significant average fuelwood savings per week using the Mirt stove. Across all rounds we estimate that the Mirt stove saved about 10.55 kilograms of fuelwood per week since households received their stoves. Average savings increased steadily across rounds, reaching 12.40 kilograms per week saved by round 4. By round 2 of the study, user experience and frequency of use had increased substantially and as shown in Table 1 by fall 2013 seventy-five percent of users had positive measured fuelwood savings. By round 4 in summer 2014, virtually all users had positive measured fuelwood savings using the Mirt stove.

Table 2 Fuelwood Savings Overall and by Round (in Kilograms per Week per Mirt Stove)

	Overall Sample	Round 1 (June-August 2013)	Round 2 (August - October 2013)	Round 3 (March-May 2014)	Round 4 (May – July 2014)
Mean	10.55	5.22	12.12	12.18	12.40
Minimum	-29.24	-29.24	-26.62	-25.07	-12.09
Maximum	68.59	58.94	63.06	68.59	67.86
10th percentile	-2.52	-7.95	-1.70	-1.07	-0.01
25th percentile	0	-1.31	0.155	0.012	0.01
50th percentile	3.98	1.48	5.55	4.96	4.55
75th percentile	16.60	11.05	19.00	21.26	21.01
90th percentile	38.51	22.92	39.93	41.83	39.58
Observations	1090	262	276	280	272

The review period preceding our cooking round 1 evaluation appears to have been one of startup both in terms of having the lowest average per-minute fuelwood savings according to the CCTs and lowest Mirt stove use as measured by the SUM. We therefore view average fuelwood savings from later rounds as more typical. For example, 12.20 kilograms per stove per week is a reasonable overall estimate of the fuelwood Mirt stoves save on average every week in the long-run compared with traditional stoves. As our sample covers a full year and respondents travel infrequently, the Mirt stove is estimated to save approximately 634 kilograms per year.[7]

[7] Leakage would scale this estimate down. For example, if with rebound effects net fuelwood savings is only 80% of estimated, the average savings per stove per year would be 507.20 kilograms per year. We also remind, however, that

This estimate is roughly half the 1277 kilogram estimate of Dresen et al. (2014), which used survey methods to calculate the fuelwood and carbon savings of the Mirt stove. As discussed by Thomas et al. (forthcoming), a lot of this discrepancy is likely due to overstating improved stove use. They find in Rwanda that respondents overstated their actual use in surveys by approximately 40%. We therefore believe that our estimates are more realistic.

Table 3 Mean Mirt Stove Fuelwood Savings in kilograms per Week by Regional State and Round

	Overall	Round 1 (June-August 2013)	Round 2 (August - October 2013)	Round 3 (March-May 2014)	Round 4 (May – July 2014)
Amhara	23.99	15.02	27.21	26.15	26.17
Oromiya	9.53	3.68	10.54	11.62	11.98
SNNP	3.32	1.65	3.63	4.07	3.97

Table 3 examines fuel savings by round and region. As was found in Gebreeziabher et al. (2015) and Beyene et al. (2015), significant differences exist across the regional states. In all rounds Amhara has significantly larger fuelwood savings than the other two regional states. This is because many fewer Mirt stoves in households in Amhara use more fuelwood than the traditional three-stone tripods. Whereas across the four rounds the smallest measured fuelwood savings was -0.71 kilograms (i.e. increased fuelwood use) per week in Amhara, in Oromiya and SNNP the smallest savings are -25.00 kilograms per week. As a result, while in Amhara at the 25th percentile the Mirt stove is estimated to save 10.00 kg per week, in Oromiya the value is 0.89 kilograms and in SNNP -1.91 kilograms per week. At the upper levels of the distributions, fuelwood savings are comparable across the three regional states. At this point we are not able to fully explain these inter-regional differences.

4.2 Effects of Randomized Behavioral Treatments on Fuelwood Savings

We now turn our attention to why the Mirt stove reduces fuelwood consumption, with a main emphasis on the six treatments equally implemented when the 360 stoves were randomly distributed to households. As discussed in the previous section, treatments were applied at the site

because we do not have continuous stove temperature monitoring (i.e. logging only every 10 minutes), there must be a downward bias in our cooking duration and therefore fuelwood savings estimates.

level so that within villages the distribution would be viewed as fair and equitable. For example, everyone who lived in a village paid or received the same amount of money for using the stoves. These treatments were assigned completely randomly, but it remains important to check that the intended randomization was successful.

Table 4
Kruskal-Wallis Tests that the Six Treatment Assignments Come from the Same Site Characteristic Distributions

	Households in *Kebele*[8]	Households in Forest User Group	Wealth Variance in Forest User Group	Existence of Forest Rules/ Regulations	% Forest Biomass Change Over 5 Years (Respondent Assessed)
X^2 (5)	5.246	3.085	6.166	.000	6.243
P Value	.387	.687	.290	1.000	.283

***, **, * indicate significant at the 1%, 5% and 10% levels

Table 4 presents Kruskal-Wallis test results for differences in community characteristics by treatment. In the interest of brevity, only the test statistics are presented, but mean ranks for all characteristics by treatment are available from the authors. The table assesses 5 different key community characteristics and the randomization appears to have been successful. There are no systematic differences in group size, wealth distribution, forest change, or forest management.

Table 5 SUM Monitoring Data by Treatment and Round

Treatment Description	Treatment Number	Round 1	Round 2	Round 3	Round 4
Received 50 Birr use payment, but no network	1	42	34	49	48
Received 50 Birr use payment, plus network	2	53	51	57	55
Paid 25 Birr, but received no network	3	49	49	57	56
Paid 25 Birr for the stove, plus network	4	53	55	54	51
Received stove for free and no network	5	53	52	51	51
Received stove for free, plus network	6	50	46	50	40
Total Observations by Round		300	287	318	301

Table 5 presents numbers of usable SUM observations by round and treatment out of a possible 360 SUM devices fielded in each round. SUM devices failed mainly because their temperature tolerances were exceeded, but some devices we removed by children and there was

[8] The Amharic word *kebele* translates as "peasant association." It is the lowest official administrative region in Ethiopia and is typically comprised of many villages

even a report that a chicken pecked one SUM device off a stove. We see that total SUM failures range from 42 in Round 3 to 73 in Round 2. The usable observations are therefore reasonably uniform across monitoring periods and highly uniform across treatments, with the possible exception of treatment 1, which has fewer usable observations in round 2.— Overall, these results suggest that any SUM failures were idiosyncratic and not systematic by treatment, which is very important.

Table 6 Mirt Stove Fuelwood Savings in kilograms per Week by Treatment

	Obs.	Median	Mean	S. D.	Max.
Received 50Birr use payment, but no network (Treatment 1)	157	6.87	9.91	13.52	59.41
Received 50 Birr use payment, plus network (Treatment 2)	199	4.63	13.34	17.64	67.40
Paid 25 Birr, but received no network (Treatment 3)	194	6.11	12.13	17.93	65.93
Paid 25 Birr for the stove, plus network (Treatment 4)	176	0.22	2.93	8.81	58.94
Received stove for free and no network (Treatment 5)	181	9.37	16.94	19.43	67.86
Received stove for free, plus network (Treatment 6)	179	2.33	7.50	16.84	68.59

We find that fuelwood savings differ by treatment, with households receiving treatments 2 and 5 having the highest estimated average fuelwood savings. As the distribution is right-side skewed, medians are below means and households receiving treatments 1, 3 and 5 (all without networking) have the highest estimated median fuelwood savings. These results are shown in Table 6. Relying on our randomized design, we now formally test whether households receiving different treatments have statistically different fuelwood savings. As our fuelwood savings data are skewed, we rely on nonparametric Mann-Whitney rank-sum tests. In Table 7 we also present tests of equalities in medians (to adjust for the skewed distributions). These non-parametric median equality tests are also more interpretable than the rank-sum tests.

Table 7
Rank-Sum Mann-Whitney Tests that Households Receiving Different Treatments have Fuelwood Savings that Come from Same Populations and have Equal Medians(p Values)

	Treatment1	Treatment 2	Treatment 3	Treatment 4	Treatment 5
Treatment 2	0.12 (0.83)				
Treatment 3	0.04** (0.24)	**0.02**** **(0.00***)**			
Treatment 4	**0.00*** **	**0.00*** **	**0.00*** **		

	(0.00*)**	**(0.01***)**	**(0.00***)**		
Treatment 5	0.00*** (1.00)	**0.00*** (0.00)*****	**0.00*** (0.00***)**	**0.00*** (0.00***)**	
Treatment 6	0.92 (0.66)	0.95 (0.14)	0.74 (0.14)	0.00*** (0.64)	**0.05** (0.00***)**

Numbers in parentheses are the P Values based on continuity-corrected Pearson $X^2(1)$ test of equal medians. Refer to Table 6 for median and mean values. ***, **, * indicate significant at the 1%, 5% and 10% levels. Cases of agreement between equal median and Mann-Whitney rank-sum tests are shown in bold.

We see in Table 7 that treatments are important for fuelwood savings and most treatment combinations result in fuelwood savings that are statistically distinct. In general, though not always, rank-sum Mann-Whitney and equal median test results correspond. These cases are shown in bold. We see that treatment 5 (no-cost of Mirt stove and no network support) stands out as giving very different fuelwood savings results than the other treatments. Except for treatment 1, which has statistically similar median fuelwood savings, we estimate that treatment 5 is superior to all other treatments in terms of fuelwood savings.

Similarly, treatment 4 (pay ETB 25, plus network support) is dominated by all other treatments except for the equality of medians test vis-à-vis treatment 6, which has statistically equivalent median fuelwood savings. These preliminary results suggest that households with treatments that include networking generally have lower levels of fuelwood savings. An exception is in the comparison of treatments 1 and 2, which are statistically equivalent based on both tests. Other generalizable inferences are difficult to draw.

4.3 Propensity Score Matching and the Effect of Treatments on Fuelwood Savings

We use propensity score matching as an important robustness check on the results presented in section 4.2. Propensity score matching utilizes observables to estimate probit models of the probability that households receive a particular treatment. The predicted values of these probit models are then used as propensity scores that match treatment and control households. The independent variables included in the probit model are ones that are likely similar for members in a particular community, but differ across villages. These variables are respondent education, distance from the respondent's household to the nearest road and respondents' religions. For some treatments and pairwise treatment comparisons all three variables are included, but in some models with all three variables included treatment and control groups were not balanced. Fewer than the three independent variables were therefore used to estimate propensity scores, but distance to the nearest road (an important continuous variable and measure of isolation) is included in all models.

We begin by focusing on each treatment in turn, with all others as "controls." This is followed by analyses of all possible treatment combinations as was presented in Table 7. Also included in each cell are (i) the number of actual matched treatment and control households; (ii) the two-tailed p-value based on the estimated t-statistic and (iii) whether the average treatment effect is significant at the 1%, 5% and 10% levels. Additional information on the number of blocks, region of common support and other information for matched treated and control households are available from the authors.

Average treatment effects are estimated after neighbor propensity score matching with standard errors bootstrapped (500 repetitions). In the models where each treatment is compared to all others, all treated observations were matched. This finding is not surprising given our randomized treatment assignment, which helps assure that treated and control observations are comparable. For the models where individual treatments are compared (e.g. treatment 6 versus treatment 1), most, but not all treated observations are matched.

In all cases matching is restricted to the region of common support, which assures that truly comparable observations are matched, but also introduces a constraint if observations are very different. As shown in Table 8, the quality of treatment and control matches is very high. Though the region of propensity score common support was generally 0.15 to 0.45, after matching, as shown below, the average absolute difference in propensity scores was extremely small and for some treatments virtually zero.

As shown in Table 8, households with treatments 1 and 6 glean about 3.15 to 3.50 fewer kilograms of fuelwood savings per week from their Mirt stoves than those with other treatments. Those with treatment 4 are estimated to have almost 10 kilograms fewer wood savings than those who received other treatments. Treatment 5 dominates other treatments with over 5 kg more fuelwood savings than households receiving other treatments. Based on these findings we estimate that if all households had received treatment 5, overall fuelwood savings would be at least 15 kilograms per Mirt stove per week (806 kg per year) rather than the 12.20 per week actually observed.

Table 8: Average Effect on Fuelwood Savings of Treatments vis-à-vis all other Treatments (kilograms per week) Using Nearest Neighbor Propensity Score Matching

	ATT (p value)	Total Treated (Total Matched)	Average Absolute Propensity Score Difference (treated vs. control)	Ave. Outcome of Matched Treated	Average Outcome of Matched Untreated
Treatment 1	-3.50** (0.02)	209 (209)	0.003	7.86	11.35
Treatment 2	0.45 (0.74)	247 (247)	0.0001	10.86	10.41
Treatment 3	-1.25 (0.35)	247 (247)	0.00001	10.83	11.34
Treatment 4	-9.97*** (0.00)	213 (213)	0.0011	2.20	12.16
Treatment 5	5.276** (0.02)	243 (243)	0.003	14.00	8.73
Treatment 6	-3.151 (0.04)**	218 (218)	0.0003	7.731	10.88

***, **, * indicate significant at the 1%, 5% and 10% levels. Matching only over region of common support.

Table 9 Average Treatment Effects of Row Treatment vis-à-vis Column Control Using Nearest Neighbor Propensity Score Matching (Kilograms Fuelwood Saved per Week)

	Treatment1	Treatment 2	Treatment 3	Treatment 4	Treatment 5
Treatment 2	4.745** (0.04) [215/212]				
Treatment 3	4.778** (0.04) [211/209]	-0.243 (0.93) [211/205]			
Treatment 4	-6.418*** (0.00) [177/166]	-10.431*** (0.00) [177/163]	-6.824*** (0.00) [177/161]		
Treatment 5	8.102*** (0.00) [207/207]	6.015*** (0.005) [207/191]	7.063*** (0.001) [207/182]	15.528*** (0.00) [207/203]	
Treatment 6	-0.767 (0.74) [182/172]	-0.365 (0.86) [175/182]	-2.234 (0.37) [175/182]	3.077** (0.045) [182/139]	-1.150 (0.65) [182/167]

ATT in bold. Positive numbers indicate the row treatment has a larger effect on fuelwood savings per week than the column treatment. P values in parentheses. Number of matched treated/number of control in square brackets. ***, **, * indicate significant at the 1%, 5% and 10% levels. Matching only over region of common support.

Table 9 presents pairwise average effects (in terms of fuelwood saving per weeks) of row treatments vis-à-vis column treatments. Results are very similar to those of the Mann-Whitney and median equality tests in Table 7. We see that using propensity score nearest neighbor matching treatment 5 dominates all other treatments, except treatment 6 (no statistical difference), with additional savings in the 6 to 15 kilogram range. Treatment 4 is found to be the least effective, with statistically fewer fuelwood savings compared with all other treatments. Treatment 4 is estimated to elicit 3 to 15 kilograms less average fuelwood savings depending on the treatment comparator.

The pairwise comparisons of treatments 2, 4 and 6, which all have the network treatment aspect, and 1, 3, 5, which do not have networking, are especially instructive. For those with the networking aspect we find that <u>those who paid for the stove rather than receiving for free or receiving an incentive have statistically lower fuelwood savings</u>, but free distribution and receiving a small incentive for use during the first monitoring period are statistically indistinguishable. Comparisons of treatments 1, 3, 5 are somewhat less clear. We find that those who received their Mirt stoves for free have much larger fuelwood savings than the other two monetary treatment aspects. Those who paid 25 birr for their stoves (treatment 3) are estimated, however, to save 4.8 kilograms more per week than those who received the first-period use incentive. This finding corresponds with those from the Mann-Whitney tests in Table 7.

We are unable to draw robust conclusions about the networking treatment aspect independent of the financial terms. Comparing treatments 1 and 2 (small incentive payment in period 1) we find that households receiving the network treatment saved an average of 4.7 kilograms more wood per week. Those who received treatment 4 (paid ETB 25 with networking) had fewer fuelwood savings than all other treatments, including treatment 3 (Paid ETB 25, no networking). Finally, the effects of treatments 5 and 6 (free distribution with and without networking) are found to be statistically indistinguishable.

4.4 Pooled OLS Regression Results[9]

We now discuss pooled OLS results with district (i.e. *woreda*) fixed effects and errors clustered by household. The dependent variable is quantity of fuelwood saved due to use of Mirt stove in kilograms per week. We also include interactions between treatment and round dummies where the interaction between treatment 5 and round 1 is the omitted category. We estimate five models, which differ in terms of the number of explanatory variables included in the analysis. Under the null hypothesis that $R^2 = 0$, the F statistic follows the F distribution with 58 and 749 degrees of freedom. We find for Model 5, which is the most comprehensive in terms of explanatory variables

[9]The random effects estimation results with district (i.e. *woreda*) fixed effects are found in Appendix A.

included, that we can reject the null hypothesis that the explanatory variables have no impact on the quantity of fuelwood saved from using the Mirt stove at less than 1% level (F=12.04).[10] We therefore focus our discussion on Model 5. Our interaction terms between round and treatment are our variables of primary interest. Marginal effects of treatments and round together with results of hypothesis tests are presented in Appendix B.

Table 10. Pooled OLS regression results of the determinants of fuel savings

	Model_1	Model_2	Model_3	Model_4	Model_5
treat1_round1	29.079**	28.624**	28.165**	35.089**	46.609***
	(13.142)	(12.983)	(12.933)	(14.952)	(15.320)
treat2_round1	36.596***	37.275***	37.159***	45.580***	69.348***
	(13.134)	(13.070)	(13.073)	(14.874)	(16.789)
treat3_round1	32.098**	31.294**	31.194**	37.109***	45.501***
	(13.133)	(12.903)	(12.824)	(13.199)	(14.771)
treat4_round1	39.992***	38.680***	38.446***	42.597***	43.153***
	(14.191)	(13.918)	(13.834)	(14.716)	(12.615)
treat6_round1	17.609	18.734	18.713	28.467*	36.166**
	(13.876)	(13.906)	(14.008)	(16.630)	(16.559)
treat1_round2	37.499***	37.808***	37.522***	45.345***	55.741***
	(13.364)	(13.196)	(13.124)	(15.118)	(15.481)
treat2_round2	40.183***	41.398***	41.280***	49.911***	74.065***
	(13.157)	(13.145)	(13.141)	(14.944)	(16.874)
treat3_round2	37.183***	36.748***	36.765***	41.450***	47.709***
	(13.257)	(13.219)	(13.190)	(13.906)	(15.850)
treat4_round2	31.741**	30.325**	30.092**	33.262**	33.120***
	(13.090)	(12.726)	(12.703)	(13.451)	(11.833)
treat5_round2	42.028***	41.786***	41.828***	44.554***	44.429***
	(13.793)	(13.846)	(13.868)	(14.687)	(14.807)
treat6_round2	28.046**	28.336**	29.821**	40.695**	51.787***
	(14.166)	(14.272)	(14.273)	(16.657)	(16.972)
treat1_round3	42.365***	42.882***	42.440***	50.410***	61.407***
	(13.185)	(12.910)	(12.831)	(14.889)	(15.476)
treat2_round3	43.720***	44.219***	44.152***	52.862***	75.275***
	(13.395)	(13.445)	(13.467)	(15.080)	(16.591)
treat3_round3	39.693***	38.370***	38.379***	42.746***	48.494***
	(13.387)	(13.286)	(13.277)	(13.725)	(15.782)
treat4_round3	32.293**	31.120**	30.983**	34.958**	36.006***
	(13.085)	(12.711)	(12.694)	(13.809)	(12.317)
treat5_round3	48.275***	48.920***	48.781***	50.522***	49.918***
	(13.958)	(14.221)	(14.151)	(14.480)	(14.353)
treat6_round3	49.972***	51.378***	50.489***	62.415***	76.473***
	(13.738)	(13.921)	(13.985)	(16.464)	(17.470)
treat1_round4	40.991***	41.455***	41.257***	48.425***	59.559***
	(13.153)	(12.903)	(12.825)	(14.782)	(15.272)
treat2_round4	40.621***	40.736***	40.732***	49.306***	71.626***

[10] $F = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$ where *k* is the number of parameters and *n* is the number of observations.

	(13.184)	(13.130)	(13.187)	(14.812)	(16.468)
treat3_round4	39.790***	39.923***	39.862***	44.774***	51.440***
	(13.212)	(13.158)	(13.125)	(13.861)	(15.883)
treat4_round4	32.235**	30.928**	30.764**	34.240**	34.892***
	(13.084)	(12.691)	(12.676)	(13.524)	(11.953)
treat5_round4	47.875***	48.539***	48.382***	50.344***	49.817***
	(14.058)	(14.300)	(14.233)	(14.553)	(14.393)
treat6_round4	45.796***	45.992***	44.833***	57.601***	72.253***
	(13.832)	(13.800)	(13.888)	(16.446)	(17.634)
age_resp		-0.047	-0.054	-0.079	-0.017
		(0.050)	(0.050)	(0.055)	(0.048)
sex_res		-6.525**	-7.069**	-7.951***	-9.717***
		(2.697)	(2.868)	(2.840)	(2.968)
Education		1.810	1.475	-0.318	-2.208
		(1.782)	(1.767)	(1.704)	(1.510)
Marital		4.319*	4.222*	5.885**	5.056**
		(2.335)	(2.344)	(2.460)	(2.410)
Ade_TOT_HH		0.637	0.713	-0.034	-0.095
		(0.669)	(0.653)	(0.719)	(0.688)
religion_dummy		5.309**	5.279**	-0.478	-1.460
		(2.348)	(2.327)	(1.847)	(2.917)
Livestock_Tot_HH		0.453	0.444	0.169	0.488
		(0.296)	(0.296)	(0.302)	(0.331)
Children_15		-0.139	-0.200	0.055	-0.234
		(0.595)	(0.618)	(0.608)	(0.558)
Dist_road		0.018**	0.019**	0.016*	0.025**
		(0.008)	(0.008)	(0.009)	(0.012)
particpation			-2.159	-2.680	-2.417
			(2.150)	(2.006)	(1.805)
av_equivalent			0.271	0.329	0.689
			(0.785)	(0.780)	(0.658)
Average_temp				-0.652	-2.176*
				(0.509)	(1.248)
Injera_no				0.347***	-0.052
				(0.117)	(0.103)
Stove_use				4.887*	1.972
				(2.831)	(2.605)
Qunatity_flour				0.775	0.615
				(0.475)	(0.405)
Flour_kind2				0.894	-2.833
				(2.631)	(2.603)
Stove_install				0.886	2.250
				(1.778)	(1.762)
Main_fuel				-7.118***	3.681
				(2.127)	(3.105)
Constant	-30.597**	-37.404***	-36.830***	-24.592**	4.142
	(13.079)	(13.421)	(13.037)	(9.948)	(26.065)
N	1139	1115	1103	1079	1079

Note: *** = Significant at 1%, ** = Significant at 5%, * = Significant at 10%, District dummies are included in model 5 but are not reported for the sake of economizing space.

The results for the six sets of hypotheses presented in Appendix B are briefly discussed below. The first set of hypotheses focus on the fuel saving implications of paying for the stove compared to getting it for free. A first test of this is to examine equality of the marginal effect of treatment 3, which involves payment of 25 Birr by the household (T3), with treatment 5 where the household receives the improved stove for free (T5). These two were not found to be significantly different. On the other hand, the marginal effect of treatment 4 (T4), which involves payment of 25 Birr combined with social network, were found to be significantly lower than those of treatment 6 (T6), where the household receives the improved stove for free and there is networking. What this suggests is that fuel savings is in fact higher for households that received the stove for free together with social networking compared with those who paid 25 Birr together with social network.

The second set of hypotheses relate to the effect of paying 25 Birr for the stove or receiving 50 Birr after using the stove. As the results show in Appendix B, there is no statistically significant difference between receiving money for stove use (T1) and paying for the stove (T3). However, when combined with social networking, receiving money for stove use (T2) increases fuel savings compared with paying for the stove (T4).

The third set of hypotheses we test is about the overall effect of treatments that involve monetary cost or reward on fuel savings. Similar to other tests above we test the equality of T1 (received 50 Birr for stove use), T3 (paid 25 Birr for the stove) and T5 (received the stove for free) as none of these involve networking. The results suggest that these three treatments are significantly different from each other. These three treatments are also significantly different from each other when each of them is combined with networking (i.e. a test of equality of T2, T4 and T6). This finding suggests the importance of incentives to receive and use the stove.

The fourth set of hypotheses tests focus on the role of networking. We test the equality of three pairs of treatments, with each pair being the same except that one of the two treatments involves networking, while the other does not (T1=T2; T3=T4; and T5=T6). The results show statistical differences for each of the three tests. When comparing the marginal effects, we find that networking increases fuel savings when it is combined with receiving 50 Birr for stove use or receiving the stove for free. However, networking decreases fuel savings when households pay Birr 25 for the stove.

The fifth set of hypotheses we test is about the effect of treatments over time. The results show that each treatment is statistically different when compared over the four rounds of data collection. Looking at the marginal effects of each treatment over time from the pooled regression in Table 10, we note that while there is a general tendency for an increase in fuel savings over time, in some cases the increase is either not much in the last round or there is a decrease, while in other cases there appears to be a fluctuation in the effect of treatments over time.

The sixth set of hypotheses we test is the effect of time on fuel savings. Tests for differences in marginal effects across rounds show that they are statistically different. The magnitudes show an increase over time, except for the last round where there is some decline. This suggests that Mirt stove use requires learning and fuelwood savings results do not come immediately.

Few covariates are not found to be significant determinants of fuelwood savings, though gender, marital status, distance to road and average temperature are significant at least at the 10% level. We note that it is not surprising that few covariates are significant, because of our randomized treatment trial approach. Some of the district/woreda dummies are significant, suggesting variation in fuel savings across sites. This is also not surprising as there are differences in culture, fuel availability and agro-climatic conditions.

4.5 Emissions Reduction in $CO_{2}e$

We follow the method described in section 3 based on Lee et al. (2013) to calculate greenhouse gas emissions reductions from using Mirt instead of the traditional three-stone stove. As discussed above, we estimate that the Mirt stove on average saves 634 kilograms of fuelwood. To convert these into carbon emissions reductions we use a net calorific value of 15 MJ/kg of woody biomass (Hall et al. 1994) and 112 g of CO_2 per MJ of fuelwood (UNFCCC, 2012). Considering the total woody biomass as non-renewable, 1.065 tons of CO_2 would be saved per Mirt stove per year, which is about half of what Dresen et al. (2014) estimate.

If we instead use an 88% non-renewable biomass percentage, which is the CDM default value for Ethiopia (UNFCCC, 2012), we estimate 0.94 tons of CO_2 saved per household per year. The CO_2e corresponding to 9.4 million improved cookstoves, which is the official distribution goal of the Ethiopian government, assuming that all ICS distributed perform like the Mirt stove, would then be about 8.8 million tons of CO_2e per year. Using the May 2015 California auction price of US$13.93 per ton of CO_2,[11] these CO_2e savings would be worth about US$123 million per year.

[11] See http://www.arb.ca.gov/cc/capandtrade/auction/may-2015/summary_results_report.pdf downloaded May 29, 2015.

5. Implications of Findings and Conclusions

This paper uses a randomized experimental design, real-time electronic stove temperature measurements and controlled cooking tests to estimate the fuelwood and CO_2 savings from a Mirt cookstove program in the process of being implemented in Ethiopia. Given the current objective of the Ethiopian government to disseminate such stoves widely, understanding whether, how much and why improved cookstoves are used and how much wood they save is important. Knowing more about how households interact with improved cookstoves is also important, because cooking – and especially *injera* cooking – uses a majority of the fuelwood in the country and therefore is a critical determinant of greenhouse gas emissions and indoor air pollution reductions. Confirming stove use and fuelwood savings is, therefore, a critical aspect of crediting ICS climate change benefits under REDD+ or other offset programs.

Our findings suggest that if put into the field at the relatively low market price of approximately $12 per stove, people will use the stoves and over time save a substantial amount of fuelwood. We estimate that on average one Mirt stove saves approximately 634 kilograms of fuelwood per year, which is about half of previous estimates. We find that fuelwood saved per week per household varies by region and round and the first round had the lowest average fuelwood savings, suggesting household learning costs. In all rounds Amhara has significantly larger fuelwood savings than the other two regional states, Oromiya and SNNP. These three regions make up about 80% of Ethiopia's population. We estimate that the CO_2e reduction corresponding to the 9.4 million improved cookstove distribution goal announced by the federal government of Ethiopia would be about 8.8 million tons of CO_2e per year not counting any black carbon reductions, which could be worth about US$123 million per year. Thus, we find that the Mirt stove could have a significant effect on reducing carbon emissions and if buyers could be found, could generate significant carbon offset revenues.

On the methodological side, we note that our estimates combine a pure field experiment focusing on stove use and duration with a more controlled field-based lab experiment that precisely estimates per-minute fuelwood and cooking time. We would like to suggest that this approach offers advantages over either the survey-based analysis – which tends to overstate both stove use and fuelwood savings – and kitchen performance tests, which have their own set of disadvantages.

The study contributes to the literature on behavioral treatments in randomized trials and particularly the literature on payment terms and network effects. Creating local Mirt stove networks generally appears to increase fuelwood savings, though in a number of models these effects depend on payment terms. Monetary costs and rewards to households also appear to have effects on fuel savings, with the most robust positive effects coming from free distribution.

Though we estimate that fuelwood savings would have been greater if all stoves were distributed for free rather than using our other monetary treatments, the existing imperfect distribution generates significant carbon value. Using the May 2015 California CO_2e auction price of \$13.39/ ton, each stove saves 0.94 tons per year, which is worth \$12.59. Thus, we estimate that using only carbon market offset revenues, the Mirt stove covers costs within one year.

References

Accenture Development Partnerships, *Ethiopia Market Assessment Report: Enhancing Markets for Delivery of Improved Cookstove Development and Promotion Support in Ethiopia; Market Analysis, Recommendations and Program Plan*, Final report, developed for the Ethiopian Federal Ministry of Water and Energy in partnership with the BARR Foundation and with the support of the Global Alliance for Clean Cookstoves, Accenture (UK) Limited.

Adkins, E., Tyler, E., Wang, J., Siriri, D., Modi, V. 2010. "Field testing and survey evaluation of household biomass cookstoves in rural sub-Saharan Africa. *"Energy for Sustainable Development* 14, 172 - 185.

Allcott, H. and T. Rogers. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence form Energy Conservation." *American Economic Review* 104: 3003-3037.

Andersson, C., A. Mekonnen, J. Stage. 2011. "Impacts of the Productive Safety Net Program in Ethiopia on Livestock and Tree Holdings of Rural Households." *Journal of Development Economics* 94: 119-126.

Bailis, R, R. Drigo, A. Ghilardi, O. Masera. 2015. "The Carbon Footprint of Traditional Woodfuels," *Nature Climate Change*, Vol. 5: 266- 272.

Bensch, G. and J. Peters. 2013. "Alleviating Deforestation Pressures? Impacts of Improved Stove Dissemination on Charcoal Consumption in Urban Senegal." *Land Economics*, Vol. 89: 676-698.

BenYishay, A. and A.M. Mobarak. 2014. "Social learning and communication." NBER working paper 20139.

Beyene, Abebe, Randy Bluffstone, Zenebe Gebreegziabher, Peter Martinsson, Alemu Mekonnen, Ferdinand Vieider, 2015. "The Improved Biomass Stove Saves Wood, But How Often Do People Use It?" Evidence from a Randomized Treatment Trial in Ethiopia, draft.

Beyene, A, R. Bluffstone and A. Mekonnen. "Community Forests, Carbon Sequestration and REDD+: Evidence from Ethiopia," Forthcoming *Environment and Development Economics*

Bizzarri, M. 2010. *Safe Access to Firewood and Alternative Energy in Ethiopia: An Appraisal Report.* Prepared for the World Food Program. [http://www.genderconsult.org/uploads/publications/doc/SAFE_Ethiopia_Appraisal_Report_Final_Draft_2.pdf Accessed May 12](http://www.genderconsult.org/uploads/publications/doc/SAFE_Ethiopia_Appraisal_Report_Final_Draft_2.pdf), 2015.

Bluffstone, R. 1998."Reducing Degradation of Forests in Poor Countries When Permanent Solutions Elude Us: What Instruments Do We Really Have?" *Environment and Development Economics*, 3 (3): 295-318. 1998.

Brouwer, R. & Falcão, M.P. 2004. "Wood Fuel Consumption in Maputo, Mozambique. *Biomass & Bioenergy* 27: 233-245.

Burwen, J. and Levine, D. I. 2012. A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *Energy for Sustainable Development* 16(3):328-338.

Charness, G. and U. Gneezy. 2009. Incentives to Exercise. *Econometrica*. 77: 909-931

Cohen, J., and P. Dupas. 2010. "Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment." *Quarterly Journal of Economics*. 124: 1-45.

Dresen, Elisabeth, Ben DeVries, Martin Herold, Louis Verchot and Robert Müller, 2014. "Fuelwood Savings and Carbon Emission Reductions by the Use of Improved Cooking Stoves in an Afromontane Forest, Ethiopia," *Land* , *3*, 1137-1157; doi:10.3390/land3031137

Duflo, E., R. Glennerster, and M. Kremer. 2008. *Using Randomization in Development Economics Research: A Toolkit, in: Handbook of Development Economics*, edited by Paul Schultz and John Strauss, 3895-3962. North Holland.

FDRE, 2011. *Ethiopia's Climate-Resilient Green Economy*: *Green Economy Strategy*.

FDRE 2014. *Ethiopia's Climate-Resilient Green Economy Climate Resilience Strategy: Water and Energy*, Draft, Ministry of Water, Irrigation and Energy, Federal Democratic Republic of Ethiopia (FDRE).

Heltberg, R. 2004. "Fuel Switching: Evidence from Eight Developing Countries." *Energy Economics* Vol. 26 (5):869-887.

Gebreegziabher, Z, A. Beyene, M., A. Mekonnen, Toman, R. Bluffstone, S. Dissanayake, P. Martinsson. 2015. "Can Improved Biomass Cookstoves Contribute to REDD+ in Low-Income Countries? Results from a Randomized Control Trial in Ethiopia." Mimeo available from the first author.

Gordon, H.S. 1954. "The Economic Theory of a Common-Property Resource: The Fishery," *Journal of Political Economy* 62(2): 124–42.

Hall, D.O., Rosillo-Calle, F, Woods, J. 1994. Biomass Utilization in Households and Industry—Energy Use and Development. *Chemosphere 29:* 1099–1119.

IEA 2012. *World Energy Outlook 2010*. International Energy Agency. Paris: IEA.

Jeuland, M. A, Pattanayak, S. and Bluffstone, R. Forthcoming. "The Economics of Household Indoor Air Pollution," *Annual Review of Environmental and Resource Economics*.

Jeuland, M. A. and S. Pattanayak. 2012. "Benefits and Costs of Improved Cookstoves: Assessing the Implications of Variability in Health, Forest and Climate Impacts." *PloS one,* 7**,** e30338.

Johnson, Michael, Rufus Edwards, Adrian Ghilardi, Victor Berrueta Dangillen, Claudio Alatorrefrenk, and Omar Masera, 2009. "Quantification of Carbon Savings from Improved Biomass Cookstove Projects," *Environmental Science & Technology*, 43, 2456–2462.

Jones, D. 1988. "Some Simple Economics of Improved Cookstove Programs in Developing Countries." *Resources and Energy*, 10: 247-264.

Lee, Carrie M., Chelsea Chandler, Michael Lazarus and Francis X. Johnson, 2013. "Assessing the Climate Impacts of Cookstove Projects: Issues in Emissions Accounting," SEI Working paper 2013-01.

Lim, S, T. Vos, A. D. Flaxman., G. Danaei, K. Shibuya, H. Adair-Rohani, . A. Almazroa, M. Amann, H. R. Anderson, and K. G. Andrews. 2013. "A Comparative Risk Assessment of Burden of Disease and Injury Attributable to 67 Risk Factors and Risk Factor Clusters in 21 Regions, 1990–2010: a Systematic Analysis for the Global Burden of Disease Study 2010." *The Lancet,* 380**,** 2224-2260.

Masera, O., Edwards, R., Arnez, C.A., Berrueta, V., Johnson, M., Bracho, L.R., Riojas-Rodriguez, H., Smith, K.R. 2007. Impact of Patsari improved cookstoves on indoor air quality in Michoac·n, Mexico. *Energy for Sustainable Development*, Vol. 11:45 - 56.

Mekonnen, A, and R. Bluffstone, "Increasing forest cover in Ethiopia: Lessons from Economics and International Experience." In *Policies to Increase Forest Cover in Ethiopia*, Proceedings of a Policy Workshop Organized by the Ethiopian Development Research Institute 18 – 19 September 2007, Addis Ababa.

Miller, G. and A.M. Mobarak. 2014. "Learning about new technologies through social networks: Experimental evidence from nontraditional stoves in Banglaesh." *Marketing Science* 1-20.

Munshi, K. 2004. "Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution." *Journal of Development Economics*, 73, 185-215.

Mobarak, Ahmed Mushfiq, Punee Dwivedi, Robert Bailis, Lynn Hildemann, and Grant Miller, 2012. "Low Demand for Nontraditional Cookstove Technologies," *Proceedings of the National Academies of Sciences*. 109: 10815-10820.

Nepal, M, A. Nepal and K. Grimsrud, 2010. "Unbelievable but improved cookstoves are not helpful in reducing firewood demand in Nepal, *Environment and Development Economics*, Vol. 16: 1-23.

O'Haver, Tom, 1997. "A Pragmatic Introduction to Signal Processing with applications in scientific measurement: An illustrated essay," Department of Chemistry and Biochemistry, University of Maryland at College Park; http://terpconnect.umd.edu/~toh/spectrum/IntroToSignalProcessing.pdf

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

Pattanayak, S., Jeuland, M., Lewis, J., Bhojvaid, V., Brooks, N., Kar, A., Morrison, L., Patange, O., Philippone, L., Ramanathan, N., Rehman, I., Thadani, R., Vora, M. & Ramanathan, V. 2014. "Cooking up change in the Himalayas: Evidence from Mixing Quasi-Experiments with an Experiment on Cookstove Promotion." *Duke University Working Paper.* Durham, USA.

Practical Action Ethiopia. *Healthy Hoods Ethiopia Fact Sheet*, https://practicalaction.org/healthy-smoke-hoods-tool/country_data/ethiopia.pdf Accessed May 12, 2015.

Smith, K.R., K. Dutta, C. Chengappa, P. P. S. Gusain, O. Masera, V. Berrueta, R. Edwards, R. Bailis, K. N. Shields. 2007. "Monitoring and evaluation of improved biomass cookstove programs for indoor air quality and stove performance: conclusions from the Household Energy and Health Project," *Energy for Sustainable Development.* Vol. 11: 5 -18.

Shankar, A., M. Johnson, E. Kay, R. Pannu, T. Beltramo, E. Derby, S. Harrell, C. Davis, H. Petach. "Maximizing the Benefits of Improved Cookstoves: Moving from Acquisition to Correct and Consistent Use." *Global Health: Science and Practice* 2: 268-274.

Thakuri, M. B. M. 2009. "Revising the Need of Improved Stoves: Estimating Health, Time and Carbon Benefits," SANDEE Working Paper No. 44-09. Kathmandu.

Thomas, E., C. Barstow, G. Rosa, F. Majorin and T. Clasen. Forthcoming. "Use of Remotely Reporting Electronic Sensors for Assessing Use of Water Filters and Cookstoves in Rwanda." *Environmental Science & Technology*, www.doi.org/10.1021/es403412x

UNFCCC, 2012. "Default Values of Fraction of Non-Renewable Biomass for Least DevelopedCountries and Small Island Developing States." 67th Meeting Report, Annex 22. United Nations Framework Convention on Climate Change, Executive Board of the Clean Development Mechanism, Bonn. http://cdm.unfccc.int/DNA/fNRB/index.html.

Wolfram, C., Shelef, O., and Gertler, P., 2012. "How will Energy Demand Develop in the Developing World? *Journal of Economic Perspectives* 26(1): 119-138.

Appendix A. Random Effects Model of Fuelwood Savings

	Model_1	Model_2	Model_3	Model_4	Model_5
treat1_round1	28.975**	28.562**	28.083**	35.089**	46.609***
	(13.146)	(12.978)	(12.925)	(14.956)	(15.320)
treat2_round1	36.533***	37.242***	37.126***	45.608***	69.348***
	(13.137)	(13.079)	(13.086)	(14.884)	(16.789)
treat3_round1	32.896**	31.855**	31.879**	37.364***	45.501***
	(13.133)	(12.921)	(12.852)	(13.216)	(14.771)
treat4_round1	39.910***	38.629***	38.392***	42.588***	43.153***
	(14.194)	(13.915)	(13.831)	(14.717)	(12.615)
treat6_round1	17.178	18.480	18.412	28.409*	36.166**
	(13.887)	(13.920)	(14.024)	(16.642)	(16.559)
treat1_round2	37.448***	37.764***	37.465***	45.325***	55.741***
	(13.363)	(13.187)	(13.115)	(15.118)	(15.481)
treat2_round2	40.192***	41.407***	41.298***	49.953***	74.065***
	(13.159)	(13.151)	(13.151)	(14.954)	(16.874)
treat3_round2	37.027***	36.700***	36.703***	41.440***	47.709***
	(13.257)	(13.225)	(13.199)	(13.911)	(15.850)
treat4_round2	31.628**	30.266**	30.027**	33.247**	33.120***
	(13.093)	(12.725)	(12.702)	(13.451)	(11.833)
treat5_round2	42.112***	41.845***	41.897***	44.576***	44.429***
	(13.821)	(13.865)	(13.891)	(14.695)	(14.807)
treat6_round2	27.953**	28.343**	29.846**	40.733**	51.787***
	(14.179)	(14.292)	(14.293)	(16.674)	(16.972)
treat1_round3	42.281***	42.820***	42.370***	50.401***	61.407***
	(13.188)	(12.908)	(12.830)	(14.896)	(15.476)
treat2_round3	43.655***	44.193***	44.125***	52.877***	75.275***
	(13.398)	(13.451)	(13.475)	(15.087)	(16.591)
treat3_round3	39.596***	38.356***	38.362***	42.745***	48.494***
	(13.390)	(13.297)	(13.291)	(13.729)	(15.782)
treat4_round3	32.272**	31.096**	30.955**	34.959**	36.006***
	(13.088)	(12.710)	(12.692)	(13.810)	(12.317)
treat5_round3	48.032***	48.753***	48.583***	50.479***	49.918***
	(13.938)	(14.204)	(14.138)	(14.482)	(14.353)
treat6_round3	50.146***	51.508***	50.653***	62.489***	76.473***
	(13.749)	(13.937)	(14.003)	(16.480)	(17.470)
treat1_round4	40.931***	41.401***	41.192***	48.417***	59.559***

	(13.156)	(12.897)	(12.817)	(14.785)	(15.272)
treat2_round4	40.607***	40.741***	40.738***	49.338***	71.626***
	(13.187)	(13.140)	(13.198)	(14.821)	(16.468)
treat3_round4	39.758***	39.941***	39.887***	44.792***	51.440***
	(13.215)	(13.167)	(13.138)	(13.865)	(15.883)
treat4_round4	32.231**	30.928**	30.767**	34.253**	34.892***
	(13.089)	(12.694)	(12.678)	(13.526)	(11.953)
treat5_round4	47.635***	48.370***	48.184***	50.292***	49.817***
	(14.031)	(14.281)	(14.216)	(14.554)	(14.393)
treat6_round4	45.814***	46.110***	44.964***	57.643***	72.253***
	(13.849)	(13.834)	(13.926)	(16.469)	(17.634)
age_resp		-0.048	-0.056	-0.080	-0.017
		(0.051)	(0.050)	(0.055)	(0.048)
sex_res		-6.447**	-6.969**	-7.908***	-9.717***
		(2.709)	(2.884)	(2.845)	(2.968)
Education		1.861	1.538	-0.303	-2.208
		(1.796)	(1.787)	(1.710)	(1.510)
Marital		4.246*	4.128*	5.838**	5.056**
		(2.344)	(2.355)	(2.466)	(2.410)
Ade_TOT_HH		0.624	0.697	-0.040	-0.095
		(0.675)	(0.661)	(0.723)	(0.688)
religion_dummy		5.268**	5.227**	-0.509	-1.460
		(2.362)	(2.345)	(1.855)	(2.917)
Livestock_Tot_HH		0.441	0.429	0.163	0.488
		(0.299)	(0.300)	(0.304)	(0.331)
Children_15		-0.122	-0.179	0.059	-0.234
		(0.603)	(0.629)	(0.612)	(0.558)
Dist_road		0.019**	0.019**	0.017*	0.025**
		(0.008)	(0.008)	(0.009)	(0.012)
particpation			-2.172	-2.689	-2.417
			(2.172)	(2.013)	(1.805)
av_equivalent			0.264	0.332	0.689
			(0.785)	(0.781)	(0.658)
Average_temp				-0.659	-2.176*
				(0.513)	(1.248)
Injera_no				0.348***	-0.052
				(0.117)	(0.103)
Stove_use				4.925*	1.972
				(2.844)	(2.605)
Qunatity_flour				0.769	0.615
				(0.475)	(0.405)
Flour_kind2				0.882	-2.833
				(2.648)	(2.603)
Stove_install				0.859	2.250
				(1.782)	(1.762)
Main_fuel				-7.123***	3.681
				(2.130)	(3.105)
Constant	-30.507**	-37.233***	-36.584***	-24.316**	4.142
	(13.082)	(13.418)	(13.048)	(9.928)	(26.065)
chi2	166.759	174.628	170.351	232.149	698.278
N	1139	1115	1103	1079	1079

Note: *** = Significant at 1%, ** = Significant at 5%, * = Significant at 10%, District dummies are included in model 5, but are not reported in the interest of brevity.

Appendix B. Marginal effects and summary of hypothesis test results for effects of treatments and period on fuel savings

Marginal effects of treatments (T)	T1=223.3; T2=290.3; T3=193.1; T4=147.2; T5=144.2; T6=236.7		
Marginal effects of period (P)	R1=240.8; R2=306.9; R3=347.6; R4=339.6		
Hypothesis	**Null Hypothesis**	**F-Value**	**P-value**
A. There is no effect of paying for the stove	T3=T5	2.69	0.102
on fuel savings compared to getting it for free	T4=T6	9.11	0.003
B. There is no effect of (dis) incentives for using stove on fuel savings	T1=T3	1.71	0.192
	T2=T4	22.3	0.000
C. An overall test of effect of			
monetary cost/reward on fuel savings	T1=T3=T5	5.58	0.004
	T2=T4=T6	15.16	0.000
D. There is no effect of social network on fuel savings	T1=T2	32.46	0.000
	T3=T4	3.00	0.084
	T5=T6	8.88	0.003
E. There is no change in effect of a treatment over time	T1 in P1=T1 in P2=T1 in P3=T1 in P4	10.2	0.000
	T2 in P1=T2 in P2=T2 in P3=T2 in P4	6.35	0.000
	T3 in P1=T3 in P2=T3 in P3=T3 in P4	3.84	0.005
	T4 in P1=T4 in P2=T4 in P3=T4 in P4	3.34	0.011
	T5 in P1=T5 in P2=T5 in P3=T5 in P4	-	-
	T6 in P1=T6 in P2=T6 in P3=T6 in P4	9.93	0.000
F. There is no effect over time	P1=P2=P3=P4	13.79	0.000

*T=treatment; P=Period or round