WPS7297

Policy Research Working Paper 7297

The Improved Biomass Stove Saves Wood, But How Often Do People Use It?

Evidence from a Randomized Treatment Trial in Ethiopia

Abebe D. Beyene
Randall Bluffstone
Zenebe Gebreegziabher
Peter Martinsson
Alemu Mekonnen
Ferdinand Vieider

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

Policy Research Working Paper 7297

Abstract

This paper uses a randomized experimental design and real-time electronic stove use monitors to evaluate the frequency with which villagers use improved biomass-burning Mirt injera cookstoves in rural Ethiopia. Understanding whether, how much, and why improved cookstoves are used is important, because use of the improved stove is a critical determinant of indoor air pollution reductions, and reduced greenhouse gas emissions due to lower fuelwood consumption. Confirming use is, for example, a critical aspect of crediting improved cookstoves' climate change benefits under the United Nations Reducing Emissions from Deforestation and Forest Degradation Programme. The paper finds that Ethiopian households in the study area do use the Mirt stove on a regular basis, taking into account regional differences in cooking patterns. In general, stove users also use their Mirt stoves more frequently over time. Giving the Mirt stove away for free and supporting community-level user networks are estimated to lead to more use. The study found no evidence, however, that stove recipients use the stoves more if they have to pay for them, a hypothesis that frequently arises in policy arenas and has also been examined in the literature.

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 corresponding author may be contacted at bluffsto@pdx.edu.

The Improved Biomass Stove Saves Wood, But How Often Do People Use It? Evidence from a Randomized Treatment Trial in Ethiopia

Abebe D. Beyene[1]

Randall Bluffstone [4]

Zenebe Gebreegziabher[1,2]

Peter Martinsson [5]

Alemu Mekonnen[3,1]

Ferdinand Vieider[6]

Keywords: improved biomass cookstoves, RCT, fuelwood, indoor air pollution, Ethiopia.

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

1. Environmental Economics Policy Forum for Ethiopia/Ethiopian Development Research Institute, Ethiopia
2. Mekele University, Ethiopia
3. Addis Ababa University, Ethiopia
4. Portland State University, USA
5. University of Gothenburg, Sweden
6. Risk & Development Unit, WZB Berlin Social Science Center, Germany

The Improved Biomass Stove Saves Wood, But How Often Do People Use It? Evidence from a Randomized Treatment Trial in Ethiopia[1]

1. Introduction

Humans have cooked food for several hundred thousand years and while the origin of cooking is debated, evidence may date back almost 1 million years, which is before *Homo Sapiens* (Wrangham et al., 1999). Currently, about three-fifths of the human population regularly uses gas or electricity for their cooking (Smith et al., 2013; IEA, 2012), but the remaining 2.8 billion people often find commercial fuels too expensive or too irregularly supplied to use for cooking and heating. Instead, they rely on solid fuels like coal, fuelwood, dung and charcoal that are combusted inside their homes (Jeuland and Pattanayak, 2012, Grieshop et al., 2011, Smith et al., 2013). Without major policy changes, the global number of people depending on such fuels is projected to remain close to constant at least through 2030 (IEA, 2006).

Biomass fuels are a particularly important class of solid fuels that are often self-collected and are easy to use in traditional and inexpensive stoves. This paper focuses on these cooking technologies and particularly on an improved wood-burning stove that in laboratory tests and in controlled field tests uses less wood. The Federal Government of Ethiopia is promoting fuel-efficient biomass cookstoves and has declared its intention to distribute 9.4 million stoves within five years. Achieving this goal would imply that roughly half the households in Ethiopia would use improved biomass stoves. The so-called MIRT *injera*[2] stove, which is currently being promoted in Ethiopia, and is the focus of this paper, has been found to use approximately 30% less wood to cook a typical meal in the field (Gebreegziabher et al., 2014).

---

[1] The authors are listed in alphabetical order and authorship is shared. They would like to thank the World Bank for financial support of the research through the Knowledge for Change Trust Fund and the Trust Fund for Environmentally and Socially Sustainable Development. We would also like to acknowledge the Berkeley Air Monitoring Group for supplying and advising on the use of the electronic stove use monitors used in this study and to the supervisors and enumerators who conducted the field research. Ermias Dessie provided expert data processing that allowed the data to be analyzed.

[2] *Injera* is the main staple bread in Ethiopia, and represents the end-use for a majority of fuelwood consumed in the country.

The continued dependence on biomass fuels and the technologies in which they are burned are of concern for three main reasons. First, combustion of such fuels emits several indoor air pollutants (Smith et al., 2013). When inhaled, these pollutants are known to cause a variety of diseases, the most important of which are lower respiratory infections and chronic obstructive pulmonary disease. Household air pollution from solid fuels is estimated to have killed 3.5 million people in 2010 (Lim et al., 2013) and constitutes the 4th largest risk factor for the global burden of disease.

Second, biomass burning can contribute to outdoor air pollution and climate change. If biomass (e.g. fuelwood) is not sustainably harvested, as is often the case in developing countries, sequestered carbon is on net released into the atmosphere. Indeed, deforestation and forest degradation account for between 12% and 20% of annual greenhouse gas emissions (Saatchi et al., 2011; van der Werf, 2009).

Indoor burning of biomass contributes about 16% of total outdoor particulate concentrations (Smith et al., 2013) and a potentially important part of these particulates is black carbon. Black carbon is a short-lived greenhouse gas that while it is in the atmosphere is a particularly strong climate forcer. Fossil fuel CO_2 emissions cause 40% of anthropogenic climate change, but black carbon is second at 18% and approximately 25% of the black carbon contribution may be from small biomass stoves, mainly in low-income countries (Bond et al., 2013). Certain types of improved biomass cookstoves have been found to reduce black carbon emissions compared with traditional stoves (Kar et al., 2012).

Finally, collecting fuelwood can be time-consuming for households and in some circumstances the scarcity of fuelwood could have important livelihood implications, including reduced agricultural output due to labor and land diversions, increased labor burden on women and children and reduced consumption (Cooke et al., 2008). In general, though, the literature does not support the hypothesis of major welfare effects from biomass scarcity, particularly via household labor budgets (e.g. Amacher et al., 2004). A key reason for this general finding is likely that household production systems and prices - where they exist - adjust to scarcities in sophisticated ways, mitigating major declines in welfare (Bluffstone, 1995).

Improved biomass cookstoves ultimately are only useful for reducing indoor air pollution, mitigating climate change and improving livelihoods if people substitute them for more traditional cooking methods. Substituting, of course, at minimum implies that improved stoves are used. This paper focuses on the actual use of Mirt stoves over more than one year in 36 sites in rural Ethiopia. The study uses real-time stove use monitoring data derived from a 360-stove randomized trial to evaluate the degree to which – under six treatments –people use the Mirt stove. These results can, as discussed above, have important implications for climate change, human health and livelihoods.

There may also be carbon finance implications. The UN Collaborative Programme on Reducing Emissions from Deforestation and Degradation in Developing (REDD+) can credit reduced fuelwood consumption if it is documented that cooking technologies actually reduce deforestation and forest degradation. About a quarter of stove programs in 2010 indeed received or were planning to receive voluntary carbon market or CDM resources (Gifford, 2010).

Whether, how frequently and why low-income country residents actually use more efficient biomass cookstoves are very much open questions that go to the heart of whether stoves are truly "improved." As discussed in the following section, in the past more efficient biomass cookstoves promoted in low-income countries have often not been accepted or used by cooks. Better understanding whether particular stoves really have the potential to address critical challenges – through use – is therefore of prime importance.

The following section discusses the literature on improved biomass cooking stoves, the potential linkages with climate change and forestry policies and the use of field experiments for evaluating the efficacy of such technologies. Section 3 presents the sampling and experimental designs used to collect the primary data. Part 4 of the paper reviews the variables and empirical methods used and section 5 presents the results. Section 6 concludes the paper and draws key policy implications.

2. Literature Review

REDD+ is a still-evolving program to provide incentives for FCCC non-Annex 1 countries to reduce deforestation and forest degradation. Resources have been assembled to support forest management programs, including the World Bank Forest Carbon Partnership Facility and UN-

REDD, which is a United Nations implemented initiative funded by the Government of Norway. The idea behind REDD+ is to provide financial and technical support to developing countries so they can reduce forest biomass loss, sequester more carbon and improve livelihoods. Reducing biomass loss includes forest restoration, rehabilitation, sustainable management and/or reforestation, which means that countries with low rates of deforestation and degradation or have increasing forest stocks will be able to participate in REDD+; REDD+ is therefore more than just avoided deforestation.

REDD+ credits are available for measurable and verifiable emissions reductions as well as sustainable management of forests, and conservation and enhancement of forest carbon stocks (UN-REDD, 2010). But monitoring and verification of carbon sequestration and avoided forest degradation can be a challenge in developing countries, like Ethiopia. Forests and villages may be remote and in many developing countries local groups and communities make decisions about how forests are used. Under such situations it can be very difficult to measure baseline forest carbon and the increments villagers in communities sequester.

A number of authors have suggested that better management of community forests can support REDD+ objectives (e.g. Agrawal and Angelsen; 2009 and Skutsch and McCall, 2012). An estimated 25% of developing country forests (World Bank, 2009; Economist, 2010) and 15.5% of global forests (RRI, 2014) are under the control of communities. Many of these forests are degraded and have significant upward potential. REDD+ could provide incentives to realize this potential while also supporting rural livelihoods in low and lower-middle income countries. This is, of course, the "win-win" potential of REDD+.

What monitoring and commitment mechanisms can help realize these goals when direct monitoring and verification is difficult? Input-based compliance mechanisms for which GHG impacts can be reliably estimated are one potential approach. Such a strategy can be an attractive way to credit offsets, because it may not involve measuring and monitoring actual carbon sequestration in forests. If it is known, for example, that a particular technology is preferred, used and reduces demand for fuelwood, all that has to be monitored is whether subsidies from carbon finance actually reach households and if the technologies are installed.

Improved biomass cooking stoves, such as the Mirt stoves, are technologies that, among other possible improvements, could reduce fuelwood use and forest degradation. As evaluated in a small case study in Mexico by Johnson et al. (2009), proving such claims can be difficult. They note that generating REDD+ or other carbon credits involves understanding stove emission factors, fuel consumption and the portion of fuelwood used that is unsustainably harvested. Documenting that improved stoves with fewer emissions in areas with unsustainably harvested fuelwood are regularly used is critical to documenting reduced fuelwood use, but is not fully sufficient. Rebound effects could even increase fuelwood consumption if users use two stoves more extensively than one stove. Beltramo and Levine (2010), for example, find that actual use of solar ovens in Senegal was low, but when it was used households continued to use their traditional stoves.

As discussed in Gebreegziabher (2015), the history of improved biomass cookstoves is not exactly one of complete success. Criticisms, particularly on behavioral grounds, in the past were widespread. Manibog (1984) took aim at sloppy implementation and Gil (1987) pointed to stove designs that did not meet user expectations. Others focused on what today would be called rebound effects (e.g. Jones, 1988) and Barnes et al. (1993) noted that a key reason for disappointing results was that fuelwood savings were overestimated. While lab tests suggested savings of about 50%, in the field, such savings were unusual. Goals of 25% fuelwood savings were more reasonable, they said.

The importance of the almost 3 billion people who cook with wood on a regular basis remains, however, and in recent years improved biomass cookstoves have re-emerged as a potential solution to important problems. Of special significance is the Global Alliance for Clean Cookstoves, which was founded in 2010 and seeks to foster adoption of clean cookstoves and fuels by 100 million households by 2020. Improved cookstove programs are also increasingly funded by carbon finance (Lewis and Pattanayak, 2012; Gifford, 2010).[3]

At least with regard to per-meal fuelwood savings, authoritative evidence is emerging that low wood-using biomass cookstove designs have improved. Bensch and Peters (2013) conduct an *ex post* evaluation of the Jambar charcoal stove in Senegal. Using OLS with and without propensity

[3] See www.projectsurya.org for an interesting example of a project relying on carbon finance.

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 (Bensch and Peters, 2012). A randomized controlled cooking test was used by Burwen and Levine (2012) to evaluate a locally made and designed woodburning 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 the stoves are used very frequently. Gebreegziabher et al. (2015) find that in field controlled cooking tests Mirt stove wood savings per meal averaged 26% across two rounds of tests and users were satisfied with the technology.

Such field experiments can be very helpful for distinguishing short and long-term effects of interventions.[4] Long-run effects are typically the most important from a welfare perspective (e.g. Alcott and Rogers, 2014), but this depends on the context (e.g. Charness and Gneezy 2009) and the results of interventions often hinge on the terms of those interventions. Dupas (2012), for example, shows that a one-time subsidy can boost the purchase of mosquito-repellent bed nets, but Cohen and Dupas (2010) did not find a significant difference in actual use of the bed nets between households who paid compared with those who received them for free.

Ashraf et al. (2010) find in an experiment on the use of chlorine in drinking water that higher prices indeed *increase* the probability that people use chlorine. Charness and Gneezy (2009) paid students (i.e. a negative price) to exercise at nearby gyms. In the treatment where students were paid to go eight times over a period of one month, they find a long-term effect on gym visits compared to a treatment where students were paid only once. They conclude this difference is because it takes time for habits to change.

This literature on behavioral treatments and technology use is very relevant for our paper, because we examine the effect of randomly assigned treatments on Mirt stove use within the context

[4] We use “field experiment” interchangeable with “randomized treatment trial.” For an overview of the use of field experiments in developing countries, please see Duflo et al. (2008)

of a field experiment. Users received Mirt stoves under three randomly assigned price schemes and two networking options that are discussed in the following section. We then examine the extent of stove use over the course of more than one year.

3. Sampling and Experimental Design

3.1 Sample site selection

The sample sites are chosen from that used in a 2012 study by the Environmental Economics Policy Forum for Ethiopia (EEPFE) based at the Ethiopian Development Research Institute (EDRI) that identified 110 sites from which forestry and related data were collected. It was the community level information collected from these study sites that was used to assign treatments and test randomization.

Out of the 110 sites, we remove 15 sites that were covered during a pilot survey conducted to inform our research. We also remove all sites from Tigrai Regional State as open three-stone traditional stoves, which are used throughout most of the rest of Ethiopia, are not used in this region. We also remove three sites from Borena District, because *injera* is not typically eaten. From the 81 remaining sites we select 36 sites at random using proportionate random sampling based on regional state forest cover. Forest cover was used, because most fuelwood in the sample areas come from forests. Forested area also proxies well for total regional state population and area. Based on this criterion 20% of the observations come from Amhara region, 50% from Oromiya region and 30% from SNNP. The selection of sites within each regional state was done using simple random sampling.

A site in this study is a village (locally called *Got* or sub-*Kebele*). A list of households for each site was obtained from the local administration. All sites have formal or informal forest user groups. A total of 504 households were selected from the 36 villages (14 households from each site) using systematic random sampling. A total of 360 households (10 households from each site) received the improved Mirt stove and the rest were control households who did not receive the Mirt stove. In this paper, which deals with Mirt stove use frequency, we do not use results from the controls.

3.2 Improved Mirt Stove Distribution, Stove Use Monitors (SUM) and Information Provided to Respondents

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

Respondents were informed that if they agreed to participate, their stove use would be monitored and enumerators would come to their houses to download data a few weeks after installation or re-installation of the stove use monitor (SUM) devices. We rely on automatic temperature measurement rather than respondent reporting to identify cooking events, which we believe is more reliable. Thomas et al. (forthcoming) indeed found that users in Rwanda overstated their use by 40% compared with their actual use frequency.

The SUMs measure stove temperature and are approximately the size of a watch battery. They were purchased from Berkeley Air Monitoring Group of Berkeley, California. The recording intervals on the SUMs are adjustable and the memory on 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 data are downloaded for analysis.

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 the only activity the devices would perform is to periodically record the temperature of the stoves.[5]

[5]Fieldworkers also 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 SUMS 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.

Fieldworkers informed participants that stove use is not required to receive a stove, but they were 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.

Households were typically invited to a centralized location, such as a school or *kebele* 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. Typically, the following day enumerators came to install the stove in the kitchen area, either in the main home (47%) or in a separate kitchen (53%) by mudding together the six concrete pieces. They also gave the cooks training on how to use the stove and the SUM device was installed on the stove using heat-resistant tape.

The SUM was placed exactly at the back of the stove, because it is the coolest location on the Mirt stove. It was initially believed that the front of the stove was the coolest and several SUM devices were ruined in the first round, because their temperature tolerances were exceeded. Because of such malfunctions, the panel data used in this study are not balanced, with the second monitoring period having the fewest observations. Over the four rounds a maximum of 1,440 observations (360 per round) are possible, but because of stove use monitor failures, observations in each round are about 300, with a total of 1,209 observations across all rounds. There was no attrition in the sample[6]

[6]A subset of 108 households also participated in three controlled cooking tests. These tests are not part of this study, but in the analysis we take account of which households participated in this additional research.

Supervisor Demonstrating Placement of Clay Plate Respondent Takes 2 of 6 Mirt Stove Parts Home
Source: Randall Bluffstone Source: Randall Bluffstone

In the home or separate kitchen 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 a 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 a SUM device recorded data for between 37 and 56 days. The median recording period was 49 days.

The traditional technology in most of Ethiopia is the three-stone tripod fire. It is not technologically possible to reliably monitor this type of stove using the SUMs. The use of the traditional stove was therefore not monitored and we must imperfectly infer the degree to which the Mirt stove is exclusively used and cannot say much about possible rebound effects in which efficiency gains are eroded by increased cooking.

Two features of the Mirt stove likely mitigate large rebound effects. First, the Mirt stove is highly specialized for *injera* baking, which makes it very unlikely that it will be shifted to other uses (a potential rebound effect). As will be discussed below, we find that the Mirt stove is on average used 2-3 times per week, which is a very traditional and typical interval for *injera* baking. It is also almost exactly the mean reported baking interval reported in Gebreegziabher (2015). It is therefore very likely that on average most if not all *injera* baking is done on the Mirt stove and therefore captured by our monitoring.

Second, the Mirt stove has the capacity to do cooking with waste gases before they are vented and over 80% say they use the second burner for cooking stews and coffee. This feature offers additional efficiencies that likely reduce the need to cook on two stoves simultaneously. In sum, while we cannot rule out major rebound effects, there is little reason to expect they would be large.

3.3 Treatments

The participating households received the stoves under 6 randomized 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 in those sites 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 stove use frequency.

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 those who pay for their stoves use them more frequently. Other payment treatments were not used, because of budgetary limitations and it is possible that if, for example, full cost was charged, some respondents would have refused to take the stoves.

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 were also 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 the stove use 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 stove use frequency.

All households in a village who were randomly selected received only one of the six treatments below, which remained constant across all four monitoring periods. Households receiving other treatments are used as controls for the particular treatments being analyzed.

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.

One treatment is assigned to 6 sites making a total of 6 sites receiving the same treatment. Households that received treatments 1 and 2 received 50 Birr if they used the Mirt stove twice per week during the first monitoring period in addition to receiving the stoves for free. Households that received treatments 3 and 4 paid 25 Birr for the stove. Respondents who paid for stoves did not simultaneously receive incentives for using their stoves, because this combination was judged to be confusing for users. Those in treatment groups 5 and 6 received their stoves for free. Treatments 2, 4 and 6 received the network treatment aspect in addition to one of the three financial treatment aspects.

Table 1 presents the numbers of usable observations by round and treatment. We see that the availability of usable stove use frequency data are quite uniform across monitoring periods and treatments, with treatment 1 being a few percentage points lower than the other treatments. These results suggest that any SUMs failures were not systematic.

Table 1 Available Cooking Event Observations 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

It is important that the treatments are randomized across sites. For this randomization we use variables/indicators from community level data collected at the selected sites. These indicators include wealth variance, existence of forest rules and regulations, and percent of biomass change over 5 years, distance to market and percent of households with access to piped water. These variables were derived from the community survey conducted by EEPFE in 2012.

3.4 Data Management

The SUMs generate an enormous amount of temperature data for each household, but our interest is only in the frequency of stove use, which is defined as the number of times that cooking events occur per week. We define a cooking event as having taken place if the recorded temperature exceeded 40 degrees Celsius. The number of peaks is calculated with the help of signal processing and analysis (O'Haver, 1997) and a cooking event is one where there is a peak above 40 degrees Celsius. The use of 40 degrees Celsius helps remove the possibility of counting peaks that are really due to normal temperature variation not related to cooking. The maximum ambient temperature recorded in the sample was 35 degrees Celsius and the average was 25 degrees.

To measure the frequency of use during a monitoring period, we simply count the number of times the stove surface temperature exceeds the critical value. The average stove use in a round is therefore total frequency/total days temperature was measured. We express this value per week, because *injera* tends to not be cooked every day. The key outcome variable is the number of times the Mirt stoves are used per week.

Using this experimental design and data management system, the key outcome data were generated and assembled. Other data on household characteristics, demographics and cooking situations were collected using a household survey. The following section discusses the empirical methods used to analyze these data.

4. Variables and Empirical Methods

There are two major steps to the analysis. Because our stove distribution was random at the household level, we begin our empirical analysis with use frequency descriptive statistics and rank-sum tests of differences that exploit this randomness. Second, we evaluate the factors affecting the number of Mirt stove cooking events that occur on average per week during the four monitoring periods.

A key objective of the analysis is to test whether the 6 treatments applied affect stove use frequency. That is, we attempt to determine the effect of each treatment on stove use had the stove been distributed without that treatment. For example, we might want to test the effect of our network treatment combined with free distribution (i.e. treatment 6) compared with free distribution and no network treatment <u>to the same household</u> (treatment 5). 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 1, 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.

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

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. If differences are observable, they can be controlled for in regressions or used to match households and estimate average treatment effects. Both these methods are used and presented below.

Because some differences may not be observable and therefore they cannot be controlled for, without simultaneous access to parallel universes, random distribution of treatments is the best context in which to use $E(Y_j|T_j=0)$ as a proxy for $E(Y_j|T_i=1)$. In other words, successful and full randomization is perhaps the most effective way to assure that treatment i and control j groups are fully comparable.

As will be established and discussed in detail below, using Kruskal-Wallis tests based on 7 key community characteristics we confirm the hypothesis that the treatments are from the same population; randomization therefore appears to have been largely successful, though it is not possible to test successful balance based on unobservables.

Because randomization seems to have been successful, in the following section we present 15 non-parametric Mann-Whitney rank sum tests. We also use regression methods to test the robustness of these findings to avoid confounding treatment effects with the influence of other variables and to evaluate the effects of variables other than our 6 treatments on the frequency of stove use. For this purpose we first present pooled OLS results with round and district (i.e. *woreda*) fixed effects and errors clustered by household. We use pooled OLS instead of a random effects model, because only our stove use data vary over time. We therefore do not have a true panel. However, we also provide the results from random effects estimation in Appendix B.

Table 2. Descriptive Statistics of Covariates used in the Pooled OLS Models

Variable	Obs	Mean	Sd. D.	Min	Max
Age of respondent	1436	42.134	13.159	20	90
Sex of respondent 1 if male, 0 if female	1436	0.880	0.325	0	1
Marital of respondent 1 if married, 0 otherwise	1440	0.900	0.300	0	1
Education- 1 if illiterate, 0 if literate	1440	0.383	0.486	0	1
Family Size in adult equivalent	1440	4.888	1.859	0	10.88
Religion -1 if Christian, 0 Muslim	1436	0.710	0.454	0	1
Livestock in tropical livestock units (TLU)	1424	5.077	3.711	0	26.23
Children under 15	1316	3.036	1.548	1	11
Walking Distance from household to nearest road (two-way) in minutes	1436	63.214	105.565	0	840
Participation in controlled cooking test, 1 if yes, 0 if no	1440	0.300	0.458	0	1
Risk coefficient	1420	3.815	1.212	0.08	6.58
Average temperature in ^{0}C	1409	24.982	3.296	12	32.5
Average number of injeras baked at a time	1436	19.721	10.292	0	55
Use the of improved stove for any other purpose other than baking 1 if yes, 0 if No	1440	0.333	0.472	0	1
The average quantity of flour in kg used per cooking	1428	4.677	2.375	0.75	15
Type of flour, teff=1, mixed or no teff=0	1440	0.156	0.363	0	1
Place of stove installed - 1 if inside the house, 0 separate kitchen	1432	0.360	0.480	0	1
Main fuel used for baking - 1 if fuel wood, 0 otherwise	1440	0.747	0.435	0	1

In these models we adjust for key household level variables that may confound our interpretation of the treatment effects. These variables are given in Table 2 above. Most variables are self-explanatory, except the risk coefficient. Risk preferences are measured by eliciting participants' preferences between lotteries and a list of sure amounts between the low and high amounts of the lottery. The prize to be won at the lottery was fixed at 40 Birr and the low amount at 0, with the sure amounts increasing from 0 Birr to 40 Birr in steps of 1 Birr.

The point at which a participant switched from preferring the lottery to preferring the sure amount of money was encoded as the certainty equivalent for the participant, i.e. the amount for which the participant is indifferent between the sure amount and playing the lottery. There were in total seven such choice tasks, with the probability of winning the prize ranging from 1/8 to 7/8 and the tasks administered in random order. The measure used in this paper is simply the average certainty equivalent for the seven choice tasks. For a detailed description of the data, see Vieider et al. (2014).

As a robustness check, we utilize 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.[7] The predicted values of these probit models are then used as propensity scores that match treatment and control households. All possible treatment combinations are estimated for a total of 15 propensity score matching results. We estimate propensity for observations only within the region of common support. The independent variables included in the Probit model are ones that are likely similar for members in a community. These variables are respondent education, distance from the respondent's household to the nearest road and respondents' religions. For some treatment combinations all three variables are included (e.g. comparison of treatment 1 versus treatment 6), but for some propensity score models with all three variables included the treatment and control groups were not balanced. Fewer than the three independent variables were therefore used in the propensity score estimation, but distance to the nearest road (a continuous variable) was always included. Details on the propensity score models are available from the authors.

5. Results

We now present the results of our empirical analysis, which seeks to evaluate whether people actually use the Mirt stoves at levels consistent with minimal stove stacking. We begin with a discussion of the frequency with which the Mirt stoves are used based on the SUMs that were attached to the stoves for between 39 and 58 days a maximum of four times per household. These results are followed by analyses of stove use based on the treatments applied. We test for whether the treatment applications come from the same population and therefore whether the randomization was successful. We then present the results of our pooled OLS regressions, followed by nearest neighbor propensity score matching

5.1 Basic Stove Use Frequency

[7] 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.

We find that across the four rounds of our sample users who received the Mirt stove on average used it 2.27 times per week. This result is in accord with our household survey and Gebreegziabher (2015) (for the subsample that participated in the controlled cooking test or CCT), which suggests that respondents cook *injera* an average of about 2.5 times per week. We therefore find that on average households use the Mirt stove frequently and seemingly appropriately.

Table 3 Frequency of Mirt Stove Use Overall and by Round (Average Times per Week)

	Overall Sample	Round 1 (June-August 2013)	Round 2 (August - October 2013)	Round 3 (March-May 2014)	Round 4 (May – July 2014)
Mean	2.27	1.863	2.078	2.722	2.395
Minimum	0	0	0	0	0
Maximum	21.367	9.545	12.778	21.367	19.121
10th percentile	0.132	0.151	0.132	0.172	0.00
25th percentile	0.686	0.515	0.738	0.829	0.663
50th percentile	1.846	1.483	1.846	2.086	1.935
75th percentile	3.144	2.80	3.024	3.728	3.484
90th percentile	4.644	4.048	4.106	6.223	4.800
Observations	1209	301	287	319	302

Table 3 presents the mean stove use per week by round and breaks the distribution down into 10th, 25th, 50th, 75th and 90th percentiles. Looking not only at the mean, but at the whole distribution is important, because we find based on Shapiro-Wilk tests for normality that frequency of stove use is not normally distributed (Z=12.078; Prob.>Z=0.000). By round and regional state, frequency is also not normally distributed and in all cases the P value is approximately 0.000.

We find that across the four rounds only 8.85% of observations had zero use, suggesting that virtually all households in all rounds actually used the stove. About 80% of observations indicate average stove use at least once every two weeks and 68% of observations averaged at least one cooking event per week. Table 3 indicates that in later rounds households generally used the stove more frequently than in early rounds, with median stove use plateauing at about 2.0 times per week. The pattern is not completely monotonic, however, with a dip in average frequency between rounds 3 and 4, which could be due to reduced cooking during the warmest season, but it seems that as cooks gain experience with the stove they generally use it more frequently.

In the latter three rounds 25% of households use the stove an average of at least three times per week and except for in Round 4, 90% of users average positive usage. Select users cook with the Mirt stove an average of 2 to 3 times per day, suggesting they use it for purposes in addition to cooking *injera*. This finding is also in accord with a satisfaction survey we conducted that is discussed in Gebreegziabher et al. (2015) in which 84% of users report using the stove for purposes in addition to cooking *injera*, such as baking bread, making coffee and cooking stews.

Table 4 Mean Cooking Events 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	2.640	2.381	2.500	2.901	2.731
Oromiya	2.500	1.976	2.140	3.000	2.83
SNNP	1.672	1.350	1.691	2.170	1.480

Table 4 presents the mean number of cooking events per week by region and round. Consistent with the previous table we find roughly increasing average usage across rounds regardless of region, with a dip between rounds 3 and 4, perhaps due to seasonality. Amhara and Oromiya Regional State have higher average usage than SNNP, with Oromiya having the highest average number of cooking events in rounds 3 and 4. SNNP is lowest, because the cultural practice is not to eat *injera* at every meal. As discussed above, the Mirt stove, while being capable of cooking other foods, is primarily designed to cook *injera*.

5.2 Effect of Treatments on Stove Use Frequency Based on Randomized Distribution

In the above discussion we have shown that households use the Mirt stove on average consistent with its expected purpose to cook *injera*. We now turn our attention to a preliminary analysis of why the stove is used, with a main emphasis on the six treatments equally applied when the 360 stoves were randomly distributed to households. As discussed in the previous section, treatments were applied at the site 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 5
Kruskal-Wallis Tests that the 6 Treatment Assignments Come from the Same Site Characteristic Distributions

	Households in Kebele	Households in Forest User Group	Wealth Variance in Forest User Group	Existence of Forest Rules/ Regulations	% Forest Biomass Change Over 5 Years (Respondent Assessed)	Distance to Market in Minutes)	% Households with Access to Piped Water
X^2 (5)	5.246	3.085	6.166	.000	6.243	10.997	20.147
P Value	.387	.687	.290	1.000	.283	.051**	.001***

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

Table 5 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 7 different key community characteristics. The randomization appears to have been successful. There are no systematic differences in group size, wealth distribution, forest change or forest management as assessed by respondents. The Kruskal-Wallis test for distance to market is significant at approximately the 5% level and access to piped water is significant at more than the 1% level. We do not view this as a major problem, though, because it is the community variable that is least directly related to improved stove use, which we evaluate in this paper.

Table 6 provides key descriptive statistics on the frequency of stove use by treatment. We find that treatment 5 in which stoves are distributed free, with no other treatment aspect offered resulted in the highest average stove use of about 2.7 times per week, followed by treatment 1, which has the largest maximum use of three times per day. Measured cooking events per week are all, however, well within one-half standard deviation of each other, raising questions of statistical significance.

Table 6 Frequency of Mirt Stove Cooking Events per Week by Treatment

	Obs.	Mean	S. D.	Max.
Received 50Birr use payment, but no network (Treatment 1)	172	2.447	2.828	21.367
Received 50 Birr use payment, plus network (Treatment 2)	216	2.098	1.836	10.047
Paid 25 Birr, but received no network (Treatment 3)	211	2.215	1.859	8.888

Paid 25 Birr for the stove, plus network (Treatment 4)	213	1.944	2.058	12.778
Received stove for free and no network (Treatment 5)	207	2.653	2.320	13.806
Received stove for free, plus network (Treatment 6)	186	2.381	2.277	10.501

As was mentioned above, stove use frequency is not normally distributed. We therefore use Mann-Whitney (i.e. Wilcoxen rank sum) tests to evaluate whether there are statistically significant differences in stove use based on treatment. Table 7 presents these test results, with means given in Table 6. In the table row treatments are compared with column treatments, implying that positive numbers indicate the row treatment has a larger estimated treatment effect than the comparable column treatment.

Table 7 Rank Sum Mann-Whitney Test P Values by Treatment

	Treatment1	Treatment 2	Treatment 3	Treatment 4	Treatment 5
Treatment 2	-0.41				
Treatment 3	0.99	0.33			
Treatment 4	-0.03***	-0.17	-0.05**		
Treatment 5	0.07*	0.01***	0.03**	0.00***	
Treatment 6	-0.90	0.40	0.82	0.04**	-0.25

Positive numbers indicate the row treatment has a larger effect on frequency of use than the column treatment. Refer to Table 6 for mean values. ***, **, * indicate significant at the 1%, 5% and 10% levels

The results suggest that stove users often decide their stove use on bases other than the treatments they received, because only seven of 15 possible unpaired differences are statistically significant. That said, more than half of the statistically significant differences involve treatment 5, which is statistically more effective than treatments 1, 2, 3 and 4, but not 6. These results suggest that those who paid for their stoves or conversely were paid to use their stoves a minimum of twice per week during the first round on average use their stoves <u>less</u> than if stoves are offered for free.

The network treatment in which users were assembled into groups for additional training and Q&A shows mixed results. Treatment 6 dominates treatment 5, but those who received treatment 3 used their stoves more than those who received treatment 4. The statistically significant comparison of treatments 6 and 4 is also interesting, because both have the network treatments, but different randomized monetary treatments. Those who received their stoves for free used them more than those who paid 25 birr, conditional on also receiving the network treatment. These results suggest that we can reject the hypothesis that paying for the stove spurs more frequent use than when stoves are distributed for free.

5.3 Propensity Score Matching and Treatments

We now turn our attention to the findings from our nearest neighbor propensity score matching models. Table 8 presents the average effect of each treatment on frequency of Mirt stove use per week. These effects are compared with the counterfactual of one of the other treatments. Also included in each cell are the number of actual matched treatment and control households, the two-tailed p-value based on the estimated t-statistic and 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, absolute propensity score difference between treated and controls and average frequency of stove use for matched treated and control households are available from the authors.

Table 8 Average Treatment Effects of Row Treatment vis-à-vis Column Control

	Treatment1	Treatment 2	Treatment 3	Treatment 4	Treatment 5
Treatment 2	**0.01** (0.99) [236/135]				
Treatment 3	**0.11** (0.77) [240/165]	**0.23** (0.30) [240/179]			
Treatment 4	**-0.86** (0.03)** [200/162]	**-0.53** (0.04)** [200/181]	**-0.12** (0.62) [200/201]		
Treatment 5	**0.57** (0.03)** [240/130]	**0.59** (0.01)*** [240/370]	**0.28** (0.28) [240/165]	**0.89** (0.00)*** [240/146]	
Treatment 6	**0.43** (0.33) [236/140]	**0.53** (0.02)** [236/196]	**0.01** (0.99) [236/202]	**0.17** (0.51) [236/126]	**0.50** (0.07)* [236/185]

ATT in bold. Positive numbers indicate the row treatment has a larger effect on frequency of use 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

Results in Table 8 are consistent with results already discussed, but utilize explicitly matched samples. We find that those who receive stoves under treatment 4 use their Mirt stoves less than other treatments, except for treatments 3 and 6, which are statistically insignificant. The treatment 4 effect is statistically significant at least at the 5% level for treatments 1, 2 and 5. Marginal effects of treatments 1 and 5 versus treatment 4 are particularly large at 0.86 and 0.89. These effects are about 1/3 of the pooled sample mean.

The average treatment effect of treatment 5 is also large vis-à-vis treatments 1 and 2 at 0.57 and 0.59 more times per week (about 25% of the sample mean). In the matched sample we find that households receiving stoves under treatment 6, which includes the network treatment along with free distribution, on average use the Mirt stove 0.50 times (22%) more per week than those receiving treatment 5 (no network). In the matched sample those receiving treatment 6 also on average use their stoves 0.53 times more per week than those receiving treatment 2 (50 Birr use incentive, plus network).

5.4 Pooled OLS Regression Results[8]

We now present our pooled OLS results with round and district (i.e. *woreda*) fixed effects with errors clustered by household. The dependent variable is the number of times per week respondents use the Mirt stoves. In these models, though our randomization appears to have been successful, we adjust for key household level variables that could confound our interpretation of the treatment effects. The regression models also allow us to analyze the effects of non-treatment variables on stove use frequency. The results of 6 models are presented in Table 9.

Under the null hypothesis that $R^2 = 0$, the F statistic follows the F distribution with 60 and 984 degrees of freedom. We find for model 5, which is the focus of our interpretation, that we can reject the null hypothesis that the explanatory variables have no impact on the stove frequency at the 1.0% level (F=16.4).[9] In model 6 we include interactions between treatments and rounds to see whether treatments have different effects over time, however when we compare model 5 (without interaction) with model 6 (with interaction), we find that the F statistic is insignificant. This result suggests that the interaction terms do not improve the explanatory power of the regression.

Based on model 5, we find that those who received stoves under treatment 2 and 6 use their Mirt stoves more frequently than treatment 5 (the omitted treatment). The dominance of treatment 6 vis-à-vis treatment 5 is consistent with the other empirical methods used. As the only difference between the treatments is the network component, the result indicates that without payment incentives, networking spurs actual use of the Mirt stoves. The two previous methods indicated that treatment 2 spurred stove use no better (and compared with treatment 5 worse) than other treatments. The pooled OLS results suggest the opposite conclusion and is statistically significant at

[8] The random effects estimation results with round and district (i.e. *woreda*) fixed effects are found in Appendix B.

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

the 1% significance level for both treatments 2 and 6 in Models 5 and 6, which include district fixed effects. Consistent with previous methods, treatment 3, payment of ETB 25 and no network, has a negative impact on stove use frequency compared to treatment 5. These results therefore also contradict the hypothesis that paying for the Mirt stove promotes use.

Treatments 2, 3 and 6 are insignificant without the district fixed effects. These pooled OLS results therefore suggest that once unobserved locational differences are taken into account, treatments that include the network aspect may be important for the use of Mirt stoves, though the insignificant effect of treatment 4 makes this conclusion less clear. Consistent with results in table 3, we find that households on average increased their use frequency over time. Compared with round 1, which is the base category, we find that stove use frequency is greater in the second, third and fourth rounds than in round 1. Consistent with our other results, we also find the largest effect in round 3, which was measured during the March to May 2014 period.

Table 9 Regression results from pooled OLS. Dependent variable is number of times per week Mirt stove is used

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Age of respondent	-0.010	-0.007	-0.005	-0.007	-0.004	-0.004
	(0.007)	(0.007)	(0.009)	(0.009)	(0.007)	(0.007)
Sex of respondent	-0.155	-0.104	-0.331	-0.380	-0.167	-0.184
	(0.367)	(0.367)	(0.392)	(0.387)	(0.344)	(0.348)
Marital status is married	0.604*	0.612*	0.685*	0.634*	0.291	0.308
	(0.331)	(0.334)	(0.373)	(0.344)	(0.316)	(0.321)
Education is illiterate	0.131	0.133	-0.081	-0.107	-0.090	-0.090
	(0.212)	(0.213)	(0.213)	(0.207)	(0.190)	(0.192)
Family size	0.288***	0.271***	0.191**	0.200**	0.157**	0.158**
	(0.088)	(0.088)	(0.086)	(0.086)	(0.075)	(0.076)
Religion is Christian	-0.691***	-0.670***	-0.825***	-0.822***	0.001	0.002
	(0.256)	(0.256)	(0.300)	(0.300)	(0.347)	(0.343)
Livestock in TLU	-0.047*	-0.050*	-0.063***	-0.075***	-0.083***	-0.085***
	(0.026)	(0.025)	(0.022)	(0.024)	(0.026)	(0.026)
Children under 15	-0.099	-0.070	-0.037	-0.041	-0.006	-0.004
	(0.085)	(0.083)	(0.080)	(0.077)	(0.072)	(0.073)
Distance to road	-0.002***	-0.002***	-0.002***	-0.002***	-0.002*	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Round2	0.272**	0.268**	0.271**	0.284**	0.319***	0.483***
	(0.114)	(0.116)	(0.118)	(0.119)	(0.118)	(0.167)
Round3	0.861***	0.867***	0.894***	0.891***	0.904***	0.782***
	(0.160)	(0.160)	(0.164)	(0.163)	(0.169)	(0.293)
Round4	0.552***	0.563***	0.593***	0.592***	0.613***	0.604**
	(0.147)	(0.148)	(0.150)	(0.148)	(0.149)	(0.241)
Participation in CCT		0.350*	0.304	0.302	0.328**	0.331**
		(0.206)	(0.196)	(0.190)	(0.164)	(0.164)
Risk Coefficient		0.052	0.113	0.129	0.152**	0.149**
		(0.083)	(0.080)	(0.080)	(0.073)	(0.073)
Average temperature in ^{0}C			0.125***	0.108***	0.145*	0.145*
			(0.037)	(0.034)	(0.080)	(0.084)
Number of Injera baked			0.037***	0.039***	0.028*	0.027*
			(0.010)	(0.010)	(0.015)	(0.015)
Stove use for other purpose			0.418	0.471*	0.476	0.479

	(1)	(2)	(3)	(4)	(5)	(6)
			(0.258)	(0.244)	(0.304)	(0.308)
Quantity of flour			0.027	0.030	-0.023	-0.023
			(0.055)	(0.056)	(0.054)	(0.054)
Flour kind is teff			-0.266	-0.296	-0.176	-0.163
			(0.216)	(0.212)	(0.261)	(0.264)
Stove installed inside			-0.349	-0.311	-0.461	-0.463
			(0.245)	(0.252)	(0.340)	(0.343)
Main fuel is fuelwood			-0.349	-0.411	-0.207	-0.192
			(0.263)	(0.289)	(0.247)	(0.248)
treatment_1				0.125	-0.225	-0.634*
				(0.354)	(0.532)	(0.382)
treatment_2				-0.130	0.977***	1.320***
				(0.235)	(0.271)	(0.322)
treatment_3				-0.007	-0.749**	-0.747**
				(0.246)	(0.337)	(0.342)
treatment_4				-0.660***	-0.301	-0.498
				(0.229)	(0.642)	(0.604)
treatment_6				0.173	0.832***	1.078***
				(0.270)	(0.296)	(0.322)
treat1_round2						0.180
						(0.470)
treat2_round2						-0.432*
						(0.241)
treat3_round2						-0.406*
						(0.242)
treat4_round2						0.153
						(0.358)
treat6_round2						-0.252
						(0.282)
treat1_round3						0.804
						(0.598)
treat2_round3						-0.181
						(0.458)
treat3_round3						0.147
						(0.393)
treat4_round3						0.347
						(0.443)
treat6_round3						-0.361
						(0.486)
treat1_round4						0.643
						(0.526)
treat2_round4						-0.661*
						(0.365)
treat3_round4						0.271
						(0.344)
treat4_round4						0.282
						(0.367)
treat6_round4						-0.392
						(0.419)
Constant	1.537***	1.079*	-2.225*	-1.609	-2.028	-1.983
	(0.554)	(0.616)	(1.337)	(1.237)	(2.070)	(2.131)
R2	0.0728	0.0784	0.1442	0.1570	0.2499	0.2577
N	1078	1066	1045	1045	1045	1045

Note: *** = Significant at 1%, ** = Significant at 5%, * = Significant at 10%, District dummies are included in both model 5 and 6 but are not reported for the sake of economizing spaces. The result shows that there are differences in stove use among the different districts.

Several control variables are also found to be important. We find that those respondents who participated in the CCT use the stove more frequently. This result appears to be evidence of a "Hawthorne" effect in which more intensive intervention changes respondent behavior. We also find that households with higher temperatures at the time of the household survey use their Mirt stoves more than those in cooler areas. This result is expected as warm temperature allows *injera* to be stored for fewer days. The number of *injeras* baked per cooking session is positively related to stove use frequency. Similarly, family size is consistently positive and significant. On the other hand, those who live in more remote locations (i.e. farther from roads) and who have more livestock use their stoves less than those closer to roads and who invest less in livestock.

An interesting result is that those who are found to be less risk averse in the field-based risk experiment use their Mirt stoves more than those who are more risk averse. This result may indicate that those with greater risk tolerance could have greater willingness to experiment. If successful, this experimentation could then lead to higher levels of usage. We do not suggest that this is necessarily the right explanation. As discussed in Yesuf and Bluffstone (2009) and Vieider et al. (2014) estimated risk aversion is a function of wealth, income and market failures. The negative coefficient estimate may therefore be picking up such omitted variables.

6. Conclusions

This paper uses a randomized experimental design and real-time electronic stove use monitors to evaluate the frequency with which villagers use improved biomass-burning Mirt *injera* cookstoves in rural Ethiopia. Understanding whether, how much and why improved cookstoves are used is important, because stove use is a critical determinant of total greenhouse gas and indoor air pollution reductions. Confirming use is, for example, a critical aspect of crediting improved cookstoves' climate change benefits under REDD+.

We find that Ethiopian households in the study area use the Mirt stove in accord with its primary purpose to cook *injera*, which is estimated to use more than half the fuelwood in the country. Reducing fuelwood use for *injera* baking is therefore of keen importance. Households on average use the stove 2.4 times per week in the final round of monitoring, but there is significant spatial variation, with households living in SNNP Regional State on average using the stoves just over half as often as households in Amhara and Oromiya. We believe this result is related to SNNP having a less strong tradition of *injera* baking than the other two regions. This result echoes the literature, which emphasizes the importance of local context and the difficulty – even with a relatively specialized cuisine and a technology designed for that type of cooking – of developing one-size-fits-all improved stoves. A homogeneous approach worked poorly in the past and we see that in 2014 local conditions continue to matter.

In general, stove users over time more frequently use their Mirt stoves. Though there appears to be a seasonal element, with most frequent cooking occurring during the March to May 2014 reporting period, as users gain experience they use the stove more frequently, which is a very good indication that the stoves are appropriate for most of the local conditions.

Our results suggest that the conditions under which stoves are given matter, but our estimates and therefore conclusions about treatments vary somewhat by estimation method and whether regional effects are explicitly accounted for. Unobserved district effects indeed pick up a lot of the variation in the pooled OLS models, a feature that is not explicitly considered in the Mann-Whitney tests and propensity score matching models.

Giving the Mirt stove away for free and providing network treatments for users (treatment 6) is found in a number of models to lead to more use and in the pooled OLS model treatment 2, which also includes a network aspect, also spurs more frequent use. We find no evidence, however that stove recipients use the stoves more if they pay for them. Indeed, if anything, treatments 3 and 4, which include 25 birr payments by respondents, result in less frequent stove use. Free distribution seems to be the most robust monetary treatment if the policy goal is to promote stove use. Results are a bit less clear with regard to the networking, though the pooled OLS results adjusting for region fixed effects suggest – as long as recipients are not required to pay for their stoves – that networking increases use frequency.

Is the Mirt stove a technology that could potentially be creditable under REDD+? In other work (Gebreegziabher, 2015) we are able to document that on average Mirt stoves use less wood to cook typical meals and users say they like them. The current paper indicates that if put into the field, people will use the stoves and just need to be given the stoves. No additional financial incentives seem to be needed, which could make the stove a potentially cost-effective REDD+ compliance mechanism. These results together suggest that in the field under uncontrolled conditions the Mirt stoves reduces fuelwood consumption.

Using the evaluation framework of Johnson et al. (2009), open questions remain about the portion of the fuelwood that is unsustainably harvested and the emission factor of the technology. The greenhouse gas emission factor depends only on the technology rather than behavior. In previous literature (e.g. Megen Power, 2008) the stove was found to emit fewer gases, but more needs to be done particularly on the indoor air quality implications of the stove. The importance of unsustainable harvest percentage is increasingly discussed in the literature and though the Clean

Development Mechanism has adopted a default unsustainable percentage for Ethiopia of 88% (UNFCCC, 2012) little is probably known about the details of the Ethiopian context. Perhaps this is the most important arena where local context matters, but it may be that such granularity is not necessary if the first two conditions are met. After all, are only areas with known and continued unsustainable fuelwood harvests to be rewarded with better technologies? Such an approach appears counterproductive and potentially could create incentives for unsustainable harvest.

References

Agrawal, A., and Angelsen, A. 2009. "Using Community Forest Management to Achieve REDD+ Goals." In A. Angelsen (ed.), *Realising REDD+: National Strategy and Policy Options*. Bogor, Indonesia: Centre for International Forestry Research.

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.

Amacher, G., L. Ersado, D. L. Grebner and W. F. Hyde. 2004. "Disease, Microdams and Natural Resources in Tigray, Ethiopia: Impacts on Productivity and Labour Supplies." *Journal of Development Studies*: 40: 122-145.

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.

Ashraf, N., J. Berry and J. Shapiro. 2010. "Can higher prices stimulate product use? Evidence from a field experiment in Zambia." *American Economic Review*. 100: 2383-2413.

Beltramo, T. and D.I. Levine. 2013. "The Effect of Solar Ovens on Fuel Use, Emissions and Health: Results from a Randomised Controlled Trial." *Journal of Development Effectiveness*. 5: 178–207.

Barnes, D, K. Openshaw, K. Smith and R. van der Plas. 1993. "The Design and Diffusion of Improved Cooking Stoves." *World Bank Research Observer*, Vol. 8: 119-141.

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.

Bluffstone, R. 1995. "The Effect of Labor Markets on Deforestation in Developing Countries Under Open Access: an Example from Rural Nepal." *Journal of Environmental Economics and Management*. 29: 42-63.

Bond, T., Doherty, S., Fahey, D., Forster, P., Bernsten, T., DeAngelo, B., Flanner, M., Ghan, S., Karcher, B., Koch, D., Kinne, S., Kondo, Y., Quinn, P., Arofim, M., Schultz, M., Schulz, M., Venkataraman, C., Zhang, H., Zhang, S., Bellouin, N., Guttikunda, S., Hopke, P., Jacobsen, M., Kaiser, J., Klimont, Z., Lohmann, U., Schwarz, J., Shindell, D., Storelvmo, T., Warren, S. and Zender, C. 2013. "Bounding the Role of Black Carbon in the Climate System: A

Scientific Assessment." *Journal of Geophysical Research Atmospheres*, 118: 5380–5552. doi: 1002/jgrd.50171. Available at http://onlinelibrary.wiley.com/doi/10.1002/jgrd.50171/pdf

Burwen, J. and Levine, D. I. 2012. "A rapid Assessment Randomized-Controlled Trial of Improved Cookstoves in Rural Ghana," *Energy for Sustainable Development*. 16:328-338.

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.

Cooke, P. 1995. *Household Heterogeneity, Time Allocation and the Use of Environmental Products: Responses to Deforestation by Rural Nepali Households*, Unpublished Ph.D. Dissertation, University of Washington, Seattle, WA.

Cooke, P., G. Köhlin, W. F. Hyde. 2008. "Fuelwood, Forests and Community Management – Evidence from Household Studies." *Environment and Development Economics*. 13: 103-135.

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.

Economist. 2010. "Not a Small Problem: Will REDD Trample on the Rights of Traditional Forest Folk?" *The Economist*, September 25, 2010.

Gebreegziabher, Z, A. Beyene, R. Bluffstone, S. Dissanayake, P. Martinsson, M Toman. 2015. "Can Improved Biomass Cookstoves Contribute to REDD+ in Low-Income Countries? Results from a Randomized Control Trial in Ethiopia." World Bank Working Paper.

Gifford, M. L. 2010. *A Global Review of Cookstove Programs*, Unpublished Masters Thesis, UC-Berkeley, downloaded 2/28/14 from mlgifford.wordpress.com

Gil, J. 1987. "Improved Stoves in Developing Countries: a Critique." *Energy Policy*, April: 135-144.

Grieshop, A. P., P. J. D. Marshall and M. Kandlikar. 2011. "Health and Climate Benefits of Cookstove Replacement Options." *Energy Policy,* 39: 7530-7542.

IEA 2006. *World Energy Outlook 2004: Energy for Cooking in Developing countries (ch. 15*). International Energy Agency. Paris: IEA.

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

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, M, R. Edwards, A. Ghilardi, V. Berrueta, D. Gillen, C. A. Frenk and O. Masera. 2009. "Quantification of Carbon Savings from Improved Biomass Cookstove Projects. *Environmental Science and Technology.* 43: 2456-2462.

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

Kar, A., I. Rehman, J. Burney, S. P. Puppala, R. Suresh, L. Singh, V. K. Singh, T. Ahmed, N. Ramanathan and V. Ramanathan. 2012. "Real-Time Assessment of Black Carbon Pollution in Indian Households Due to Traditional and Improved Biomass Cookstoves." *Environmental Science & Technology.* 46: 2993-3000.

Lewis, J. and S. Pattanayak. 2012. "Who Adopts Improved Fuels and Cookstoves? A Systematic Review. *Environmental Health Perspectives*, 120: 637-644.

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.

Manibog, F. 1984. "Improved Cooking Stoves in Developing Countries: Problems and Opportunities." *Annual Review of Energy*, 9: 199-227.

Megen Power Ltd. 2008. *Final Report: Impact Assessment of MIRT Improved Biomass Injera Stoves [sic] Commercialization in Tigray, Amhara and Oromiya National Regional States*, Submitted to the MoARD/GTZ SUN Energy Programme, Addis Ababa.

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

Rights and Resources Initiative. 2014. *What Future for Reform? Progress and Slowdown in Forest Tenure Reform Since 2002*. Washington DC.

Saatchi SS, Harris NL, Brown S, Lefsky M, Mitchard TAE, Salas W, Zutta BR, Buermann W, Lewis SL, Hagen S, Petrova S, White L, Silman M and Morel A. 2011. "Benchmark Map of Forest Carbon Stocks in Tropical Regions across Three Continents." *PNAS*, 108(24): 9899-9904.

Skutsch, M and M. K. McCall. 20012. "The Role of Community Forest Management in REDD+," *Unasylva*, 63: 51-56.

Smith, K. R, H. Frumkin, K. Balakrishnan, C. D. Butler, Z. A. Chafe, I. Fairlie, P. Kinney, T. Kjellstrom, D. L. Mauzerall and T. E. McKone. 2013. "Energy and Human Health." *Annual Review of Public Health,* 34**,** 159-188.

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 NationsFramework Convention on Climate Change, Executive Board of the Clean DevelopmentMechanism, Bonn. http://cdm.unfccc.int/DNA/fNRB/index.html.

UN-REDD. 2010. *The UN-REDD Programme Strategy 2010 – 2015*. UN-REDD Programme 5th Policy Board Meeting. Washington, DC, 4-5 November.

van der Werf, G. R.; D. C. Morton; R. S. DeFries; J. G. J. Olivier; P. S. Kasibhatla; R. B. Jackson; G. J. Collatz; and J. T. Randerson. 2009. "CO2 Emissions from Forest Loss," *Nature Geoscience* 2: 737 – 738.

Vieider, F. M., A. Beyene, R. Bluffstone, S. Dissanayake, Z. Gebreegziabher, P. Martinsson, and A. Mekonnen. 2014. "Measuring Risk Preferences in Rural Ethiopia: Risk Tolerance and Exogenous Income Proxies. World Bank Policy Working Paper WPS 7137, December.

Wrangham, R.W., J.H. Jones, G. Laden, D. Pilbeam and N.L. Conklin-Brittain. 1999. "The Raw and the Stolen: Cooking and the Ecology of Human Origins." *Current Anthropology* 40: 567–594.

World Bank. 2009. *Forests Sourcebook: Practical Guide for Sustaining Forests in International Cooperation*. World Bank: Washington, D.C.

Appendix A

Improved Cookstove Randomized Controlled Trial (RCT)

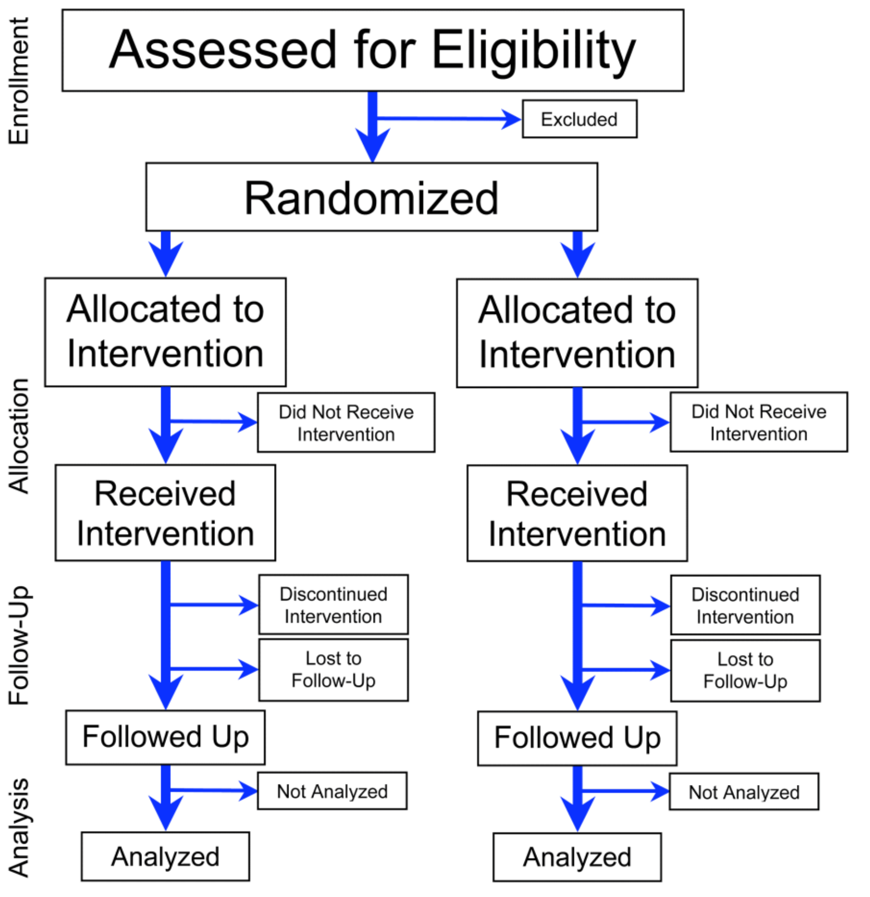
Forest Sector Institutional Reform and REDD+ in Ethiopia: Making Participatory Forest Management Pro-Poor Carbon Sequestration Policy

<u>Key Steps in RCTs</u>

Before beginning the RCT, conduct the household survey, including the two choice experiments (CE) so respondents have a chance to think and learn about improved stoves.

The diagram below lists the main steps in any RCT. In our case, we have several interventions that we have been calling treatments

Source: Wikipedia



Detailed Steps Associated with the RCT

1. Randomly select 36 forest user groups (FUGs)/villages from population of 84 FUGs already identified (the number was 110 but after omitting Tigray sites and also FGD sites it goes down to 84).
2. Randomly apply 6 treatments to villages.
3. Randomly select 10 households from each village, all of whom receive Mirt stoves and treatments.[10] All 10 treated households in each village receive the same treatment.
4. Recruit respondents. Tell respondents that they were chosen randomly to receive a stove under the same terms as others in their village.
5. Give respondents full information on the stove features using material from the improved stove CE preamble (see next page). Respondents would have already completed the improved cookstove CE.
6. Inform respondents that if they agree to participate, their stove use will be monitored and enumerators will come to their houses twice to download data.
7. Show the stove use monitor (SUMS) device and inform respondents that these devices only record their use of the stoves.
8. SUMS will be placed by enumerators and respondents are requested not to touch them. If they are moved, respondents should put them back in the same spot on the stove using the heat resistant tape.
9. Inform respondents that the SUMS 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 SUMS 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.
10. Answer any questions about the SUMS.
11. Present the treatment randomly assigned to the FUG/village (see page 4). Ask if respondent has any questions.
12. Inform participants that stove use is not required to receive a stove, but they are encouraged to use it.
13. Obtain formal oral informed consent. Note on survey form that consent was received.
14. If respondent agrees to treatment, proceed with stove installation. If respondent does not agree, administer non-participant questionnaire, thank them for their time and randomly select a replacement household from the village.
15. Install stove according to technical requirements in the place requested by respondent.
16. Initialize stove use monitors according to instructions if it has not been done already. Be sure to record the installation and start times on questionnaire.

[10] 4 control households are also chosen in each village, but these are mainly related to the choice and behavioral experiments.

17. Train respondents in stove use. Respondents receiving treatments 2, 4, 6 also receive network building and group training (see page 5 for components).
18. Inform respondents that enumerators will return to collect the SUMS data after 6 weeks and again after 12 weeks. Provide respondents with a mobile telephone number to call if they have any concerns.

Stove Features and Performance Information

"The Mirt stove, which you are offered is being promoted by the Government of Ethiopia as part of its National Improved Cookstoves Program. The objective of the program is to support the dissemination and adoption of 9 million improved cookstoves in Ethiopia.

The use of these improved cookstoves is believed to have the following advantages: (i) reduced deforestation from reduced fuelwood consumption; (ii) reduced smoke and reduced respiratory disease infections and deaths; and (iii) possibly less time collecting fuel.

These benefits come from more efficient burning of wood. The Mirt stove has been found to reduce fuel consumption by about half and the release of harmful gases was found to decline almost to zero (decline of about 90%)."

Randomly Assigned Treatments[11]

	Attributes		
Treatment	ETB 25 Payment by Respondent to Receive Stove	ETB 50 Total Incentive for Using the Stove at Least 2 Times Per Week for First 40 Days	Network Building and Group Training
Treatment 1	No	Yes	No
Treatment 2	No	Yes	Yes
Treatment 3	Yes	No	No
Treatment 4	Yes	No	Yes
Treatment 5	No	No	No
Treatment 6	No	No	Yes

Description of Attributes Making up Treatments to Tell Respondents

Enumerators should always mention the payment terms to respondents (i.e. receive free or pay ETB 25).

Please say "We offer a Mirt stove to you for free (at a cost of ETB 25) payable today or at latest tomorrow. This money will be used to further the goals of our research project. I will not keep this money."

[11] Treatments 7 and 8 are eliminated to assure that 6 villages receive each treatment.

The other two attributes should be discussed only if they are active (i.e. incentives for use and/or network building/group training are offered).

- **ETB 50 Incentive for Using the Stove for First 6 weeks**.
  "If you accept the Mirt stove you also have the opportunity to receive an incentive for actually using the stove to cook injera. You will receive a payment of ETB 50 if you use the stove at least 2 times per week for the next 6 weeks as measured by the stove use monitor. That is, if over the next 6 weeks you use the stove 12 times, you will receive a payment of ETB 50. The use each time should be for a period of time appropriate to cook injera for your family."

- **Network Building/Group Training**

  "If you accept the Mirt stove you will also need to attend a meeting lasting no more than 2 hours with all other villagers who are receiving these stoves. The meeting will take place at ____________________(time) and __________________________ (place). During the meeting we will provide additional information about the Mirt stove and answer any questions."

After all information is given, please ask respondents if they understand everything and whether they agree to the terms under the treatment. If respondents agree, proceed with installation, taking payment if appropriate, etc. Otherwise administer the short questionnaire for those who refused to participate.

Stove Use Training Information for Individual Training

From Ministry of Agriculture and Rural Development/GTZ training materials. **Please add and use after stove is installed.**

Network and Group Training Attribute Implementation

For Enumerators:

The purpose of the network and group training attribute is to evaluate whether when promoting such technologies creating a network or "club" of users supports good outcomes like easier adoption. In villages receiving treatments 2, 4 and 6, after all 10 respondents are recruited and stoves and SUMS are installed in all 10 houses in each village, a group meeting should be called. The goals of the meeting, which should last 30 minutes to 2 hours, are the following:

1. Be sure all households receiving Mirt stoves are known to each other.
2. Encourage respondents to interact and discuss issues that come up regarding stoves during the coming months.
3. Provide some additional details on potential benefits other users have received using the 2008 stove program impact assessment or other documents.[12] Information on the climate and other environmental and community benefits of Mirt stoves should be discussed.
4. Answer any questions.
5. Provide some snacks.
6. Create some esprit de corps.

NOTE: During the household survey, for respondents receiving Mirt stoves, please mention the names of the 10 villagers who will receive stoves. Ask each respondent to name up to 3 of their friends from this list. There is a place to record these names in question 4.0.

[12] Please assure that all enumerators give the same information to respondents. Of course, questions and therefore answers will differ.

Appendix B

Random Effects Model of Mirt Stove Use Frequency

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Age of respondent	-0.011	-0.009	-0.005	-0.008	-0.005	-0.005
	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)
Sex of respondent	-0.069	-0.025	-0.223	-0.290	-0.184	-0.205
	(0.362)	(0.362)	(0.378)	(0.380)	(0.355)	(0.361)
Marital status is married	0.507	0.521	0.557	0.507	0.230	0.242
	(0.327)	(0.330)	(0.360)	(0.339)	(0.317)	(0.325)
Education is illiterate	0.081	0.082	-0.135	-0.158	-0.171	-0.175
	(0.206)	(0.207)	(0.206)	(0.202)	(0.196)	(0.199)
Family Size	0.266***	0.250***	0.178**	0.191**	0.161**	0.163**
	(0.087)	(0.087)	(0.085)	(0.086)	(0.077)	(0.078)
Religion is Christian	-0.715***	-0.694***	-0.877***	-0.862***	-0.083	-0.090
	(0.267)	(0.267)	(0.314)	(0.309)	(0.399)	(0.397)
Livestock in TLU	-0.051**	-0.056**	-0.071***	-0.084***	-0.087***	-0.090***
	(0.026)	(0.026)	(0.024)	(0.026)	(0.027)	(0.027)
Children under 15	-0.078	-0.054	-0.025	-0.027	-0.004	-0.002
	(0.087)	(0.086)	(0.082)	(0.079)	(0.074)	(0.075)
Distance to road	-0.002**	-0.002**	-0.002***	-0.002**	-0.002*	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Round2	0.298***	0.293**	0.292**	0.295**	0.303**	0.435***
	(0.113)	(0.114)	(0.118)	(0.118)	(0.118)	(0.167)
Round3	0.869***	0.872***	0.892***	0.890***	0.893***	0.650**
	(0.159)	(0.160)	(0.164)	(0.165)	(0.167)	(0.284)
Round4	0.612***	0.616***	0.637***	0.636***	0.636***	0.509**
	(0.141)	(0.142)	(0.146)	(0.146)	(0.147)	(0.232)
Participation in CCT		0.333	0.267	0.268	0.266	0.265
		(0.209)	(0.200)	(0.191)	(0.173)	(0.173)
Risk Coefficient		0.048	0.112	0.123	0.158**	0.154**
		(0.082)	(0.079)	(0.079)	(0.073)	(0.073)
Average temperature in ^0C			0.107***	0.088**	0.047	0.031
			(0.036)	(0.034)	(0.084)	(0.088)
Number of Injera baked			0.036***	0.038***	0.030**	0.030**
			(0.010)	(0.010)	(0.015)	(0.015)
Stove use for other purpose			0.403	0.477*	0.423	0.423
			(0.263)	(0.250)	(0.308)	(0.314)
Quantity of flour			0.007	0.014	-0.037	-0.037
			(0.054)	(0.054)	(0.055)	(0.055)
Flour kind is teff			-0.210	-0.245	-0.201	-0.186
			(0.212)	(0.210)	(0.262)	(0.266)
Stove is installed inside			-0.378	-0.332	-0.411	-0.408
			(0.240)	(0.249)	(0.339)	(0.343)
Main fuel is fuel wood			-0.389	-0.453	-0.193	-0.172
			(0.255)	(0.278)	(0.253)	(0.254)
treatment_1				0.034	-0.338	-0.817**
				(0.330)	(0.530)	(0.387)
treatment_2				-0.153	0.946***	1.290***
				(0.236)	(0.294)	(0.341)
treatment_3				-0.025	-0.653*	-0.707**
				(0.242)	(0.352)	(0.352)
treatment_4				-0.713***	-0.615	-0.882
				(0.220)	(0.571)	(0.540)
treatment_6				0.263	0.813**	0.903**
				(0.275)	(0.323)	(0.359)
treat1_round2						0.158
						(0.457)

treat2_round2				-0.453*		
				(0.248)		
treat3_round2				-0.350		
				(0.242)		
treat4_round2				0.115		
				(0.358)		
treat6_round2				-0.069		
				(0.275)		
treat1_round3				0.976		
				(0.607)		
treat2_round3				-0.104		
				(0.444)		
treat3_round3				0.233		
				(0.379)		
treat4_round3				0.456		
				(0.443)		
treat6_round3				-0.190		
				(0.483)		
treat1_round4				0.739		
				(0.521)		
treat2_round4				-0.613*		
				(0.350)		
treat3_round4				0.415		
				(0.331)		
treat4_round4				0.299		
				(0.361)		
treat6_round4				-0.090		
				(0.388)		
Constant	1.681***	1.263**	-1.464	-0.784	0.785	1.332
	(0.556)	(0.631)	(1.294)	(1.228)	(2.409)	(2.437)
chi2	50.564	57.456	125.197	136.839	282.583	326.525
N	1078	1066	1045	1045	1045	1045

Note: *** = Significant at 1%, ** = Significant at 5%, * = Significant at 10%, District dummies included in both model 5 and 6 are not reported for the sake of economizing space.