

**Who Adopts Matters:  
Evaluating the Health Impacts  
of Improved Stoves in China  
using Matching Methods**

Valerie Mueller\*  
The Earth Institute at Columbia University

Alexander Pfaff\*  
Duke University

John Peabody  
University of California at San Francisco

Yaping Liu  
First Hospital of Tsinghua Hospital, Beijing, China

Travis Riddell  
Boston Children's Hospital

Kirk R. Smith  
University of California at Berkeley

*Keywords*

indoor air pollution, stoves, health, valuation, propensity score matching

*JEL codes*

I1, Q25, Q42

---

\* Co-lead authors. Mueller is corresponding author: vam2105@columbia.edu;  
B-19 Hogan Hall, 2910 Broadway, New York, NY 10025.

## 1. Introduction

Over 3 billion people in developing countries face health risks associated with biomass use for energy, e.g. the burning of wood, dung, and crop residues (Reddy et al. 1996, World Bank 2003). Trends predict that the number exposed to such risks will increase (WHO 1997, WRI et al. 1998).<sup>1</sup> The World Health Organization found that indoor smoke already accounts for 3.7% of the burden of disease in developing countries, ranking indoor air pollution (IAP) 4<sup>th</sup> among all sources, following malnutrition, unprotected sexual relations, and poor water quality and sanitation (WHO 2005). In other terms, 1.6 million people die annually from IAP. A problem of this magnitude demands accurate evaluation of key interventions. This paper demonstrates both a bias relevant to most existing evaluations and one way to address it.

Smith et al. 2004 emphasizes why indoor air pollution is so important. The combination of high exposure with high pollutant concentration is the key issue. Given indoor concentrations and person-time spent indoors, exposure to air pollution is largely indoors (even if more pollution is emitted outdoors). Biomass fuels are often used in poorly ventilated places with open fires or inefficient stoves, yielding pollutant levels well above living in a dirty city (Smith 1993). Smith et al. 2004 also emphasizes that such exposure is not likely to be uniform across the members of a household. Men will spend more time outdoors, and less time cooking, while on average children will spend more time indoors with women.

The severity of the problem has inspired studies evaluating health risks associated with biomass fuel use. Published evidence suggests that changes in what biomass is used and how could reduce health risks. For instance, Bruce et al. 1998 find fewer respiratory symptoms when women use plancha stoves with enclosed combustion chambers and chimneys instead of open fires, while McCracken et al. 2007 examines impacts of plancha stoves upon blood pressure, an important indicator for cardiovascular disease (Brook et al. 2003, He et al. 1999, Thun et al. 1997). Along these lines one might see also, for example, research by Bruce et al. 2000, Boy et al. 2002, Bruce et al. 2004 and Mishra et al. 2004.<sup>2</sup>

Ezzati and Kammen 2002's review summarizes that studies have shown that biomass and coal smoke contribute to the incidences of ARI, middle ear infection, COPD, lung and other cancers, asthma, TB, low birth weight, eye diseases (e.g., Florig, 1997; Smith et al., 2000; Bruce et al., 2000; Ezzati and Kammen 2001a 2001b), where the main focus has been on ALRI, middle ear infection, and COPD (Smith et al. 2000, Bruce et al. 2000). They note further that, in 2000, 1.5 to 2 million deaths or 3 or 4% of total mortality were attributable to these risks (see Smith and Mehta 2000 and von Schirnding et al. 2002). Such research and such syntheses have inspired projects worldwide to spread and to commercialize emission-reducing stoves (as discussed in, for instance, both Holm-Nielsen 2000 and UNDP 2006).

This paper examines a data set on approximately 3,500 Chinese households, representing a very significant data collection effort at the household level, in particular to include multiple health outcomes. We note that Dasgupta et al. 2006 examine over 30,000 households in Bangladesh to explain air quality, not a health outcome, while Boy et al. 2002 and Mishra et al. 2004 also study thousands of households and both do feature a health outcome, birth weight, albeit a single one in contrast to this China data set.

---

<sup>1</sup> Given recent attention to mitigation of the emissions thought to cause long-run climate change, while we focus on direct local health impacts it is worth noting that biomass emissions become outdoor and atmospheric pollution too. The IPCC's 4<sup>th</sup> Assessment Report, in Chapter 8 "Human Health" of Working Group II's work on impacts, adaptation and vulnerability, notes that household use of biomass and coal for energy on a wide scale contributes to global carbon emissions. FAO 2005 (and FAO Newsroom December 9, 2005 for instance) suggests that the 25 percent of carbon dioxide emissions caused by deforestation may increase with rising global energy demand.

<sup>2</sup> This list is incomplete and focuses on developing countries, given our focus on China. Other actions may well be recommended in developed settings. For instance, Moshhammer et al. 2006 suggest good room ventilation following their analysis of the impacts on children's lung function of cooking with natural gas (a clean fuel in our analyses!).

Given our focus on China, we want to highlight in particular other analyses of Chinese data. Mumford et al. 1987 and 1989 show associations between lung cancer rates and the percentages of households using smoky coal in communes within Xuan Wei province. Luo et al. 1996 compare data on lung cancer cases from Fuzhou with a randomly stratified control sample to suggest that burning coal indoors is a determinant of lung cancer. In addition, Lan et al. 2002 examine the impacts of chimneys on lung cancer rates. Finally, we highlight Peabody et al. 2005, which analyzes the same data studied here.

Peabody et al. 2005 compare other stoves to coal, finding significant effects of improved biomass stoves in reducing respiratory disease, COPD and exhaled CO, and in increasing FVC. Improved biomass also does better than traditional biomass in terms of respiratory disease, COPD, exhaled CO, and FVC. The history of asthma was not found to be a significant determinant in either of those stove comparisons.

We improve upon prior work by addressing the non-random distribution of cleaner stoves, a fact that can arise when households' characteristics, in any of various ways, can affect who uses what stoves.<sup>3</sup> As will be seen below, we will also examine different health outcomes from Peabody et al. 2005 and in different comparisons of stove types. However, our main point is a method to address estimation biases.

Much of the existing literature supporting such stove interventions is subject to selection biases (Heckman and Smith 1995). The reason is that which stove an individual uses may be correlated with her health outcomes for reasons other than the impact of the stove on health. Any such correlation confounds the accurate estimation of the cleaner stove's causal impact on health. For example, cleaner stoves may be more likely to be adopted by, e.g., those with poor housing ventilation or those with generally poor health already. In each case, non-linear responses (of air quality to emissions and/or of health status to exposure) may imply that those groups' marginal benefits of lowered stove emissions are greater. If that is so, and if in addition the characteristics of poor ventilation<sup>4</sup> and poor prior health are otherwise associated with worse measured health outcomes, then it is easy for one to underestimate the cleaner stove's benefits.<sup>5</sup>

Our analysis of this large cross-sectional survey demonstrates how health impact estimates of stove interventions can be biased in that way. We show that the groups who own the different stoves differ in terms of characteristics which themselves matter for health outcomes. That creates biases.

To address them we use 'matching' techniques, specifically both propensity score matching (PSM) and covariate matching (CM) approaches, to estimate the effects on health outcomes of moving from traditional biomass (TB) and from coal stoves to improved biomass (IB), and clean-fuel stoves. The idea, as is well known in many policy evaluation settings, is to 'compare apples to apples' by making sure to compare the outcomes for those with cleaner stoves to *otherwise similar* households with other stoves.

Matching approaches are used in various settings to identify the effect of policy programs. For instance, job training (Dehejia and Wahba, 1999; Heckman et al., 1997) and health (Hill et al. 2003) and forest conservation (Andam et al. 2008; Pfaff et al. 2008a 2008b) have all been studied. We use matching to estimate the effect of cleaner stoves on the number of respiratory symptoms reported and the incidence of specific symptoms (exactly which symptoms is discussed below<sup>6</sup>) and who benefits most from stoves.<sup>7</sup>

---

<sup>3</sup> Others have started to implement another approach to this problem for evaluation of stove interventions, that being to randomize stove distribution (Kirk Smith in Guatemala and lately economists working in India and Bangladesh). Theirs are a tiny fraction not only of current studies but also, we believe, even now, of planned analyses. It may also be worth noting that such work corrects without demonstrate the magnitude of the bias from non-random allocation. Finally, as is clear in our model, household choice can still confound evaluation with randomized stove allocation.

<sup>4</sup> Bruce et al. 1998 for Guatemala and Dasgupta et al. 2006 for Bangladesh link dwelling characteristics with health.

<sup>5</sup> A bias in the opposite direction is also possible (see Pitt et al. 2006 and the discussion within Section 2 below).

<sup>6</sup> Previous studies have drawn upon the association between a number of similar symptoms and outdoor air pollution to measure the benefits of reducing outdoor air pollution (Alberini and Krupnick, 1998; Alberini et al., 1997).

We find that clean-fuel and IB stoves reduce the number of symptoms relative to, respectively, the IB and the TB stoves. In looking at specific symptoms underlying that aggregate outcome, we find that for most stove comparisons the gains are largely in terms of headaches and dizziness. Further, for both the number of symptoms and specific symptoms, we show that correcting for non-random stove distribution matters. Our matching estimates increase the estimated benefits of clean-fuel stoves relative to TB stoves. Usefully, they also eliminate a spurious result that IB stoves are worse for health than coal stoves, which arises because those using the improved biomass (IB) stoves are more prone to symptoms. Matching also simply confirms many estimates in cases where who uses what stove does not create bias. Finally we also find that, across all of these results, women are the main beneficiaries of cleaner stoves.

## **2. Background & Model**

### **2.1 Stove Distribution In China Over Time And Space**

In the early 1980s, the Chinese government funded a National Improved Stove Program (or NISP) to provide more efficient biomass stoves and improved coal stoves (see Sinton et al. 2004). The Ministry of Agriculture was responsible for this program, which included direct subsidies to households where the fraction of cost subsidized depended on stove and county. NISP support to 860 counties total (of 2,126 countrywide). Some counties took their own initiative in promoting improved stoves. Some agencies would develop the stoves while the households would pay for the materials and installation.

In NISP's first phase, which lasted until 1992, the program subsidized rapid dissemination. Counties applied to participate and were chosen based on criteria including energy shortages and energy dependence as well as the local willingness to share the cost burden. The latter criterion led the program to wealthier areas (Sinton et al. 2004, p. 37).

In NISP's second phase, which occurred during 1990 – 1995, the distribution of improved biomass stoves no longer counted on household subsidies. However, the rural energy industries were given tax and loan benefits, training, and administrative support. In addition, the Ministry of Health started a program in the 1990s to improve kitchens in poor regions, targeting fluorosis.

NISP's third phase was after 1995 and included a response to the flooding of the Yangtze River in 1998 by the former State Development Planning Commission. The Yangtze River Valley Environmental Protection Project was intended to reduce soil erosion by supporting reforestation, and included in addition the promotion of improved stoves in order to reduce fuelwood demand. There was also a promotion of rural coal markets to convert previous biomass users to coal use.

With all of this stove guidance from above, nonetheless households' choices concerning whether or not to adopt a cleaner stove appeared to remain paramount. Sinton et al. 2004 conveys, for instance, that: "Unlike many improved stove programs in other countries (such as India), households bore most of the direct costs of stove purchases." (p.39); and also "Households paid about 94% of all costs" (p. 40).

### **2.2 The Implications Of A Stove Being A Choice**

#### *2.2.1 Household Choice Model*

---

<sup>7</sup> We have tried to explain blood pressure but believe this is an exceptionally noisy health measure. The associations of our two measures (diastolic and systolic) with possible covariates often are counterintuitive in terms of their sign (and in a way not explained by who uses what stove) and further are often reversed in sign across the two measures.

We provide a simple one-period model that offers one perspective upon who ends up with cleaner stoves.<sup>8</sup> An individual maximizes her utility, which is a function of her health, subject to a budget and a time constraint. Health,  $h(d,b)$ , responds to effective exposure to airborne pollutants  $d$  and medical intervention  $b$ . Effective exposure to pollutants  $d(s_q,s_t)$  depends on  $s_q$ , the quantity of clean air which the individual purchases, and  $s_t$ , the amount of time spent obtaining clean air, e.g. by relocating oneself relative to a source of pollutants. Having a distinction between the  $s_q$  and  $s_t$  allows the cost of obtaining clean air to have components with different prices, here those being literally the market price of clean air and the opportunity cost of time. Following the idea of such a distinction, we will re-interpret each of these cost components below. We can write health  $h(s_q,s_t,b)$ .

Clean air may require both monetary expenditure and an investment of time. In our context, the purchase of clean air may be thought of as expenditure on a cleaner stove that is installed and maintained (for a household with multiple stoves, if the fuel for the cleaner stove is more expensive this could also be thought of expenditures on the cleaner fuel). The opportunity cost of avoiding the pollutants in order to breathe cleaner air may literally be time incurred in avoidance. It could also represent other non-monetary costs of exposure avoidance such as suboptimal food preparation on a stove left unattended for periods of time so that the person doing the cooking could spend time away from the most direct emission of pollutants from the stove. While we do not present a model of the intra-household allocation of time to cooking (see, e.g., Pitt et al. 2006), we would expect higher exposure for those who cook, which in our model implies an inability to avoid pollutants.

The individual's utility maximization problem with self-protection choices is the following:

$$\begin{aligned} & \max_{s_q, s_t, b, f, X} U(X, f, h) \\ \text{s.t.} \quad & I + p_w(T - f - h - s_t) \geq X + p_{s_q} s_q + p_b b \end{aligned}$$

Here  $X$  is a composite private good that is purchased by the individual, while  $f$  is leisure time,  $T$  is the total amount of time available,  $I$  is the level of income independent of the time spent working,  $p_w$  is the wage,  $p_{s_q}$  is the price of a cleaner stove or cleaner fuel and finally  $p_b$  is the price of medical services.

Note that the expression  $(T - f - h - s_t)$  represents the total amount of time spent earning income.

The first-order conditions for clean air purchases and the time spent obtaining clean air are

$$\frac{\partial U}{\partial h} - \lambda p_w = \frac{\lambda p_{s_q}}{\partial h / \partial s_q} \quad (1)$$

$$\frac{\partial U}{\partial h} - \lambda p_w = \frac{\lambda p_w}{\partial h / \partial s_t} \quad (2)$$

where  $\lambda$  is the Lagrange multiplier on the budget constraint.

This optimization yields the demand function for  $s_q$ ,  $s_q^*(I, p_w, p_{s_q}, p_b)$ , and analogously for  $s_t$ . Empirically, we will consider the impact of the use of cleaner stoves on health outcomes. While relevant details of the Chinese stove program are numerous, that household preference and choice were critical in the allocation of stoves is the key issue for this model. That only a subset of the households will prefer to invest in the cleaner stoves is also the critical issue motivating the empirical approach taken in this paper.

---

<sup>8</sup> The model is essentially the one-period version of Grossman's (1972) model presented in Freeman (1993).

## 2.2.2 Implications For Stove Impact Evaluation

The model of clean-stove adoption above describes the first stage of the process we wish to analyze. It shows that as the parameters of the model vary, so will stove choice. That is a key point missing in the stove-evaluation literature, which generally assumes stoves are distributed randomly.

Individuals vary in many dimensions relevant for stove choice. Some of these were shown above. For instance, households facing higher prices for cleaner stoves (and/or fuels) will be less likely to adopt. Further, other characteristics (denoted  $\theta$ ) could affect the value of a cleaner stove. One example is poor general health. This might raise the net benefits of a cleaner stove if the dose-response function  $h(d)$  were more in fact  $h(d, \theta)$  and vulnerability to  $d$  rose with poor health (an element of  $\theta$ ). On the other hand, in the spirit of Pitt et al. 2006 where poor health might lower agricultural productivity and thus lower the cost of exposing some household members to stove emissions, perhaps poor health could lead to lower net benefits of stove adoption. Certainly there must be multiple possible priors to explore empirically.

Consider the implications of such  $\theta$  for the relationship that we wish to estimate. We examine the impact of stoves on health, studying  $h(s_q, s_b, b)$ . We do not have information about medical interventions. We focus upon the effect of  $s_q$ , the expenditures on clean air, which empirically will be represented by the possession and the use of a cleaner stove. Having done this, we also examine the effect on stoves' impacts of gender, a proxy for  $s_i$  in that women cook more and thus are less able to avoid pollutants.

To spell this out further, in order to motivate our empirical approach, we will examine each of the stove types' impacts upon our measured health outcomes by regressing the health outcomes  $h$  on whether the respondent has a cleaner stove (in fact we will compare outcomes across stove-types pairs, given that there are multiple stove options that are difficult to compare clearly all at once). For instance:

$$h_i = s_{qi}\alpha + s_{ti}\beta + s_{qi}s_{ti}\gamma + b_i\sigma + \varphi_i\delta + \varepsilon_i \quad (3)$$

The key issue here is that  $\theta$ , e.g. poor general health, affects the measured health outcome  $h$ . That need not be the case, yet it might be and several observable characteristics should be empirically examined. We also note at this point the error  $\varepsilon$ , as there are of course unobservable dimensions that also affect  $h$ .

The motivation for our approach becomes apparent when the  $s_q$ \* function is paired with (3):

$$s_{qi} = I_i\zeta + p_{wi}\tau + p_{sqi}\nu + p_{bi}\mu + \varphi_i\zeta + z_i\psi + \eta_i \quad (4)$$

Focusing on  $h$ ,  $s_q$  and  $\theta$  illustrates the challenges for stove evaluation. That  $\theta$  has a role both within (3), in explaining stove adoption, and within (4), in health, can bias estimates of impacts of stoves  $s_q$  on health  $h$ . If  $\theta$  is poor general health, for the moment assume that higher  $\theta$  both shifts dose-response functions so net benefits of stove adoption rise, i.e.  $\zeta > 0$ , and increases  $h$ , i.e. worsens health (recalling that  $h$  is lost work) such that  $\delta > 0$ .<sup>9</sup> Thus high  $\theta$ , i.e. poor general health, makes one both more likely to suffer ill health  $h$  and more likely to have a clean stove  $s_q$ . This need not but can bias estimated impacts of cleaner stoves, e.g. up towards zero if  $\alpha < 0$ , and even yield a spurious estimate of stove damage (i.e., estimated  $\alpha > 0$ ).

If we knew of variables  $z$  in (4) but not in (3), then we could estimate  $\alpha$  correctly anyway (the  $z$  could include exogenous distribution following a program design such as randomization). Alternatively, as discussed in Section 4 on methods, we will try to identify the observed variables that play the role of  $\theta$  above, i.e. that affect both stove adoption and health, and use matching to control for their influences.

---

<sup>9</sup> The errors in (3) and (4),  $\varepsilon_i$  and  $\eta_i$ , may well be correlated. While using instruments, such  $z$  variables in (4) above, could address that issue, the matching approach that we employ does not eliminate this potential source of bias.

### 3. Data

#### 3.1 Household Survey

We use a cross-sectional survey of approximately 3,500 households in three provinces of China (Shaanxi, Hubei, and Zhejiang). The data were collected during the years 2001 - 2003 to help to evaluate China's National Improved Stoves Program (NISP -- see Sinton et al. 2004 for a detailed description of the survey). The survey includes information on health outcomes for adults (ages 18 and over), household demographics, fuel use, and use of stoves by type and characteristics. We focus on the health impact of stoves used primarily for cooking meals, assuming cooking stoves provide the greatest amount of exposure to air pollutants. There are four types of cooking stoves used in our sample: traditional biomass (16%), improved biomass (48%), coal (30%), and clean-fuel (6%). Stoves are defined by the type of fuel they use. Traditional biomass (TB) and improved biomass (IB) stoves both use wood, crop residues, or dung. Improved biomass stoves have the additional structural feature of having at least a flue and a grate. Coal stoves use coal, coke, or lignite. Clean-fuel stoves include electric, liquefied petroleum gas, or biogas. Peabody et al. 2005 has ranked the stoves in terms of association with worst health outcomes in the following fashion: coal (worst), traditional biomass, improved biomass, and clean-fuel (best).

#### 3.2 Significant & Meaningful Differences Among Groups Owning Each Stove

For our matching approach to stove-impact estimation to make any difference, it would have to be the case that the groups who use each of the stove types differ from one another. In particular, they would have to differ along dimensions that are themselves significant for our health outcomes.

In the following subsection 3.3, we will examine regressions that explain our health outcomes. Here we examine first the group differences in terms of some variables that matters in those regressions. Specifically, we have tested for statistical differences in the average covariate values across stoves pairs. Tables 1 - 5 display the percentage differences in stove-group averages, with t-statistics to test difference, for some variables that are relevant for our "number of symptoms" health outcome. The groups do differ.

To show that these differences are not only statistically significant but also meaningful for health, we have placed within the same tables the marginal effects of each variable on the health outcome, from regressing "number of symptoms" on a stove treatment dummy variable and the covariates (3.3 below).

Combining the sign of the difference in the covariate average across groups with the sign of the impact of that covariate on the health outcomes in question indicates the direction in which the difference in the groups should bias the estimated effect of the cleaner stove relative to the other stove in the pairing. For example, from Table 1, improved biomass (IB) users are more likely to be in bad health and, further, we see that to be in bad health correlates with reporting more symptoms. Thus, this difference between the IB- and the TB-using groups will bias downward estimates of the reduction in symptoms due to IB.

#### 3.3 Regressions For Health Outcomes

While quite imperfectly to this point, we appeal to the public health literature in choosing which covariates reflect the effect of emissions' exposure upon health outcomes. For example, Dasgupta et al. 2006 find that ventilation, which is promoted by wall and roof materials, openings in the kitchen, and possession of open air or detached kitchens, affects individual exposure. We account for the impact of housing ventilation on our health outcomes by including the number of kitchen openings, and whether the dwelling has an additional open air kitchen. Also, Ezzati et al. (2000), and Ezzati and Kammen (2001a, 2001b) suggest the time in the particular microenvironment proximate to cooking changes exposure. We include the number of minutes an individual spends cooking in a given day the week prior to the survey.

In addition, we note that age, gender, wealth, and smoking status are often incorporated in the aforementioned studies to explain average levels of health outcomes. We include age categorical variables (26-40, 41-55, >55 years old), a male dummy, whether the household income is more than 12,000 Yuan, whether the house has a washing machine, whether the individual self-diagnoses himself as having fair or poor health, whether the individual is a smoker, and whether the individual smoked more than 40 years.<sup>10</sup>

Table 6 presents the ordinary least squares (OLS) results for the number of symptoms and for some specific symptoms that we also examine stoves' impacts on below. We exclude stove treatment dummies here. Many of the variables have the anticipated effect on the number of symptoms. Age has an increasing, monotonic effect, wealth and possession of a washing machine have negative effects, cooking exposure has a small positive effect, and bad health and whether the individual is a smoker have positive effects (though smoker is insignificant) on the number of symptoms that are reported by the individual.

The number of kitchen openings and the possession of additional open air kitchen variables both positively affect the number of symptoms, however, which contradicts our expectations given prior work. All of the covariates with the exception of an age categorical variable, gender, smoker, and a province dummy variable significantly impact the number of symptoms outcome at the five percent critical level. The rather low adjusted-R square of 0.11, however, suggests that our model could greatly be improved.

## 4. Methodology

### 4.1 Matching With Multiple Treatments

The importance of accounting for households non-randomly selecting which stove to use when measuring the stove effect on health can be demonstrated in the single treatment setting. Consider the bivariate regression of a health outcome  $H$  on a clean stove treatment  $S$ :

$$H_i = \beta_0 + \beta_S S_i + \varepsilon_i. \quad (5)$$

Suppose individuals with increasing levels of some characteristic tend to use the stove, e.g., those with poor health increase their frequency of clean stove use because their marginal benefits from clean stove use are greater than the alternative stove use. This suggests that the treatment variable  $S$  and the error  $\varepsilon$  are correlated, i.e.  $E(\varepsilon_i|S_i) \neq 0$ . The estimate of the stove impact will be biased *and* inconsistent, because the stove treatment variable reflects the decision to use a particular stove which is based on individual behavior. In practice, the bias direction will depend on how multiple variables influence the decision to use a particular stove and the health outcome. The impact of the bias on the treatment parameter estimate can be generalized by taking the probability limit of the OLS estimate of  $\beta_S$ :

$$p \lim \hat{\beta}_S = \beta_S + \Gamma_{SX}' \beta, \quad (6)$$

where  $\Gamma_{SX}$  is a vector of coefficients from bivariate regressions of each element of  $X_i$  on  $S$ , and  $\beta$  is a vector of coefficients of the omitted variables in (1) (Angrist and Krueger, 1999). Applying (6) to our example, we would expect  $\Gamma_{SX} > 0$  and  $\beta < 0$  since the omitted variable, poor health, is positively correlated with clean stove use and negatively correlated with a good health outcome. Thus, our estimate of  $\beta_S$  will be an underestimate of the effect of clean stoves on health.

---

<sup>10</sup> For blood pressure, we also consider variables that are believed to affect blood pressure, such as the number of days the individual had a drink containing alcohol in the past four weeks, body mass index ( $\text{kg}/\text{cm}^2$ ), and whether the individual eats vegetables three or more times a day, representing dietary fiber intake (Whelton et al., 2005). However, as noted above we have focused on symptoms (overall and specific) given the noise in blood pressure.

There are several econometric methods that can be used to correct estimates for selection bias (see Angrist and Krueger (1999) for a discussion). Matching techniques are particularly attractive for cross-sectional analyses and can be used to both understand and take into account the non-random choice of cleaner stoves (see Morgan and Harding 2006 for a review). We apply two matching techniques in the analysis, Propensity Score Matching (PSM) and Covariate Matching (CM, Imbens 2004).

PSM (and CM) techniques exploit the selection-on-observables notion that allows for conditioning on a set of observables to eliminate the selection bias. The key identifying assumption of using these techniques is the conditional independence assumption:

$$H_i \equiv H_i(S_i) = \begin{cases} H_i(0) & \text{if } S_i = 0 \\ H_i(1) & \text{if } S_i = 1 \end{cases}, \quad (7)$$

$$(H(0), H(1)) \perp S \mid p(X)$$

where  $p(X)$  is the propensity score (or distance from the covariates for CM) that is independent of the outcomes and conditional on observables  $X$ .

For the single treatment case, PSM involves estimating probit regression with the outcome “being treated”. The regression includes observed variables  $X$  that simultaneously influence the outcome and the treatment decision. The regression yields a predicted probability of use for each individual. These probabilities, or propensity scores, are used to match the “treated” individuals, those with cleaner stoves, with ‘similar’ individuals for the purpose of the comparison which estimates the stove effect. Here ‘similar’ means a similar probability of treatment. When those without cleaner stoves are matched to similar individuals with cleaner stoves, then outcomes are compared. Ideally, this removes from the estimate the influences of the unobserved factors that affected the stove use.

Our study involves evaluating the impact of multiple treatments on health outcomes: traditional biomass (TB), improved biomass (IB), coal, and clean-fuel stoves. Imbens (2000), Lechner (2001), and Lechner (2002) demonstrate how matching can be used to evaluate multiple treatments. We adopt the PSM methodology proposed in Lechner (2002) to estimate the impact of cooking stoves on health outcomes. Lechner proposes a few ways of addressing this complexity in the form of multinomial logit, multinomial probit, and a series of binomial models. We opt for the latter application. Although estimating several binomial models can be more cumbersome, the assumptions for these models are less restrictive than the multinomial logit. The multinomial probit overcomes some of these restrictions, but requires alternative restrictions to estimate the covariance matrix. The binomial models are more computationally tractable and less sensitive to specification in comparison to the multinomial probit model (Lechner, 2002). Lechner shows all three approaches yield similar results in an empirical application.

To describe our approach, we first define stove use by variable  $S \in \{TB, IB, Coal, Clean\}$  and the health outcome  $\{H^{TB}, H^{IB}, H^{Coal}, H^{Clean}\}$ , where only one cooking stove is observed for each individual. For the multi-treatment case, we make pairwise comparisons of the effects of cooking stove  $s$  and  $r$  for users of stove  $s$ :

$$E(H^s \mid S = s) - E(H^r \mid S = s). \quad (8)$$

Lechner provides four steps for computing (8), which we will briefly describe for our own context. First, we estimate the probabilities of stove use including the subsample of users of stoves  $s$  and  $r$  using a probit regression. Since we are interested in estimating the impact of clean stoves, we restrict the number of

pairwise comparisons to those that compare IB and clean-fuel stoves to all other stove alternatives. Specifically, we estimate five probit regressions for the following pairwise comparisons where treatment is listed first and followed by the control group: i) TB/IB, ii) Coal/IB, iii) Clean-fuel/Coal, iv) Clean-fuel/TB, and v) Clean-fuel/IB. Note that the first two comparisons assume the dirty stove as the treatment, and the last three comparisons assume the clean stove as the treatment. To improve the performance of matching, we had to assign the dirty stove users as the treatment groups in comparisons i) and ii) because the samples of clean stove users were larger and thus more appropriate as control groups. For ease of interpretation, we adjust the sign of the treatment effects of stove comparisons i) and ii) in our tables of statistical results.

Next, the five probit regressions yield propensity scores that are used to match individuals in each treatment group  $s$  with similar individuals in each control group  $r$ . We employ the nearest neighbor algorithm to match a single observation in the treatment group to a similar single observation in the control group (and then five observations in the control group in a subsequent analysis for robustness). Once the matches are established, we compute the expectations of a given health outcome variable conditional on the propensity scores for stove group  $s$  and  $r$  for each five pairwise comparisons:

$$\frac{1}{n^s} \left[ \sum_i H_i^s - \sum_j w_{i,j} H_j^r \right], \quad (9)$$

where  $n^s$  is the number of people using clean stoves, or the treatment group, and  $w$  represents the number of times an observation in the control group has been matched. The variance of the treatment effect (4) is calculated using the expression below:

$$\frac{1}{n^s} \text{Var}(H | S = s) + \sum_{i \in r} \frac{w_i^2}{(n^s)^2} \text{Var}(H | S = r). \quad (10)$$

We estimate (5) and (6) to measure the impact of clean stoves on health outcome for five pairwise comparisons of stove users using the *psmatch2* procedure in Stata (Leuven and Sianesi, 2003).

We apply CM in order to provide an alternative estimate of the average treatment effect on the treated (8) for purpose of comparison. The indicator used to match treated individuals to control individuals is formulated from a distance metric based on the norm of the observables (Imbens, 2004):

$$(X_s - X_r)' \text{diag}(\Sigma_X^{-1})(X_s - X_r), \quad (11)$$

where  $X_s$  and  $X_r$  denote the vectors of covariates for a single treatment and control observation, and  $\Sigma$  signifies the covariance matrix of the covariates. Expression (11) is computed for a given treatment observation with the covariates of every control observation in the subsample, and the control with the smallest value, or fewest dissimilarities conditional on the covariates, is matched with that particular treatment for the single nearest neighbor matching case. We use Stata procedure *nnmatch* to perform CM using the method of nearest neighbor to match treatments with one and five control observations for robustness (Abadie and Imbens 2004). The procedure computes (9) and (10) for the health outcomes and various stove pairwise comparisons, after having matched treated observations with the appropriate control observations compute the effect of various stoves on health outcomes.

## 4.2 Robustness Checks

As mentioned, we use both PSM and CM approaches (and specify two different numbers of matches when using the nearest neighbor algorithm) to obtain the estimates of the cooking stove effects on health outcomes. We also perform additional robustness checks using the PSM approach to understand how sensitive these estimates are to additional bias and matching performance.<sup>11</sup>

Matching estimates can require further bias-adjustment when the matching is not exact, e.g., incorporating a large number of covariates can interfere with the performance of matching (Abadie and Imbens, 2002). We adopt the method proposed in Hill and Reiter (2006) to correct for such bias. We estimate the following regression for each subsample of stove pairs:

$$H_i = \beta_0 + \beta_S S_i + \beta_X' X + \varepsilon_i, \quad (12)$$

where  $S_i$  represents the treatment dummy variable, and  $X$  refers to the vector of covariates used in matching. The regression uses the sample of treated observations and unduplicated matched control observations. We estimate a weighted regression to place more importance on the controls that have repeated use in the matching procedure. Thus, the regression is weighted based on the normalized number of times a single control is used to match with observations in the treatment group, or simply the inverse of the number of observations in the regression for each observation in the treatment group. The estimate of parameter  $\beta_S$  in (8) is interpreted as the matched bias-adjusted estimate.

We check for covariate balancing and impose two additional restrictions to understand how matching performance affects our stove impact estimates. Covariance balancing occurs when the matched untreated group is more similar to the treated group than the group of all untreated observations. In practice, this involves confirming that the differences in the covariates' averages of the matched untreated and treated groups are statistically different (based on t-statistics). Imposing a caliper is one way of reducing the tendency for poor matches. The caliper restricts matches to some maximum distance between the propensity scores. As a robustness check, we apply a rather stringent caliper in one PSM specification, caliper is equal to 0.001. Our third check on matching performance is to impose a common support restriction in one of our PSM specifications. Some treatment cases will have no counterpart among the controls. By restricting the analysis to common support, we are restricting the analysis to the treatment cases whose scores fall between the minimum and maximum propensity scores in the control.

## 5. Results

In this section, we summarize the results from the model specifications that match each treated observations with four similar control observations ( $m=4$ ),<sup>12</sup> when using the nearest neighbor (NN) algorithm, and weighing all controls, when using the Kernel Gaussian (KG) algorithm. We also report the PSM estimates imposing caliper and common support restrictions to check on the performance of matching, and calculate bias-adjusted matching estimates to correct for consequences of covariate imbalancing. All of these matching estimates are compared to the “naïve” and “naïve bias-adjusted” estimates which measure the differences in the health outcome averages between the treated and control groups, and the differences after conditioning for other covariates using OLS regressions.

---

<sup>11</sup> We do not perform these similar checks in CM because they are not offered in the `nnmatch` procedure in Stata.

<sup>12</sup> In determining the number of matches when applying the method of NN, we evaluated how the matching estimates performed assuming the number of matches between 1 and 20 for each outcome. We chose the lowest number of matches around which the effect stabilized in order to diminish the bias from selecting poor matches. We also compared the covariate balancing for estimates that used 4 and 5 matches. There was very little difference.

## 5.1 Number of Symptoms

Our matching estimates for the effect of improved stoves upon the number of symptoms reported are presented in Table 7. The covariate balancing tests indicate that without imposing some additional restrictions on the PSM approach, the Clean-fuel/TB and Clean-fuel/IB matching quality was the highest (where 15 out of 15 of the covariates were balanced at the one percent critical level) followed by IB/TB, IB/Coal, and Clean-fuel/Coal (where respectively 12, 12, and 9 of the covariates were that well balanced).

In Table 7 we see the matching approach affects the estimate of the clean stove's impact for the IB/Coal and Clean-fuel/TB stove pairings. For the other three comparisons (i.e., Clean-fuel/IB, IB/TB, and Clean-fuel/Coal), matching confirms the naïve-bias adjusted value. In the latter pairwise comparison, the matching results confirm that there is no health impact of the cleaner stove. Note the performance of matching was also the weakest for this stove pairwise comparison, and there were some results suggesting that clean-fuel stoves relative to coal do generate benefits, although those results are not robust. In the former two comparisons, the matching validates the naïve finding of a benefit from cleaner stoves.

Matching also can remove the influence of group difference in order to eliminate spurious results. For the IB/Coal comparison, our matching approach eliminates a spurious result that IB stoves increase symptoms relative to Coal stoves. Table 5 suggests that this spurious outcome may come from IB users being older and having a greater tendency to live in Province 2. For the Clean-fuel/TB comparison, the matching approach increases the magnitude of the clean stove estimate of benefits. The exact increase in the benefits' estimate depends, not too surprisingly, upon the exact method. Drawing on the comparison of characteristic averages between clean-fuel and TB stove users (Table 4), the estimate of clean stove benefits that ignores selection bias may be driven by the tendency of clean-fuel stove users to have less ventilation in the form of reduced kitchen openings or the absence of an additional open air kitchen.

Our results support the claims that there are health benefits from using IB and clean-fuel stoves. Specifically, we find benefits in terms of reduced number of symptoms from using IB (in comparison to TB) and from using clean-fuel stoves (versus TB and IB) even with a common support restriction.<sup>13</sup> Using IB stoves (relative to TB) reduces the number of symptoms in a range of 20 to 27 percent of the average number of symptoms (0.95). Using clean-fuel stoves (relative to IB and TB stoves) reduces the number of symptoms in a range of 19 to 28 percent and 36 to 79 percent of the average number of symptoms.

## 5.2 Specific Symptoms

We calculate all of our estimates of the effects of improved cooking stoves upon specific symptoms using the matching approach as well. The results for each stove pairing are included within Tables 8 - 12. We summarize the stove effects in the table below in terms of the sign of any significant effect (benefit vs. detriment) and the impact of matching on the estimated stove effects (thus, EB = enhances benefit if the naïve showed a benefit and matching raised it, CB = confirms benefit seen in naïve, benefit disappears = BD and detriment disappears = DD are self-explanatory, and GB = generates benefit for the cases in which the naïve shows no effect but the matching finds a benefit). All the “---” cells indicate no effect.

For every stove pairing, matching technique affected the stove effect for at least one symptom. For all stove pairings except Clean-fuel/Coal, we observe benefits in reduced headaches and dizziness. This suggests that improving stoves has a robust effect on these symptoms whether the improvement is in the form of enclosing the combustion chamber, adding chimney, or using cleaner fuel. This effect does not show up for Clean-fuel versus Coal though we see an effect when comparing coal stoves to IB stoves.

---

<sup>13</sup> For the Clean-fuel/IB stove pairing, this effect persisted only when using the KG method.

|                 | Headache<br>> 4 hours   | > 5 headaches<br>in a week          | Nausea<br>> 8 hours                        | Nausea<br>Vomiting | Dizzy                             |  |
|-----------------|-------------------------|-------------------------------------|--|--------------------|-----------------------------------|--|
| IB/TB           | ---                     | EB                                  | ---  | EB                 | EB                                |  |
| IB/Coal         | ---                     | GB                                  | ---  | ---                | CD                                |  |
| Clean-fuel/Coal | ---                     | ---                                 | ---  | GB                 | ---                               |  |
| Clean-fuel/TB   | ---                     | CB                                  | EB   | ---                | EB                                |  |
| Clean-fuel/IB   | GB                      | ---                                 | ---  | ---                | CB                                |  |
|                 | Cough<br>With<br>Phlegm | Both eyes<br>irritated<br>> 4 hours | Both eyes<br>irritated > 5<br>times a week | Both eyes<br>Itch  | History<br>of eye(s)<br>allergies | Eyes stick<br>together from<br>discharge |
| IB/TB           | ---                     | CB                                  | BD   | ---                | ---                               | ---                                      |
| IB/Coal         | ---                     | ---                                 | ---  | DD                 | CD                                | CD                                       |
| Clean-fuel/Coal | ---                     | ---                                 | ---  | ---                | ---                               | ---                                      |
| Clean-fuel/TB   | ---                     | ---                                 | ---  | EB                 | EB                                | EB                                       |
| Clean-fuel/IB   | ---                     | ---                                 | ---  | ---                | CB                                | GB                                       |

### 5.3 Women Are The Main Beneficiaries

We evaluate the impact of improved stoves by gender. Ezzati and Kammen (2001a, 2001b), among others, emphasize the importance of time spent in a particular microenvironment. Proximity to stove and time spent home child rearing (in other words, time spent in an enclosed area with high indoor pollutants) will increase respiratory illness. Such time will depend on how tasks are allocated within a household (as noted above, see Pitt et al. 2006 for a relevant model focused precisely on intra-household allocations). Our data suggest (Table 13) that tasks may be allocated differently to men and women. In our sample, women spend slightly more time cooking (though perhaps less, relatively, than one might have expected).

Women and men also may have different dose-response functions, due to differences in behaviors or simply inherent differences in health risks. From Table 13, the evidence also supports that men engage in less healthy behavior than women in terms of smoking and alcohol consumption (although it is quite possible that the number of women that truly smoke could be understated within these data).

Table 14 reports the effects of improved cooking stoves on the number of symptoms by gender. With the exception of the effect of clean-fuel stoves relative to coal, improved stoves affect women more. They are the sole beneficiaries of moving from TB to IB or clean-fuel stoves, or IB to clean-fuel stoves. We note also that the importance of accounting for non-random stove choice in estimating stove impact is preserved, with the exception of the IB/Coal stove pairing, when focusing upon the female population.

### References

- Abadie, A., D. Drukker, J. L. Herr, and G. W. Imbens (2004). "Implementing matching estimators for average treatment effects in Stata," *Stata Journal* 4(3): 290-311.
- Abadie, A. and G. W. Imbens (2002), "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects," NBER Technical Working Paper no. 283.
- Alberini, A., M. Cropper, T. Fu, A. Krupnick, J. Liu, D. Shaw, W. Harrington (1997). "Valuing

- Health Effects of Air Pollution in Developing Countries: The Case of Taiwan,” *Journal of Environmental Economics and Management* 34, 107-126.
- Alberini, A. and A. Krupnick (1998). “Air Quality and Episodes of Acute Respiratory Illness in Taiwan Cities: Evidence from Survey Data,” *Journal of Urban Economics* 44, 68-92.
- Andam, K., P. Ferraro, A. Pfaff, G. Arturo Sanchez-Azofeifa, J. Robalino (2008). “Correctly Measuring The Effectiveness Of Land Use Regulations In Reducing Deforestation”. Manuscript submitted.
- Angrist, J. D., & Krueger, A. B. (1999). *Empirical strategies in labor economics*. Princeton, N.J.: Industrial Relations Section, Dept. of Economics, Princeton University.
- Boy, Erick, Nigel Bruce, Hernan Delgado (2002). “Birth Weight and Exposure to Kitchen Wood Smoke During Pregnancy in Rural Guatemala,” *Environmental Health Perspectives* 110(11), 109-114.
- Brook, R.D., J.R. Brook, and S. Rajagopalan (2003). “ Air Pollution: The “Heart” of the Problem,” *Current Hypertension Reports* 5(1): 32-39.
- Bruce, Nigel, John McCracken, Rachel Albalak, Morten Schei, Kirk R. Smith, Victorina Lopez, and Chris West (2004). “Impact of Improved Stoves, House Construction, and Child Location on Levels of Indoor Air Pollution Exposure in Young Guatemalan Children,” *Journal of Exposure Analysis and Environmental Epidemiology* 14, S26-S33.
- Bruce, Nigel, Lynnette Neufeld, Erick Boy, Chris West (1998). “Indoor Biofuel Air Pollution and Respiratory Health: The Role of Confounding Factors Among Women in Highland Guatemala,” *International Journal of Epidemiology* 27, 454-458.
- Bruce, Nigel, Rogelio Perez-Padilla, Rachel Albalak (2000). “Indoor air pollution in developing countries: a major environmental and public health challenge,” *Bull. WHO* 78, 1078-1092.
- Dasgupta, S., M. Huq, M. Khaliquzaman, K. Pandey, D. Wheeler (2006). “Indoor Air Quality for Poor Families: New Evidence from Bangladesh,” *Indoor Air* 16, 426-444.
- Dehejia, R. and S. Wahba (1999). “Causal Effects in Nonexperimental Studies: Re-evaluating the Evaluation of Training Programs,” *J. of the American Statistical Association* 94, 1053-1062.
- Diprete, T.A., and M. Gangl (2004). “Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments,” *Sociological Methods* 34, 271-310.
- Edwards, Rufus D, Yaping Liu, Gongli He, Zheng Yin, Jonathan Sinton, John Peabody, Kirk R. Smith (2007). “Household CO and PM Measured As Part of A Review of China’s National Improved Stove Program,” *Indoor Air* 17(3), 189-203.
- Ezzati, Majid and Daniel Kammen (2002). “Household Energy, Indoor Air Pollution, and Heath in Developing Countries: Knowledge Base for Effective Interventions,” *Annual Review of Energy and the Environment* 27, 233-270.
- \_\_\_\_\_ (2001a). “Indoor Air Pollution from Biomass Comubstion and Acute Respiratory Infections in Kenya: An Exposure-Response Study,” *Lancet* 358, 619-624.

- \_\_\_\_\_. (2001b). "Quantifying the Effects of Exposure to Indoor Air Pollution from Biomass Combustion on Acute Respiratory Infections in Developing Countries," *Environmental Health Perspectives* 109 (5), 481-488.
- Ezzati, M., Saleh, H., and Kammen, D. (2000). "The Contributions of Emissions and Spatial Microenvironments to Exposure to Indoor Air Pollution from Biomass Combustion in Kenya," *Environmental Health Perspectives* 108 (9), 833-839.
- Ferraro, P., C. McIntosh, M. Ospina (2007). "The effectiveness of the US Endangered Species Act: An Econometric Analysis Using Matching Methods," *Journal of Environmental Economics and Management* 54(3), 245-336.
- Florig HK (1997). "China's air pollution risks," *Environmental Science and Technology* 31, 274A-279A.
- Heckman, J.J., H. Ichimura, and P. Todd (1997). "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies* 64, 605-654.
- Heckman, J.J., and J.A. Smith (1995). "Assessing the Case of Social Experiments," *Journal of Economic Perspectives* 9(2), 85-110.
- He., J., S. Vupputuri, K. Allen, M.R. Prerost, J. Hughes, and P.K. Whelton (1999). "Passive Smoking and the Risk of Coronary Heart Disease—A Meta-Analysis of Epidemiologic Studies," *New England Journal of Medicine* 340(12): 920-926.
- Hill, J. L., J. Brooks-Gunn, J. Waldfogel (2003). "Sustained Effects of High Participation in an Early Intervention for Low-Birth-Weight Premature Infants," *Developmental Psychology* 39(4), 730-744.
- Hill, J. and J. Reiter (2006). "Interval Estimation for Treatment Effects Using Propensity Score Matching," *Statistics in Medicine* 25, 2230-2256.
- Holm-Nielsen, Lauritz B. (2000). "The World Bank's Role in Science and Technology," presented at the International S&T Cooperation for Sustainable Development, OECD Seoul Conference November 15-17, 2000, <http://www1.worldbank.org/education/tertiary/documents/OECD%20KOREA%20final.pdf>.
- Imbens, G. (2004). "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *The Review of Economics and Statistics* 86(1), 4-29.
- \_\_\_\_\_. (2000). "The Role of the Propensity Score in Estimating Dose-Response Functions," *Biometrika* 87(3), 706-710.
- Lan, Qing, Robert Chapman, Dina Schreinemachers, Linwei Tian, Xingzhou He (2002). "Household Stove Improvement and Risk of Lung Cancer in Xuanwei, China," *Journal of the National Cancer Institute* 94(11), 826-835.
- Lechner, M. (2002). "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies.

- \_\_\_\_ (2001). "Identification and Estimation of Causal Effects of Multiple Treatments Under the Conditional Independence Assumption," in *Econometric Evaluation of Labour Market Policies*, ed. By M. Lechner, and F. Pfeiffer, Physica-Verlag, Heidelberg: 1-18.
- Leuven, E. and B. Sianesi. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". <http://ideas.repec.org/c/boc/bocode/s432001.html>. This version 3.0.0.
- Luo, RX, Wu B, Yi YN, Huang ZW, Lin RT (1996). "Indoor Burning Coal Air Pollution and Lung Cancer—a case control study in Fuzhou, China," *Lung Cancer* 14 Suppl. 1, S113-119.
- McCracken, J.P., K.R. Smith, A. Diaz, M.A. Mittleman, J. Schwartz (2007). "Chimney Stove Intervention to Reduce Long-term Wood Smoke Exposure Lowers Blood Pressure among Guatemalan Women," *Environmental Health Perspectives* 115(7), 996-1001.
- Mishra, Vinod, Xiaolei Dai, Kirk Smith, and Lasten Mika (2004). "Maternal Exposure to Biomass Smoke and Reduced Birth weight in Zimbabwe," *Annals of Epidemiology* 14(1), 740-747.
- Morgan, S. and D. Harding (2006). "Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice," *Sociological Methods & Research* 35(1), 3-60.
- Moshhammer, Hanns, Hans-Peter Hutter, Manfred Neuberger (2006). "Gas cooking and reduced lung function in school children," *Atmospheric Environment* 40, 3349-3354.
- Mumford, JL, Chapman RS, Harris DB, He XZ, Cao SR, Xian YL, Li XM (1989). "Indoor air Exposure to Coal and Wood Combustion Emissions Associated with a High Lung Cancer Rate in Xuan Wei, China," *Environmental International* 15, 315-320.
- Mumford JL, He XZ, Chapman RS, Cao SR, Harris DB, Li XM, et al. (1987). "Lung Cancer and Indoor Air Pollution in Xuan Wei, China," *Science* 235, 217-20.
- Peabody, J., T. Riddell, K. R. Smith, Y. Liu, Y. Zhao, G. Jianghui, M. Milet, and J. Sinton (2005). "Indoor Air Pollution in Rural China: Cooking Fuels, Stoves, and Health Status," *Archives of Environmental and Occupational Health* 60(2), 1-10.
- Pitt, Mark, Mark Rosenzweig, and Md. Nazmul Hassan (2006). "Sharing the Burden of Disease: Gender, the Household Division of Labor and the Health Effects of Indoor Air Pollution in Bangladesh and India," presented at the 2006 Stanford Institute for Theoretical Economics Summer Workshop, [www.stanford.edu/group/SITE/Web%20Session%202/Rosenzweig.pdf](http://www.stanford.edu/group/SITE/Web%20Session%202/Rosenzweig.pdf).
- Pfaff, A., J. A. Robalino, G. A. Sanchez-Azofeifa (2008a). "Payments for Environmental Services: Empirical Analysis for Costa Rica". Duke U., Sanford Working Paper Series SAN08-05.
- Pfaff, A., J. A. Robalino, G. A. Sanchez-Azofeifa, K. Andam, P. Ferraro (2008b). "Location Affects Protection: observable characteristics drive park impacts in Costa Rica". Duke U., Sanford Working Paper Series SAN08-06.
- Reddy AKN, RH Williams, TB Johansson, eds. (1996). *Energy After Rio: Prospects and Challenges*. New York, United Nations Publications.

- Rosenbaum, P.R. (2002). *Observational Studies*, Springer, New York.
- Sinton, Jonathan E., Kirk R. Smith, John W. Peabody, Yaping Liu, Rufus Edwards, Meredith Milet, Quan Gan, Zheng Yin (2004). "Improved Household Stoves in China: An Assessment of the National Improved Stove Program (NISP)," Shell Foundation, London, 2004, <http://ehs.sph.berkeley.edu/HEM/hem/China%Stoves/Presentations/3.USA-%20Kirk%20Smith/ChinaStoves.Final%20report.2004.pdf>.
- Smith, K.R. (1993). "Fuel combustion, air pollution exposure, and health: situation in developing countries". *Annual Review of Energy & Environment* 18:529-566.
- Smith KR, Samet JM, Romieu I, Bruce N. (2000). "Indoor air pollution in developing countries and acute lower respiratory infections in children." *Thorax* 55, 518-532.
- Smith, Kirk R., Sumi Mehta, Mirjam Maeusezahl-Feuz (2004). "Indoor smoke from household solid fuels," in Ezzati M, Rodgers AD, Lopez AD, Murray CJL (eds) *Comparative Quantification of Health Risks: Global and Regional Burden of Disease due to Selected Major Risk Factors*, Geneva: World Health Organization, Vol 2 pp. 1435-1493.
- Smith, Kirk R., Nigel Bruce, Byron Arana (2006). "RESPIRE - the Guatemala Randomized Intervention Trial," Symposium M3 at the ISEE/ISEA Annual Conference, Paris, *Epidemiology* 17(6) Suppl. November, S44-46, 2006.
- Thun, M., J. Henley, and L. Apicella (1999). "Epidemiologic Studies of Fatal and Nonfatal Cardiovascular Disease and ETS Exposure from Spousal Smoking," *Environmental Health Perspectives* 107(suppl 6), 841-846.
- United Nations Development Program, 2006. The GEF Small Grants Programme (<http://sgp.undp.org>).
- von Schirnding, Y., N. Bruce, K.R. Smith, G. Ballard-Tremeer, M. Ezzati and K. Lvovsky (2002). "Addressing the impact of household energy and indoor air pollution on the health of the poor – implications for policy action and intervention measures", in *Working Group 5 (Improving The Health Outcomes Of The Poor)*. Commission on Macroeconomics and Health.
- Whelton, S., Am. Hyre, B. Pederson, Y. Yi, P. Whelton, and J. He (2005). "Effect of Dietary Fiber Intake on Blood Pressure: A Meta-Analysis of Randomized, Controlled Clinical Trials," *Journal of Hypertension* 23(3), 475-481.
- World Bank (2003). *Indoor Air Pollution at a glance Factsheet*, November issue.
- World Health Organization (WHO) (2005). *Indoor Air Pollution and Health*, Fact sheet No. 292, June 2005, <http://www.who.int/mediacentre/factsheets/fs292/en/index.html>.
- World Health Organization (WHO) (1997). *Health and Environment in Sustainable Development*. WHO/EHG/97.8. Geneva: WHO.
- World Resources Institute, United Nations Environment Programme, United Nations Development Programme, World Bank (1998). *World Resources 1998-1999*. New York: Oxford U. Press.

Table 1: Differences in Averages and Marginal Effects of Adult Number of Symptoms' Linear Regression Covariates Among IB and TB Stove Users

| Variable                       | Sample |        | Diff. in means test |      | % Change | Marginal Effect | Sign of bias |
|--------------------------------|--------|--------|---------------------|------|----------|-----------------|--------------|
|                                | TB     | IB     | t                   | p> t |          |                 |              |
| Age 26-40                      | 0.38   | 0.42   | -1.87               | 0.06 | -0.11    | 0.06            | -            |
| Age 41-55                      | 0.31   | 0.34   | -1.37               | 0.17 | -0.09    | 0.25*           | -            |
| Age >55                        | 0.26   | 0.17   | 4.66                | 0.00 | 0.32     | 0.34**          | +            |
| Male                           | 0.33   | 0.31   | 0.69                | 0.49 | 0.04     | -0.08           | -            |
| Own washing machine            | 0.40   | 0.19   | 10.78               | 0.00 | 0.51     | -0.07           | -            |
| Cooking time (minutes)         | 112.28 | 134.21 | -9.61               | 0.00 | -0.20    | 1.76E-3***      | -            |
| One kitchen opening            | 0.14   | 0.34   | -9.96               | 0.00 | -1.43    | 0.26            | -            |
| Two kitchen openings           | 0.41   | 0.32   | 4.00                | 0.00 | 0.21     | 0.15            | +            |
| More than two kitchen openings | 0.44   | 0.28   | 7.47                | 0.00 | 0.35     | 0.01            | +            |
| Additional open air kitchen    | 0.25   | 0.03   | 17.88               | 0.00 | 0.87     | 0.47***         | +            |
| Income more than 12,000 Yuan   | 0.38   | 0.20   | 9.46                | 0.00 | 0.47     | -0.20**         | -            |
| Bad health                     | 0.36   | 0.42   | -2.55               | 0.01 | -0.16    | 0.70***         | -            |
| Smoker                         | 0.23   | 0.22   | 0.35                | 0.73 | 0.03     | 0.01            | +            |
| Province 1                     | 0.33   | 0.12   | 13.00               | 0.00 | 0.65     | 0.13            | +            |
| Province 2                     | 0.10   | 0.43   | -16.37              | 0.00 | -3.48    | 0.60***         | -            |

Note: The marginal effects are from the OLS regression which includes the users of the two stove groups, the covariates, and the stove treatment dummy variable. \*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 2: Differences in Averages and Marginal Effects of Adult Number of Symptoms' Linear Regression Covariates Among IB and Coal Stove Users

| Variable                       | Sample |        | Diff. in means test |      | % Change | Marginal Effect | Sign of bias |
|--------------------------------|--------|--------|---------------------|------|----------|-----------------|--------------|
|                                | Coal   | IB     | t                   | p> t |          |                 |              |
| Age 26-40                      | 0.55   | 0.42   | 7.21                | 0.00 | 0.23     | 0.12            | +            |
| Age 41-55                      | 0.25   | 0.34   | -5.36               | 0.00 | -0.35    | 0.33***         | -            |
| Age >55                        | 0.10   | 0.17   | -5.92               | 0.00 | -0.76    | 0.49***         | -            |
| Male                           | 0.24   | 0.31   | -4.46               | 0.00 | -0.30    | -9.81E-4        | +            |
| Own washing machine            | 0.40   | 0.19   | 12.84               | 0.00 | 0.51     | -0.12*          | -            |
| Cooking time (minutes)         | 138.46 | 134.21 | 2.39                | 0.02 | 0.03     | 1.24E-3**       | +            |
| One kitchen opening            | 0.23   | 0.34   | -6.81               | 0.00 | -0.49    | 0.33**          | -            |
| Two kitchen openings           | 0.50   | 0.32   | 9.94                | 0.00 | 0.35     | 0.26**          | +            |
| More than two kitchen openings | 0.21   | 0.28   | -4.43               | 0.00 | -0.33    | 0.26**          | -            |
| Additional open air kitchen    | 0.08   | 0.03   | 6.05                | 0.00 | 0.59     | 0.41***         | +            |
| Income more than 12,000 Yuan   | 0.13   | 0.20   | -5.44               | 0.00 | -0.59    | -0.11           | +            |
| Bad health                     | 0.42   | 0.42   | -0.01               | 0.99 | 0.00     | 0.62***         | +            |
| Smoker                         | 0.18   | 0.22   | -3.12               | 0.00 | -0.26    | -0.08           | +            |
| Province 1                     | 0.70   | 0.12   | 43.04               | 0.00 | 0.84     | 0.18*           | +            |
| Province 2                     | 0.30   | 0.43   | -7.73               | 0.00 | -0.45    | 0.55***         | -            |

Note: The marginal effects are from the OLS regression which includes the users of the two

Table 3: Differences in Averages and Marginal Effects of Adult Number of Symptoms' Linear Regression Covariates Among Clean-fuel and Coal Stove Users

| Variable                       | Sample     |        | Diff. in means test |      | % Change | Marginal Effect | Sign of bias |
|--------------------------------|------------|--------|---------------------|------|----------|-----------------|--------------|
|                                | Clean-fuel | Coal   | t                   | p> t |          |                 |              |
| Age 26-40                      | 0.61       | 0.55   | 1.66                | 0.10 | 0.10     | 0.15            | +            |
| Age 41-55                      | 0.27       | 0.25   | 0.61                | 0.54 | 0.07     | 0.31**          | +            |
| Age >55                        | 0.07       | 0.10   | -1.47               | 0.14 | -0.45    | 0.58***         | -            |
| Male                           | 0.34       | 0.24   | 3.26                | 0.00 | 0.30     | 0.16            | +            |
| Own washing machine            | 0.37       | 0.40   | -0.82               | 0.41 | -0.08    | -0.21***        | +            |
| Cooking time (minutes)         | 101.03     | 138.46 | -11.08              | 0.00 | -0.37    | -2.51E-4        | +            |
| One kitchen opening            | 0.44       | 0.23   | 6.85                | 0.00 | 0.48     | 0.38**          | +            |
| Two kitchen openings           | 0.29       | 0.50   | -5.74               | 0.00 | -0.68    | 0.40**          | -            |
| More than two kitchen openings | 0.22       | 0.21   | 0.17                | 0.86 | 0.02     | 0.57***         | +            |
| Additional open air kitchen    | 0.04       | 0.08   | -2.04               | 0.04 | -0.89    | 0.39***         | -            |
| Income more than 12,000 Yuan   | 0.36       | 0.13   | 9.31                | 0.00 | 0.66     | -0.05           | -            |
| Bad health                     | 0.35       | 0.42   | -1.88               | 0.06 | -0.19    | 0.45***         | -            |
| Smoker                         | 0.22       | 0.18   | 1.55                | 0.12 | 0.20     | -0.08           | -            |
| Province 1                     | 0.12       | 0.70   | -18.47              | 0.00 | -4.67    | 0.20            | -            |
| Province 2                     | 0.41       | 0.30   | 3.45                | 0.00 | 0.28     | 0.39**          | +            |

Note: The marginal effects are from the OLS regression which includes the users of the two stove groups, the covariates, and the stove treatment dummy variable. \*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 4: Differences in Averages and Marginal Effects of Adult Number of Symptoms' Linear Regression Covariates Among Clean-fuel and TB Stove Users

| Variable                       | Sample     |        | Diff. in means test |      | % Change | Marginal Effect | Sign of bias |
|--------------------------------|------------|--------|---------------------|------|----------|-----------------|--------------|
|                                | Clean-fuel | TB     | t                   | p> t |          |                 |              |
| Age 26-40                      | 0.61       | 0.38   | 6.16                | 0.00 | 0.37     | -0.01           | -            |
| Age 41-55                      | 0.27       | 0.31   | -1.16               | 0.25 | -0.15    | 0.30            | -            |
| Age >55                        | 0.07       | 0.26   | -6.21               | 0.00 | -2.75    | 0.20            | -            |
| Male                           | 0.34       | 0.33   | 0.39                | 0.69 | 0.04     | 0.11            | +            |
| Own washing machine            | 0.37       | 0.40   | -0.86               | 0.39 | -0.09    | -0.05           | +            |
| Cooking time (minutes)         | 101.03     | 112.28 | -2.93               | 0.00 | -0.11    | 7.60E-4         | -            |
| One kitchen opening            | 0.44       | 0.14   | 10.04               | 0.00 | 0.68     | 0.11            | +            |
| Two kitchen openings           | 0.29       | 0.41   | -3.13               | 0.00 | -0.39    | 0.16            | -            |
| More than two kitchen openings | 0.22       | 0.44   | -6.06               | 0.00 | -1.01    | 0.04            | -            |
| Additional open air kitchen    | 0.04       | 0.25   | -6.95               | 0.00 | -4.75    | -0.10           | +            |
| Income more than 12,000 Yuan   | 0.36       | 0.38   | -0.40               | 0.69 | -0.04    | 0.02            | -            |
| Bad health                     | 0.35       | 0.36   | -0.27               | 0.79 | -0.03    | 0.23            | -            |
| Smoker                         | 0.22       | 0.23   | -0.29               | 0.77 | -0.04    | -0.09           | +            |
| Province 1                     | 0.12       | 0.33   | -6.12               | 0.00 | -1.64    | dropped         |              |
| Province 2                     | 0.41       | 0.10   | 11.54               | 0.00 | 0.77     | 0.07            | +            |

Note: The marginal effects are from the OLS regression which includes the users of the two stove groups, the covariates, and the stove treatment dummy variable. \*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 5: Differences in Averages and Marginal Effects of Adult Number of Symptoms' Linear Regression Covariates Among Clean-fuel and IB Stove Users

| Variable                       | Sample     |        | Diff. in means test |      | % Change | Marginal Effect | Sign of bias |
|--------------------------------|------------|--------|---------------------|------|----------|-----------------|--------------|
|                                | Clean-fuel | IB     | t                   | p> t |          |                 |              |
| Age 26-40                      | 0.61       | 0.42   | 5.49                | 0.00 | 0.31     | 0.04            | +            |
| Age 41-55                      | 0.27       | 0.34   | -2.14               | 0.03 | -0.25    | 0.29*           | -            |
| Age >55                        | 0.07       | 0.17   | -4.14               | 0.00 | -1.54    | 0.38**          | -            |
| Male                           | 0.34       | 0.31   | 0.89                | 0.37 | 0.08     | -0.09           | -            |
| Own washing machine            | 0.37       | 0.19   | 6.15                | 0.00 | 0.47     | -0.02           | -            |
| Cooking time (minutes)         | 101.03     | 134.21 | -9.27               | 0.00 | -0.33    | 1.93E-3***      | -            |
| One kitchen opening            | 0.44       | 0.34   | 3.04                | 0.00 | 0.23     | 0.19            | +            |
| Two kitchen openings           | 0.29       | 0.32   | -0.92               | 0.36 | -0.10    | 0.08            | -            |
| More than two kitchen openings | 0.22       | 0.28   | -2.09               | 0.04 | -0.30    | -0.01           | +            |
| Additional open air kitchen    | 0.04       | 0.03   | 0.78                | 0.44 | 0.23     | 0.19            | +            |
| Income more than 12,000 Yuan   | 0.36       | 0.20   | 5.84                | 0.00 | 0.45     | -0.11           | -            |
| Bad health                     | 0.35       | 0.42   | -1.94               | 0.05 | -0.19    | 0.66***         | -            |
| Smoker                         | 0.22       | 0.22   | -0.09               | 0.93 | -0.01    | -0.05           | +            |
| Province 1                     | 0.12       | 0.12   | 0.37                | 0.71 | 0.07     | 0.14            | +            |
| Province 2                     | 0.41       | 0.43   | -0.59               | 0.56 | -0.05    | 0.53***         | -            |

Note: The marginal effects are from the OLS regression which includes the users of the two stove groups, the covariates, and the stove treatment dummy variable. \*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 6: OLS Results for Health Outcomes

| Health outcome<br>Variable              | Number of Symptoms<br>Parameter | Specific Symptom X<br>T-statistic | Specific Symptom Y |
|---|---------------------------------|-----------------------------------|--------------------|
| Age 26-40                               | 0.10                            | 1.21                              |                    |
| Age 41-55                               | 0.30***                         | 3.25                              |                    |
| Age >55                                 | 0.44***                         | 4.36                              |                    |
| Male                                    | -0.02                           | -0.37                             |                    |
| Own washing<br>machine                  | -0.15***                        | -2.91                             |                    |
| Cooking time<br>(minutes)               | 9.41E-4**                       | 2.24                              |                    |
| One kitchen<br>opening                  | 0.30***                         | 2.85                              |                    |
| Two kitchen<br>openings                 | 0.25**                          | 2.41                              |                    |
| More than two<br>kitchen openings       | 0.21**                          | 2.02                              |                    |
| Additional open air<br>kitchen          | 0.55***                         | 6.44                              |                    |
| Income more than<br>12,000 Yuan         | -0.12**                         | -2.17                             |                    |
| Bad health                              | 0.60***                         | 13.21                             |                    |
| Smoker                                  | 0.03                            | 0.38                              |                    |
| Smoked more than<br>40 years            |                                 |                                   |                    |
| Days drink alcohol                      |                                 |                                   |                    |
| Body mass index                         |                                 |                                   |                    |
| Eat vegetables 3 or<br>more times a day |                                 |                                   |                    |
| Province 1                              | 0.09                            | 1.43                              |                    |
| Province 2                              | 0.48***                         | 8.53                              |                    |
| Constant                                | 0.00067                         | 0                                 |                    |
| Adjusted R-squared                      | 0.11                            |                                   |                    |
| Root MSE                                | 1.36                            |                                   |                    |
| Observations                            | 3635                            |                                   |                    |

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 7: Effect of Stove Use on Number of Symptoms Reported

| Treatment                             | IB                  | IB                | Clean-fuel          | Clean-fuel          | Clean-fuel          |
|---------------------------------------|---------------------|-------------------|---------------------|---------------------|---------------------|
| Control                               | TB†                 | Coal†             | Coal                | TB                  | IB                  |
| Naïve                                 | -0.04<br>(-0.55)    | 0.17***<br>(3.07) | -0.26***<br>(-2.58) | -0.47***<br>(-3.94) | -0.43***<br>(-3.95) |
| Naïve bias-adjusted                   | -0.26***<br>(-3.20) | 0.14*<br>(1.92)   | -0.15<br>(-1.07)    | -0.49***<br>(-3.91) | -0.22**<br>(-2.04)  |
| NN CM                                 | -0.22<br>(-1.62)    | 0.06<br>(0.72)    | -0.10<br>(-0.95)    | -0.55***<br>(-3.40) | -0.20**<br>(-2.09)  |
| Bias-adjusted NN CM                   | -0.22*<br>(-1.71)   | 0.06<br>(0.66)    | -0.06<br>(-0.62)    | -0.65***<br>(-4.05) | -0.19**<br>(-1.97)  |
| NN PSM                                | -0.16<br>(-1.27)    | 0.13<br>(1.19)    | -0.70**<br>(-1.97)  | -0.75***<br>(-4.65) | -0.18*<br>(-1.74)   |
| Bias-adjusted NN PSM                  | -0.25*<br>(-1.74)   | 0.12<br>(1.40)    | -0.34*<br>(-1.89)   | -0.66***<br>(-3.21) | -0.22**<br>(-2.30)  |
| NN PSM (caliper=0.001)                | -0.19<br>(-1.52)    | 0.03<br>(0.26)    | 0.08<br>(0.38)      | -0.34*<br>(-1.82)   | -0.21*<br>(-1.90)   |
| Bias-adjusted NN PSM (caliper=0.001)  | -0.29***<br>(-3.27) | 0.14<br>(1.45)    | 0.10<br>(0.53)      | -0.36<br>(-1.46)    | -0.24**<br>(-2.21)  |
| NN PSM (common support)               | -0.16<br>(-1.27)    | 0.13<br>(1.19)    | -0.14<br>(-0.85)    | -0.73***<br>(-4.56) | -0.18*<br>(-1.74)   |
| Bias-adjusted NN PSM (common support) | -0.25*<br>(-1.74)   | 0.12<br>(1.40)    | -0.14<br>(-0.91)    | -0.64***<br>(-3.24) | -0.22**<br>(-2.30)  |
| KG PSM                                | -0.19*<br>(-1.74)   | 0.06<br>(0.71)    | -0.83<br>(-1.23)    | -0.65***<br>(-4.72) | -0.27***<br>(-3.10) |
| KG PSM (caliper=0.001)                | -0.19*<br>(-1.74)   | 0.06<br>(0.71)    | -0.83<br>(-1.23)    | -0.65***<br>(-4.72) | -0.27***<br>(-3.10) |
| KG PSM (common support)               | -0.19*<br>(-1.74)   | 0.06<br>(0.71)    | -0.16<br>(-1.07)    | -0.64***<br>(-4.67) | -0.27***<br>(-3.10) |

Note: T-statistics in parentheses. \*\*\* P<0.01, \*\* P<0.05, \* P<0.10. † We ran TB and coal as treatments to improve the performance of matching.

Covariates in treatment regression are male, age category dummy variables (26-40, 41-55, >55), income greater than 12,000 Yuan dummy variable, smoker dummy variable, washing machine dummy variable, minutes spent cooking in a day, openings in kitchen category dummy variables (1, 2, >2), dummy variable for having an open air kitchen, dummy variable for bad health, and province fixed effects. When using the nearest neighbor algorithm, we use four matches.

Table 8: Effect of IB (vs. TB) Stove Use on Individual Symptoms

| Symptom | Headache<br>> 4 hours | > 5 headaches<br>in a<br>week | Nausea<br>> 8 hours | Nausea<br>with<br>vomiting | Dizzy              | Cough<br>with<br>phlegm | Both eyes<br>irritated<br>> 4 hours | Both eyes<br>irritated > 5<br>times a week | Both eyes<br>itch   | History<br>of eye(s)<br>allergies | Eyes stick<br>together from<br>Discharge |
|---------|-----------------------|-------------------------------|---------------------|----------------------------|--------------------|-------------------------|-------------------------------------|--|---------------------|-----------------------------------|--|
| Naïve   | 0.01<br>(0.68)        | -0.10***<br>(-6.23)           | -3.27E-3<br>(-0.36) | -0.04**<br>(-2.19)         | -0.04*<br>(-1.72)  | -0.02<br>(-0.93)        | -0.02**<br>(-2.02)                  | -0.04***<br>(-3.72)                        | -3.33E-3<br>(-0.19) | 2.14E-3<br>(0.17)                 | -0.01<br>(-0.71)                         |
| NN CM   | -0.01<br>(-0.39)      | -0.14***<br>(-4.45)           | -0.01<br>(-1.08)    | -0.06**<br>(-2.01)         | -0.09**<br>(-2.18) | 0.03<br>(1.09)          | -0.02<br>(-1.06)                    | 0.01<br>(0.37)                             | 2.50E-3<br>(0.11)   | 0.02<br>(1.04)                    | 0.03<br>(1.47)                           |
| NN PSM  | 8.91E-4<br>(0.04)     | -0.13***<br>(-4.82)           | -0.01<br>(-0.47)    | -0.07***<br>(-2.81)        | -0.07*<br>(-1.90)  | 0.01<br>(0.25)          | -0.02**<br>(-2.08)                  | -0.01<br>(-0.42)                           | -8.12E-4<br>(-0.04) | 0.02<br>(0.92)                    | 0.01<br>(0.56)                           |
| KG PSM  | 2.23E-3<br>(0.12)     | -0.13***<br>(-5.96)           | -2.05E-3<br>(-0.21) | -0.05***<br>(-2.76)        | -0.08**<br>(-2.58) | 0.01<br>(0.31)          | -0.02*<br>(-1.69)                   | -0.03<br>(-1.52)                           | 2.64E-3<br>(0.15)   | 0.01<br>(0.55)                    | 0.01<br>(0.55)                           |

Note: These are the estimates corrected for bias after the matching procedure. T-statistics are placed below impact estimates in parentheses.

We ran TB and coal as treatments to improve the performance of matching. CM and PSM indicate covariate and propensity score matching, respectively.

NN refers to nearest-neighbor, and KG refers to Kernel assuming Gaussian kernel.

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 9: Effect of IB (vs. Coal) Stove Use on Individual Symptoms

| Symptom | Headache<br>> 4 hours | > 5 headaches<br>in a<br>week | Nausea<br>> 8 hours | Nausea<br>with<br>vomiting | Dizzy            | Cough<br>with<br>phlegm | Both eyes<br>irritated<br>> 4 hours | Both eyes<br>irritated > 5<br>times a week | Both eyes<br>itch | History<br>of eye(s)<br>allergies | Eyes stick<br>together from<br>discharge |
|---------|-----------------------|-------------------------------|---------------------|----------------------------|------------------|-------------------------|-------------------------------------|--|-------------------|-----------------------------------|--|
| Naïve   | 0.01<br>(0.49)        | -0.01<br>(-0.57)              | 2.77E-3<br>(0.35)   | -0.02<br>(-1.36)           | 0.05**<br>(2.13) | 0.02<br>(1.05)          | 0.01<br>(0.73)                      | 1.87E-3<br>(0.24)                          | 0.03**<br>(2.04)  | 0.02*<br>(1.96)                   | 0.03***<br>(2.61)                        |
| NN CM   | -0.01<br>(-0.61)      | -0.04*<br>(-1.91)             | -3.99E-3<br>(-0.46) | -0.06***<br>(-3.14)        | 0.07**<br>(2.37) | 0.03<br>(1.59)          | 2.60E-3<br>(0.26)                   | -0.01<br>(-1.03)                           | 0.02<br>(1.10)    | 0.02<br>(1.40)                    | 0.03**<br>(2.24)                         |
| NN PSM  | 8.81E-4<br>(0.05)     | -0.03*<br>(-1.81)             | 2.33E-3<br>(0.32)   | -0.04**<br>(-2.36)         | 0.07**<br>(2.46) | 0.02<br>(1.27)          | 0.01<br>(0.56)                      | -0.01<br>(-0.94)                           | 0.03<br>(1.35)    | 0.02*<br>(1.89)                   | 0.03**<br>(2.04)                         |
| KG PSM  | -0.01<br>(-0.35)      | -0.02<br>(-1.27)              | -2.52E-4<br>(-0.04) | -0.04***<br>(-3.04)        | 0.04<br>(1.57)   | 0.02<br>(1.07)          | -1.08E-3<br>(-0.15)                 | -0.01<br>(-1.29)                           | 0.03<br>(1.51)    | 0.02*<br>(1.82)                   | 0.02*<br>(1.77)                          |

Note: These are the estimates corrected for bias after the matching procedure. T-statistics are placed below impact estimates in parentheses.

We ran TB and coal as treatments to improve the performance of matching. CM and PSM indicate covariate and propensity score matching, respectively.

NN refers to nearest-neighbor, and KG refers to Kernel assuming Gaussian kernel.

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 10: Effect of Clean-fuel (vs. Coal) Stove Use on Individual Symptoms

| Symptom | Headache > 4 hours | > 5 headaches in a week | Nausea > 8 hours    | Nausea with vomitin g | Dizzy              | Cough with phlegm | Both eyes irritated > 4 hours | Both eyes irritated > 5 times a week | Both eyes itch      | History of eye(s) allergies | Eyes stick together from discharge |
|---------|--------------------|-------------------------|---------------------|-----------------------|--------------------|-------------------|-------------------------------|--------------------------------------|---------------------|-----------------------------|------------------------------------|
| Naïve   | -0.02<br>(-0.59)   | -0.03<br>(-1.02)        | -1.72E-4<br>(-0.01) | 0.03<br>(1.04)        | -0.05<br>(-1.03)   | -0.05<br>(-1.61)  | 6.90E-4<br>(0.05)             | -8.75E-5<br>(-0.01)                  | -0.01<br>(-0.35)    | -4.55E-3<br>(-0.28)         | -0.02<br>(-0.86)                   |
| NN CM   | -0.01<br>(-0.38)   | 0.01<br>(0.38)          | 1.97E-3<br>(0.21)   | 0.05*<br>(1.73)       | -0.08**<br>(-2.04) | -0.01<br>(-0.52)  | 0.01<br>(1.10)                | 0.01<br>(0.50)                       | -0.02<br>(-0.74)    | -0.02<br>(-1.49)            | -7.11E-4<br>(-0.05)                |
| NN PSM  | 0.03<br>(1.18)     | 1.33E-3<br>(0.05)       | 0.01<br>(0.63)      | -0.09*<br>(-1.73)     | -0.14**<br>(-2.10) | -0.02<br>(-0.55)  | -6.29E-5<br>(-0.00)           | -0.07*<br>(-1.83)                    | -0.05<br>(-1.09)    | -2.46E-3<br>(-0.13)         | -2.23E-3<br>(-0.12)                |
| KG PSM  | 0.01<br>(0.22)     | -0.01<br>(-0.50)        | 2.20E-3<br>(0.18)   | 0.15***<br>(-3.09)    | 0.18***<br>(-3.35) | -0.04<br>(-1.39)  | 3.80E-4<br>(0.03)             | -0.01<br>(-0.90)                     | -2.72E-4<br>(-0.01) | 3.55E-3<br>(0.18)           | -0.01<br>(-0.93)                   |

Note: These are the estimates corrected for bias after the matching procedure. T-statistics are placed below impact estimates in parentheses.

We ran TB and coal as treatments to improve the performance of matching. CM and PSM indicate covariate and propensity score matching, respectively.

NN refers to nearest-neighbor, and KG refers to Kernel assuming Gaussian kernel.

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 11: Effect of Clean-fuel (vs. TB) Stove Use on Individual Symptoms

| Symptom | Headache > 4 hours | > 5 headaches in a week | Nausea > 8 hours   | Nausea with vomiting | Dizzy              | Cough with phlegm | Both eyes irritated > 4 hours | Both eyes irritated > 5 times a week | Both eyes itch     | History of eye(s) allergies | Eyes stick together from discharge |
|---------|--------------------|-------------------------|--------------------|----------------------|--------------------|-------------------|-------------------------------|--------------------------------------|--------------------|-----------------------------|------------------------------------|
| Naïve   | -0.02<br>(-0.94)   | -0.08***<br>(-2.67)     | -0.03**<br>(-2.47) | -2.39E-3<br>(-0.08)  | 0.12***<br>(-3.09) | -0.05<br>(-1.58)  | -0.01<br>(-0.74)              | -0.02<br>(-1.05)                     | -0.05**<br>(-2.26) | 0.05***<br>(-3.32)          | -0.06***<br>(-3.20)                |
| NN CM   | -0.02<br>(-0.89)   | -0.10***<br>(-3.13)     | -0.04**<br>(-2.12) | 2.36E-3<br>(0.07)    | 0.16***<br>(-3.42) | -0.05<br>(-1.50)  | -0.02<br>(-1.14)              | -0.02<br>(-1.54)                     | -0.07**<br>(-2.39) | 0.07***<br>(-3.02)          | -0.10***<br>(-3.63)                |
| NN PSM  | -0.04<br>(-1.37)   | -0.08**<br>(-2.30)      | -0.05*<br>(-1.90)  | -0.01<br>(-0.14)     | 0.15***<br>(-3.14) | -0.06<br>(-1.47)  | -0.02<br>(-0.82)              | -0.02<br>(-0.74)                     | -0.09**<br>(-2.33) | -0.06**<br>(-2.09)          | -0.09***<br>(-2.68)                |
| KG PSM  | -0.03<br>(-1.01)   | -0.09***<br>(-3.06)     | -0.04**<br>(-2.00) | 3.95E-4<br>(0.01)    | 0.15***<br>(-3.54) | -0.05<br>(-1.59)  | -0.02<br>(-0.83)              | -0.01<br>(-0.87)                     | -0.08**<br>(-2.54) | 0.07***<br>(-2.74)          | -0.08***<br>(-2.83)                |

Note: These are the estimates corrected for bias after the matching procedure. T-statistics are placed below impact estimates in parentheses.

We ran TB and coal as treatments to improve the performance of matching. CM and PSM indicate covariate and propensity score matching, respectively.

NN refers to nearest-neighbor, and KG refers to Kernel assuming Gaussian kernel.

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 12: Effect of Clean-fuel (vs. IB) Stove Use on Individual Symptoms

| Symptom | Headache > 4 hours | > 5 headaches in a Week | Nausea > 8 hours  | Nausea with vomitin g | Dizzy              | Cough            | Both eyes irritated > 4 hours | Both eyes irritated > 5 times a week | Both eyes itch     | History of eye(s) allergies | Eyes stick together from discharge |
|---------|--------------------|-------------------------|-------------------|-----------------------|--------------------|------------------|-------------------------------|--------------------------------------|--------------------|-----------------------------|------------------------------------|
| Naïve   | -0.04<br>(-1.53)   | -0.01<br>(-0.29)        | -0.01<br>(-1.12)  | 0.03<br>(1.30)        | -0.08**<br>(-2.37) | -0.01<br>(-0.49) | 6.58E-4<br>(0.05)             | -3.30E-4<br>(-0.03)                  | -0.04<br>(-1.53)   | -0.04**<br>(-2.23)          | -0.03<br>(-1.42)                   |
| NN CM   | -0.04*<br>(-1.67)  | -0.01<br>(-0.42)        | -0.01<br>(-0.63)  | 0.04*<br>(1.82)       | -0.08**<br>(-2.29) | -0.02<br>(-1.02) | -6.13E-4<br>(-0.05)           | -0.01<br>(-0.64)                     | -0.03<br>(-1.32)   | -0.02*<br>(-1.82)           | -0.02<br>(-1.15)                   |
| NN PSM  | -0.05**<br>(-2.30) | 0.01<br>(0.44)          | -0.01<br>(-1.46)  | 0.03<br>(1.18)        | -0.07**<br>(-1.99) | -0.03<br>(-1.19) | 1.65E-3<br>(0.15)             | -4.23E-3<br>(-0.41)                  | -0.03<br>(-1.63)   | 0.04***<br>(-2.77)          | -0.03*<br>(-1.87)                  |
| KG PSM  | -0.05**<br>(-2.53) | -0.01<br>(-0.73)        | -0.02*<br>(-1.79) | 0.02<br>(0.97)        | 0.08***<br>(-2.72) | -0.01<br>(-0.44) | -3.24E-4<br>(-0.03)           | -1.05E-3<br>(-0.10)                  | -0.04**<br>(-2.23) | 0.03***<br>(-3.23)          | -0.03**<br>(-2.21)                 |

Note: These are the estimates corrected for bias after the matching procedure. T-statistics are placed below impact estimates in parentheses.

We ran TB and coal as treatments to improve the performance of matching. CM and PSM indicate covariate and propensity score matching, respectively.

NN refers to nearest-neighbor, and KG refers to Kernel assuming Gaussian kernel.

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10.

Table 13: Demographic, Housing, and Behavioral Characteristics by Gender

| Gender                               | Female    | Male      |
|--------------------------------------|-----------|-----------|
| Variables                            | Mean (%)  | Mean (%)  |
| No. of symptoms                      | 0.95      | 1.00      |
| Headache > 4hours                    | (10.22)   | (8.99)    |
| >5 headaches in a week               | (9.39)    | (11.28)   |
| Nausea > 8 hours                     | (2.31)    | (1.76)    |
| Nausea with vomiting                 | (10.25)   | (9.78)    |
| Dizzy                                | (29.44)   | (30.04)   |
| Cough with phlegm                    | (9.96)    | (12.42)   |
| Both eyes irritated > 4 hours        | (2.51)    | (2.47)    |
| Both eyes irritated > 5 times a week | (2.64)    | (2.82)    |
| Both eyes itch                       | (9.29)    | (9.87)    |
| History of eye(s) allergies          | (3.57)    | (4.76)    |
| Eyes stick together from discharge   | (4.89)    | (5.55)    |
| Open Air kitchen                     | (7.97)    | (8.99)    |
| Bad health                           | (39.05)   | (40.79)   |
| Smoker                               | (2.25)    | (63.44)   |
| Vegetables more than 3 times a day   | (23.18)   | (22.29)   |
| Age                                  | 38.67     | 47.44     |
| Openings in kitchen                  | 2.08      | 2.13      |
| Cooktime (minutes)                   | 127.99    | 127.35    |
| Days drink alcohol                   | 0.75      | 7.89      |
| Body mass index                      | 2.28E10-3 | 2.23E10-3 |
| Years a smoker                       | 0.48      | 14.78     |

Table 14: Effect of Stove Use on the Number of Symptoms by Gender

| Treatment<br>Control | IB<br>TB† |             | IB<br>Coal† |             | Clean-fuel<br>Coal |             | Clean-fuel<br>TB |             | Clean-fuel<br>IB |             |
|----------------------|-----------|-------------|-------------|-------------|--------------------|-------------|------------------|-------------|------------------|-------------|
|                      | Effect    | t-statistic | Effect      | t-statistic | Effect             | t-statistic | Effect           | t-statistic | Effect           | t-statistic |
| Naïve                | -0.26***  | -3.21       | 0.14*       | 1.92        | -0.14              | -1.00       | -0.49***         | -3.93       | -0.22**          | -2.07       |
| Naïve : Female       | -0.25***  | -2.60       | 0.20**      | 2.53        | -0.15              | -0.97       | -0.62***         | -4.16       | -0.32**          | -2.42       |
| Naïve : Male         | -0.20     | -1.38       | -0.02       | -0.13       | -0.07              | -0.24       | -0.16            | -0.66       | -0.01            | -0.07       |
| NN CM                | -0.28**   | -2.11       | 0.03        | 0.38        | -0.11              | -1.05       | -0.66***         | -4.11       | -0.19*           | -1.90       |
| NN CM : Female       | -0.35**   | -2.47       | 0.12        | 1.18        | -0.10              | -0.86       | -0.77***         | -3.94       | -0.25**          | -2.26       |
| NN CM : Male         | 0.18      | 0.78        | -0.13       | -0.90       | -0.23              | -1.11       | -0.39*           | -1.68       | 1.23E-3          | 0.01        |
| NN PSM               | -0.23     | -1.62       | 0.09        | 1.06        | -0.23              | -1.19       | -0.79***         | -4.00       | -0.25**          | -2.55       |
| NN PSM : Female      | -0.33***  | -3.10       | 0.19*       | 1.95        | -0.16              | -0.86       | -0.82***         | -3.65       | -0.26**          | -2.38       |
| NN PSM : Male        | 0.18      | 0.47        | -0.11       | -0.57       | -0.32              | -0.99       | -0.29            | -1.03       | -0.06            | -0.38       |
| KG PSM               | -0.19*    | -1.80       | 0.07        | 0.87        | -0.63              | -1.39       | -0.64***         | -3.34       | -0.28***         | -3.00       |
| KG PSM : Female      | -0.18*    | -1.93       | 0.16*       | 1.76        | -0.13              | -0.23       | -0.76***         | -3.29       | -0.37***         | -3.64       |
| KG PSM : Male        | 0.11      | 0.23        | 0.16        | 1.02        | -0.84              | -1.48       | -0.46            | -0.99       | -0.10            | -0.62       |

Note: NN refers to nearest-neighbor, and KG refers to Kernel assuming Gaussian kernel.

We ran TB and coal as treatments to improve the performance of matching. CM and PSM indicate covariate and propensity score matching, respectively.

The NN estimates are adjusted for bias using a regression. KG standard error estimates are bootstrapped with 100 repetitions.

\*\*\* P<0.01, \*\* P<0.05, \* P<0.10. --- indicates to few observations to estimate effect.