University of Sussex
Economics

Working Paper Series

No. 72-2014

The Impact of Cooking with Firewood on Respiratory Health: Evidence from Indonesia

Ani Rudra Silwal
Department of Economics, University of Sussex, Falmer, UK
A.Silwal@sussex.ac.uk

Andy McKay
Department of Economics, University of Sussex, Falmer, UK
A.McKay@sussex.ac.uk

Abstract: The vast majority of households in low-income countries cook with firewood, which is known to produce various airborne toxins. We examine whether cooking with firewood results in poorer respiratory health by using a unique Indonesian household survey that collected direct measures of lung capacity. We find that individuals living in households that cook with firewood have 11.2 per cent lower lung capacity than those that cook with cleaner fuels. This impact is larger for women and children than for men. The results strongly support the international policy focus on encouraging households to switch to cooking with cleaner fuels.

Key words: Health production; Indoor air pollution; Household energy use

JEL classification: I12, Q53, O13.

Acknowledgements: We would like to thank Kathleen Beegle, Barry Reilly, Kalle Hirvonen, Pedro Rosa Dias, Maria Eugenia Jenoni, Firman Witolear, and Tsegay Tekleselassie, as well as participants at the Royal Economic Society Conference 2013, the Oxford Centre for the Study of African Economies Conference 2013, and the Sussex Economics DPhil Seminar for comments, advice, and suggestions. We are responsible for all remaining errors.

1 Introduction

About 3 billion people around the world and 91 per cent of people in low-income countries still cook with solid fuels such as firewood, charcoal, dung cakes, and crop residue (WHO, 2012). Cooking with solid fuels has various implications; firewood collection could contribute to forest depletion, households may need to spend a disproportionate amount of time on collecting these fuels, and smoke from burning solid fuels could adversely affect health. All of these facts may justify public action to encourage people to switch to cleaner fuels such as kerosene, liquefied petroleum gas, and electricity. But of particular concern is the public health burden of solid fuels, which is believed to be very high. WHO (2009) attributes 3.9 million deaths to indoor air pollution from solid fuels every year in developing countries, more than any other environmental risk and twice as much as urban air pollution. To add to this, poorer respiratory health is known to shorten lives, raise morbidity, and increase productivity (Duflo et al. (2008); Strauss and Thomas (2008)).

Early epidemiological studies prompted policymakers to recognise the magnitude of this issue by placing it on the global public health agenda (Ezzati and Kammen (2001); Smith et al. (2004)). Indicator 29 of the Millennium Development Goals now tracks the proportion of the population using solid fuels in each country. Much recent debate has revolved around the adoption of modern cookstoves that produce less indoor air pollution than their traditional counterparts (World Bank, 2011) and which have been heavily promoted by the Global Alliance for Clean Cookstoves, backed by influential supporters like Hillary Clinton and funded by various governments and multilateral donors. There is controversy around these cookstoves, with some recent experimental evidence questioning their efficacy and uptake (Hanna et al. (2012); Miller and Mobarak (2011)). But this is a wider issue than clean cookstoves. Policies encouraging households to switch from solid fuels to kerosene, liquefied petroleum gas, and electricity may be of much bigger scope and significance.

Despite high-profile debates about measures to reduce the use of solid fuels, careful evidence on the health impact of their use is in fact scarce. Duflo et al. (2008) review the literature on the health effects of indoor air pollution, raising the concern that the epidemiological literature relies primarily on observational studies that equate correlation with causation. Individuals that have switched to cleaner fuels may have better respiratory health than those that use solid fuels, but better respiratory health may also be a result of unobservable traits that are not necessarily controlled for in the analysis. Two recent studies seek to control for such factors. Pitt et al. (2010) address the endogeneity of exposure to smoke in a study in Bangladesh, and find that longer cooking times raise the likelihood that an individual will report respiratory symptoms. Making use of panel data, Gajate-Garrido (2012) also finds that cooking with firewood raises the likelihood that Peruvian children report respiratory-related illnesses.

However, Thomas and Frankenberg (2002) point out that self-reported symptoms are plagued with reporting bias since individuals from higher socio-economic groups are not only more likely to have been to the doctor for diagnosis but also more likely to be aware of such symptoms. Appropriate analysis would require a direct measure of lung health. We analyse this issue using the Indonesia Family Life Survey (IFLS), a multi-topic panel survey which collected direct measures of lung capacity from respondents. Solid fuels, almost exclusively firewood, are an important source of cooking fuel in Indonesia, used by 39 per cent of all households and 64 per cent of rural households in our sample.[1]

The primary question we consider is whether individuals that live in households that cook with

[1] These numbers are lower than those in (WHO, 2010), which reports that 55 per cent of all households and 77 per cent of rural households cook with a solid fuel in Indonesia. This difference could be due to WHO using a different sample to estimate these numbers.

firewood have poorer respiratory health than those that cook with cleaner fuels such as kerosene, liquefied petroleum gas, or electricity. We focus our analysis primarily on a direct measure of lung capacity collected by survey enumerators; we also consider respiratory and non-respiratory symptoms reported by survey respondents. The endogeneity of fuel choice is addressed by instrumenting it with the remoteness of a location as measured by the predominant road-type in the community. We argue that although the predominant road-type determines the type of fuel a household chooses, it does not directly affect the respiratory health of its members. Using standard tests, we find our instruments to be valid. We finally use propensity score matching to examine whether switching cooking fuels between IFLS waves affects lung health.

We find that individuals that live in households that cook with firewood have 11.2 per cent lower lung capacity measurement compared to those cooking with cleaner fuels. We also find a strong relationship between cooking with firewood and the self-reported symptoms of coughing or difficulty breathing experienced by the individual in the previous four weeks. We find that non-respiratory symptoms such as stomach ache and toothache are not related to cooking with firewood, discarding a spurious correlation between cooking with firewood and respiratory health. These results are corroborated by our analysis of the change in lung capacity between IFLS waves. We find that switching to a cleaner cooking fuel statistically significantly improves the lung capacity of individuals, whereas switching to a dirtier fuel worsens lung capacity.

The rest of this paper is organised as follows. The next section reviews the relevant literature chiefly in epidemiology and economics. Section 3 describes the IFLS dataset and presents characteristics of the sample that we analyse. The empirical strategy for the subsequent analysis is presented in section 4, while section 5 discusses the validity of the instrumental variables, the main results, and robustness checks. The final section discusses policy implications of the results.

2 Relevant literature

Large numbers of households in developing countries cook with solid fuels such as firewood, charcoal, and other forms of biomass because these are often more abundant than fuels they need to procure from the market. Despite their ease of access and affordability, solid fuels have many adverse consequences. Rural residents, often women or children, spend large amounts of time on firewood collection, taking time away from income-generating activities or schooling. Excessive demand for firewood can result in depletion of local forests. More importantly, noxious gases and particles emitted during the combustion of solid fuels could adversely affect health. Early evidence of this came from epidemiologists but economists have recently contributed to this literature as a part of the effort to understand the determinants of wellbeing beyond income.

The combustion of cooking fuels releases harmful toxins that are inhaled by individuals; these toxins have potentially adverse health consequences (WHO (2006); Ezzati and Kammen (2002)). Solid cooking fuels have lower combustion efficiency than kerosene or liquefied petroleum gas and as a result, emit airborne particulate matter and harmful gases such as carbon monoxide and formaldehyde in levels far higher than those recommended by the World Health Organization (Bruce et al., 2002). Indoor air pollution has been associated primarily with respiratory ailments such as asthma and chronic obstructive pulmonary disease, as well as non-respiratory ailments such as cardiovascular disease, low birth weight, and cataract (Fullerton et al., 2008). No clear evidence exists on whether genetic or environmental factors are more important in the incidence of respiratory diseases, although a plausible but complex relationship exists between these two factors (Garantziotis and Schwartz, 2010). The strength of the epidemiological literature is its emphasis on accurate measurement of indoor air pollution and health consequences; its weaknesses are small sample sizes and insufficient attention to the endogeneity of exposure to indoor pollutants (Duflo et al. (2008); Pitt et al. (2010)). The epidemiological literature relies on both self-reported symptoms and direct measures of lung functioning to assess respiratory health, although there is no unanimous measure used in the literature (Mengersen et al. (2011); Kurmi et al. (2012)). Peak expiratory flow (PEF), forced expiratory volume in one second (FEV1), and forced vital capacity (FVC) are the most commonly used measures of lung functioning.

Duflo et al. (2008) note that many studies have found a strong relationship between cooking fuel and respiratory health but most of them tend to be observational. All of them control for some confounding variables such as age, gender, and location. However, they rarely question whether this relationship is causal. Endogeneity arises from the concern that choice of cooking fuel may be correlated with unobserved health behavior that also affects health outcomes. This omitted variable problem can bias the coefficient estimates. Better respiratory health in households that cook with cleaner fuels may be due to better access to information about improving health or other unobservable health behavior. Some studies have used quasi-experimental methods and field experiments to address the endogeneity of exposure to indoor air pollutants. Pitt et al. (2010) address this challenge by instrumenting the exposure to smoke with the relationship hierarchy within the household, arguing that senior spouses are primarily responsible for cooking in Bangladeshi households. They find that greater cooking time raises the likelihood of experiencing respiratory illness. Using a two-wave panel survey of Peruvian children younger than six years old, Gajate-Garrido (2012) examines if children are more likely to report respiratory illness if their household cooks with firewood. She finds that even after the inclusion of individual fixed effects and a variety of confounding variables, children are more likely to report respiratory illnesses if their household cooks with firewood.

Recent literature on this issue focuses on modern cookstoves that still use firewood but produce less smoke than conventional stoves (Hanna et al. (2012); Lewis and Pattanayak (2012)). Several of these have used randomised control trials to examine whether cooking with modern cookstoves improves health outcomes. However, encouraging households to switch to cleaner fuels such as kerosene, liquefied petroleum gas, or electricity may have potentially bigger welfare gains for households.

We contribute to the literature in several ways. First, we analyse the respiratory health impact of cooking fuel using lung capacity measurement, an objective measure of lung health. This allows us to compare this objective measure with self-reported measures of respiratory health. Second, we address the endogeneity of fuel choice using the instrumental variables (IV) method using a rich panel dataset. Third, we examine the heterogeneous impact of cooking with firewood on respiratory health within the household.

3 Data and descriptive statistics

We conduct our analysis using data from the Indonesia Family Life Survey (IFLS), of which we initially use the 1997, 2000, and 2007 waves. IFLS was conducted in 13 out of 26 provinces and is representative of about 83 per cent of the Indonesian population (Strauss et al., 2009). The first wave in 1993 was conducted on 7,224 households and 22,019 individuals. The survey attempted to follow these individuals as well as those that were born into or joined these households. If respondents split off to a new household, they were also interviewed along with their spouses and biological children. Over 13,536 households and 44,103 individuals were interviewed in the 2007 wave. The survey has a remarkably low attrition rate, with 88 per cent of the 16,510 adult respondents from the 1993 wave interviewed in 2007 (Strauss and Thomas, 2008).

Our primary outcome measure of lung health is Peak Expiratory Flow (PEF). PEF gives a reading of a persons maximum speed of expiration in litres per minute and indicates whether there are any obstructions in the respiratory pathways. IFLS administered PEF with the Personal Best peak flow meter made by Phillips on all individuals that were nine years or older at the time of the survey. Each respondent was asked to perform the test three times of which we use the top score in our analysis.[2] The range of the peak flow meter used in IFLS is 0-999 litres/minute.

The average lung capacity of all individuals in our sample is 322 litres/minute (Table 1), and it is much higher for men (380 litres/minute) than it is for women (268 litres/minute). These numbers are much lower than range of the reference PEF values reported by Nunn and Gregg (1989) for normal non-smoking UK adult men and women over the lifetime (respectively 500-700 and 400-600 litres/minute).[3] As an alternative outcome, we examine whether an individual reported symptoms of cough or difficulty breathing in the previous four weeks. Slightly more than a third of all individuals reported coughing or difficulty breathing over this period.

Table 1: Summary statistics (IFLS waves 1997, 2000, and 2007)

Variable	**Mean**	**Std. Dev.**
<u>Respiratory health variables (adults >14yrs)</u>		
Lung capacity measurement in liters per minute	321.650	112.284
Reported cough or difficulty breathing during last 4 weeks	0.363	0.481
<u>Household variables</u>		
Household cooks with firewood	0.392	0.488
Urban	0.493	0.500
Kitchen is outside	0.307	0.461
Same room used for cooking and sleeping	0.058	0.234
Real per capita expenditure, '000 Rupiah	125.851	247.581
<u>Other individual variables</u>		
Female	0.516	0.500
Age, years	27.599	19.670
Height, meters	1.417	0.262
Education, years of school completed	6.104	7.684
Individual is a current smoker	0.293	0.455
Observations	111,714	

Notes: Notes: These statistics are based on pooled data for IFLS waves 1997, 2000, and 2007. The sample consists of 38,000 unique individuals older than 15 years 14,392 of them were interviewed in all three waves, 9,754 individuals were interviewed twice, and 13,393 individuals were interviewed only once. 15,135 unique households were interviewed during the survey. Real per capita expenditure is in August 1997 Rupiah.

Our primary explanatory variable is a dichotomous variable representing the primary cooking fuel used by the household: 1 if firewood (plus the very few cases which use charcoal; we henceforth for convenience refer to this as firewood) and 0 if kerosene, liquefied petroleum gas, electricity, or do not cook.[4] The distinction between firewood and other fuels is important because households in subsistence economies generally use firewood and other biomass for cooking but switch to market-based fuels as their income increases and more fuels become accessible to

[2]The top score is a sensible choice since PEF is a measure of the highest speed at which an individual can exhale.

[3]This gap is plausible because the average British person is most likely taller and healthier than the average Indonesian person.

[4]Burning charcoal could be potentially as harmful as burning firewood. However, only 0.23 per cent of our sample cooks with this fuel. In our section on robustness checks we find that our results are not sensitive to categorizing charcoal and firewood as the same fuel type.

them. Nearly 40 per cent of households in our sample cook with firewood (Table 1). Although the choice of main cooking fuel is not a direct measure of indoor air pollution, it is reasonable to assume that households that cook with firewood will be exposed to higher levels of indoor pollutants than households that cook with cleaner fuels. Figure 1 plots the kernel density of lung capacity of individuals that live in households that cook with firewood and compared to those using other fuels. We see that the distribution of lung capacity of the firewood users lies to the left of those using other fuels throughout the full range, suggesting that cooking with firewood is associated with lower lung capacity. This result is confirmed by Table 2, which compares various characteristics of individuals that cook with firewood with other respondents. We see that the average lung capacity of individuals that live in households cooking with firewood is lower than others (331 versus 308 litres/minute) and that this difference is statistically significant.

Figure 1: Lung capacity of individuals in households that cook with solid vs. non-solid fuels

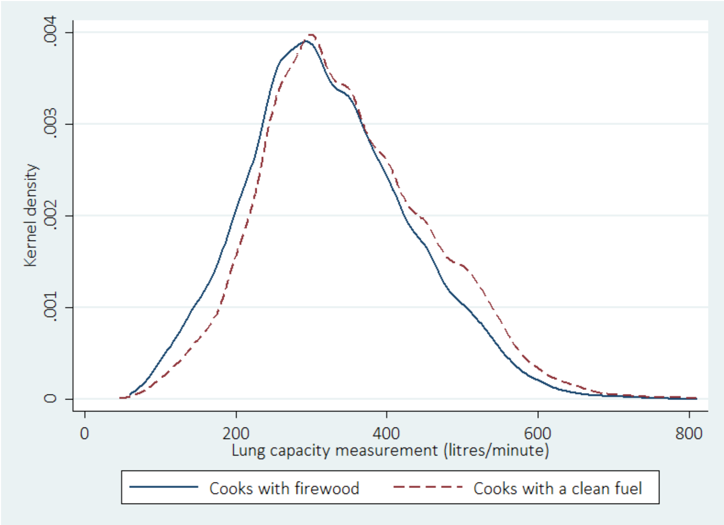

Table 2: Comparison of individuals whose households cook with firewood versus cleaner fuels

	All individuals	Cooks with non-solid fuel (a)	Cooks with solid fuel (b)	(a) - (b)
Variable	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)	Mean (t-stat)
Lung capacity measurement (litres/minute)	321.650	330.877	307.913	22.964
	(112.284)	(113.638)	(108.799)	(28.817)
Reported cough or difficult breathing	0.344	0.355	0.328	0.027
	(0.475)	(0.478)	(0.469)	(9.163)
Real per capita household expenditure	125.851	158.067	76.329	81.738
('000 Rupiah)	(247.581)	(301.350)	(108.989)	(0.054)
Years of education	6.104	7.214	4.368	2.846
	(7.684)	(8.509)	(5.754)	(59.288)
Urban	0.493	0.696	0.177	0.518
	(0.500)	(0.460)	(0.382)	(196.145)
Age	27.599	26.662	29.049	-2.386
	(19.670)	(18.653)	(21.069)	(19.831)

We control for a variety of confounding variables at the individual, household, and community levels. Existing literature suggests that household income is an important determinant of whether it moves away from firewood (Malla and Timilsina, 2014). We do not control for household per capita expenditures in our model because this variable is likely to be endogenously determined with lung health.[5] Housing conditions could also determine exposure to smoke

[5] However, we do include household per capita expenditure in our model as a robustness check. We find that

and thus respiratory health. About a third of all households have an outside kitchen, which is expected to reduce exposure to smoke relative to an indoor kitchen. About six per cent of households cook and sleep in the same room, which likely increases exposure to indoor smoke. Taller individuals typically have higher lung capacity than shorter individuals (Nunn and Gregg, 1989. Respiratory health likely has a non-linear relationship with age; childrens lungs are not as developed as adults, but the elderly are likely to have poorer lung health than adults due to the accumulated effect of age and pollution.

Table 2 compares mean characteristics of users of firewood with other individuals. We see that individuals that live in households that cook with firewood not only have lower lung capacity but also are more likely to have lower household per capita income. They are likely to be less educated and less likely to be living in urban areas. Are individuals that live in a household that switched cooking fuel between IFLS waves different from those that did not? Table 3 compares mean characteristics of individuals whose household either switched to a cleaner fuel (from firewood to another fuel), switched to a dirtier fuel (a clean fuel to firewood), or used the same fuel as in the previous wave. We see that switchers to a clean fuel did not necessarily have a higher lung capacity in the previous wave. This suggests that switchers to clean fuels may not necessarily be ex-ante healthier than switchers to firewood. However, switching to a clean fuel improved lung capacity while switching to firewood from a clean fuel lowered the lung capacity. We also see that switchers to cleaner fuel are more likely to be urban dwellers and more educated. Perhaps most importantly, they are more likely to have seen a bigger change in household income.

Table 3: Comparison of switchers versus non-switchers of cooking fuel between IFLS waves

	Switched to a dirty fuel (a)	Did not switch fuel (b)	Switched to a clean fuel (c)	(a) - (b)	(c) - (b)
Variable	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)	Mean (t-stat)	Mean (t-stat)
Change in lung capacity	6.025	14.760	21.254	-8.735	6.494
	(94.461)	(93.289)	(93.533)	(4.507)	(4.143)
Lung capacity in previous wave	319.022	319.314	310.825	-0.292	-8.489
	(106.192)	(109.538)	(105.280)	(0.132)	(5.028)
Per cent change in real per capita expenditure since last wave	35.305	48.196	94.141	-12.892	45.945
	(111.385)	(140.896)	(231.059)	(5.723)	(22.301)
Lagged per capita HH expenditure	90.216	124.797	82.803	-34.581	-41.994
	(89.301)	(343.329)	(89.225)	(6.479)	(9.829)
Average education in HH, years	5.893	7.227	6.660	-1.335	-0.567
	(5.590)	(8.231)	(8.282)	(10.188)	(5.187)
Urban	0.314	0.530	0.379	-0.216	-0.151
	(0.464)	(0.499)	(0.485)	(26.848)	(22.977)

4 Identification strategy

Let y_{it} be the respiratory health outcome measured by the lung capacity of individual i at time period t. x_{it} is a dummy variable representing whether the household cooks primarily with a solid fuel. x_{2it} is a vector of covariates that could also affect lung health. If fuel type were randomly distributed across households, we could estimate the marginal effect of fuel choice using the lung capacity by ordinary least squares estimation on the pooled data as:

$$y_{it} = \beta_0 + x_{1it}\beta_1 + x_{2it}\beta_2 + \varepsilon_{it} \quad (1)$$

doing so barely changes the magnitude or statistical significance of the coefficient on firewood. We also find that the coefficient on this variable is not statistically significant.

where observation i=1,2,...,N and t represents IFLS waves. β_1 is our coefficient of interest, and ε_{it} is an independently distributed error term assumed to be normally distributed with zero mean and constant variance. If fuel choice is not randomly distributed across households but is instead allocated due to unobservable reasons such as better access to health information or higher valuation of personal health, then the OLS coefficient of fuel choice would be biased. We control for a variety of household and individual covariates in order to account for other factors that could affect lung health. Even with these controls, unmeasured but time-varying individual and household characteristics affecting fuel choice could still be a concern.

We therefore employ the instrumental variables (IV) technique to overcome this identification challenge by instrumenting fuel choice with the predominant road type in the community. We discuss the validity of our instruments in detail in the results section. We estimate the IV model using two-stage least squares, for which equation (1) is the second stage equation. The first-stage equation is:

$$x_{1it} = \alpha_0 + z_{it}\alpha_1 + x_{2it}\alpha_2 + \mu_{it} \tag{2}$$

where z_{it} is our instrument for the endogenous variable x_{1it}, μ_{1it} is the error term, which we assume to be normally distributed with zero mean and constant variance σ^2. The identification assumption is that the IV estimate of β_1 is an unbiased estimate of the marginal effect of cooking with solid fuels on lung capacity assuming that the relevance condition holds ($\alpha_1 \neq 0$) and that the instrument is orthogonal to the error term in the second stage, i.e. $E(z'\varepsilon) = 0$. Since smoking behavior is also likely to be endogenously determined with lung capacity, we instrument it with the share of smokers in the community.

We next examine if switching cooking fuel affects the change in lung capacity of individuals over time. We could analyse this using an OLS model, but the coefficient from this model would be biased since households may self-select themselves into switchers and non-switchers based on observable or unobservable traits. A randomised control trial may be the ideal way to test if switching cooking fuel affects lung capacity, but this possibility is beyond the scope of our paper. We instead use the quasi-experimental method of propensity score matching (PSM) in order to reduce the selection bias in OLS results by matching the treatment group to a comparison group within the sample of non-participants using the propensity score. The propensity score is the predicted probability of participation given observed characteristics. Although PSM does not completely solve the problem of selection bias, it focuses the researcher on the model of treatment assignment, which is better understood than the more complex process by which outcomes are determined (Angrist and Pischke, 2009). Rosenbaum and Rubin (1983) propose that under the assumptions of selection on observables and the common support condition, the difference between the mean outcomes for treatment and control groups at each level of the propensity score provides an unbiased estimate of the average treatment effect on the treated (ATT).[6] In our case, PSM will give us an estimate of the impact of switching cooking fuel between IFLS waves on individuals lung capacity.

We then conduct various robustness checks in order to get assurance that our main results are not sensitive to various assumptions. First, we verify that our results are not sensitive to the precise definition of cooking fuel used in our paper. Second, we check to see if changing the specification of our main regression substantively affects the primary result. Third, we examine the impact of cooking with firewood on self-reported respiratory symptoms; we also examine non-respiratory symptoms to check for possible spurious correlation. We then examine the inclusion of smoking as a possible endogenous determinant of lung health. Finally, following Pitt et al. (2010), we restrict the female sub-sample to women who are likely to cook within the household the wife of the household head and daughters-in-law.

[6]The selection on observables (also called unconfoundedness) condition states that conditional on the covariates, the assignment to treatment is independent of the outcomes. The common support condition (also called the overlap condition) states that the probability of assignment to treatment is bounded between zero and one.

5 Results and robustness checks

5.1 Validity of the instruments

We instrument the households choice of cooking fuel with the predominant road type in the community. The most commonly present road type captures the remoteness of a community and is an important determinant of fuel choice because households need a steady supply of a fuel in order to use it as the primary source of cooking fuel. If infrastructure to deliver market-based fuels is weak, households may prefer to stick with firewood that is more readily available locally.

Table 4: First stage results of IV estimation

	(1)
Dependent variable →	Household cooks with firewood
Female	-0.020^{***}
	(0.005)
Urban	-0.362^{***}
	(0.013)
Number of health posts in community	-0.003^{***}
	(0.001)
Height, meters	-0.137
	(0.026)
Age	0.000
	(0.001)
Squared age	0.000
	(0.000)
Cooks and sleeps in the same room	-0.038^{*}
	(0.020)
Kitchen is outside	0.031^{***}
	(0.011)
Average household education (years)	-0.038^{***}
	(0.002)
Distance to Kabupaten capital	-0.013^{***}
	(0.005)
Predominant road-type (Reference category: Asphalt)	
Dirt	0.212^{***}
	(0.019)
Gravel	0.112^{***}
	(0.025)
Constant	1.028^{***}
	(0.039)
Observations	23,560
R^2	0.308
F-test on instrument	68.356

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4 presents estimates from the first stage of the IV estimation in which we see that the predominant road type in the community is a strong predictor of a households choice to cook with firewood. Relative to asphalt roads, gravel and dirt roads raise the probability that a household cooks with firewood. The coefficient on the road variable is significant at the 1 per cent level and the F statistic in the first stage regression is comfortably above the rule of thumb of 10 used in the literature and the Stock-Yogo critical values at the 10 per cent significance level. In addition to the predominant road-type as the instrument for fuel choice, we used distance to the market (as a proxy for remoteness) and the number of primary schools in the community (as a proxy for health information) as second instruments in our model. While we found these to be good second instruments, they do not substantively change our main findings (see Table

A.1).[7]

In order to use road type as an instrument for fuel choice, we need to be sure that the exclusion restriction holds, i.e. road type does not directly affect lung health after controlling for all other covariates. Although this assumption is not directly testable, it is plausible in our case since lung capacity is determined primarily by individual characteristics and environmental factors. While the predominant road type in the community may affect the availability of health facilities and income generation opportunities that in turn could affect lung health, we control for remoteness and availability of healthcare among factors directly in our model by including distance to the kabupaten (district) capital and the number of health posts present in the community. The exclusion restriction would be violated if individuals of varying lung capacities sorted themselves into different road types. While individuals may move towards areas with a better economic environment, there is no reason to think that they would move to an area primarily because of the predominant road type.

5.2 Main results

Table 5 presents our estimation results of the determinants of the log of lung capacity measurement. Column 1 presents the baseline results from pooled OLS estimation, column 2 presents results from the fixed-effects estimation, and column 3 addresses the endogeneity of firewood by instrumenting it. We only use the final 2007 wave for IV estimation since we believe that the quality of data on lung capacity measurement, our instrument for firewood, and other survey data in this wave are better than those in previous waves. The first row contains coefficients on whether the household cooks with a solid fuel, our main variable of interest. We see that the coefficient on firewood is negative and significant at least at the 5 per cent level in the OLS, fixed-effects, and instrumental variable models. According to the OLS estimates, cooking with firewood lowers lung capacity by 1.0 per cent. The fixed-effects model improves upon OLS estimation by controlling for unobservable time-invariant characteristics of individuals, and finds an impact of firewood of 1.6 per cent.[8] But we find that this coefficient increases to 11.2 per cent once we account for the endogeneity of firewood in column 3.[9]

The IV estimator probably better captures the full effect of cooking with firewood, having addressed the endogenous nature of the choice of cooking fuel. The fact that the OLS estimates are lower than the IV estimates suggests two things. First, unobservable factors such as lack of knowledge about health that lower lung capacity could have a positive effect on the likelihood that a household cooks with solid fuel. Second, our measure for cooking fuel may have measurement error. Having removed the attenuation bias, its coefficient may be bigger than the OLS coefficient. The IV estimate shows that cooking with firewood has a large negative impact on lung capacity.[10]

Gender has a large coefficient in all of the specifications, which makes intuitive sense since women have a lower body size on average and are more likely to be exposed to indoor smoke

[7] The p-value of the Sargan-Hansen J-statistic for the overidentification test of all instruments for our model with the distance to market as the second instrument is 0.112, suggesting that we cannot reject the null hypothesis that the instruments are valid, i.e. uncorrelated with the error term in the second stage. The similar p-value for the number of primary schools in the community as the second instrument is 0.0724.

[8] Results from the Hausman test rejected the null hypothesis that the difference in coefficients in the fixed-effects and random-effects models is not systematic, which is why we chose the fixed-effects estimates.

[9] We can rule out that this difference is due to the change in sample between OLS/FE and IV specifications. We get similar results even when we estimate these models on the same sample.

[10] It remains possible that the larger IV coefficient reflect measurement error in the solid fuel variable, given that households may use a portfolio of fuels for their cooking needs while the survey only reports on the primary cooking. For example, they may use firewood as their primary cooking fuel, but use a kerosene stove to warm up a meal or to make snacks because kerosene stoves are more convenient but expensive option.

Table 5: Determinants of the log of lung capacity measurement

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator →	OLS	FE	IV	IV	IV	IV
Sample →	Full	Full	Full	Women	Men	Children
Household cooks with solid fuel	-0.010^{***}	-0.016^{**}	-0.112^{***}	-0.132^{**}	-0.089^{*}	-0.182^{**}
	(0.003)	(0.008)	(0.041)	(0.054)	(0.048)	(0.090)
Female	-0.237^{***}		-0.299^{***}			-0.121^{***}
	(0.003)		(0.005)			(0.009)
Urban	-0.009^{**}	0.041^{**}	-0.055^{***}	-0.075^{***}	-0.033	-0.079^{**}
	(0.004)	(0.016)	(0.017)	(0.022)	(0.020)	(0.036)
Number of health posts in community	0.000	-0.000	0.001^{*}	0.001	0.001	0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Height, meters	1.005^{***}	0.911^{***}	0.599^{***}	0.384^{***}	0.705^{***}	0.490^{***}
	(0.026)	(0.061)	(0.032)	(0.039)	(0.045)	(0.073)
Age	0.024^{***}	0.047^{***}	0.030^{***}	0.025^{***}	0.035^{***}	0.025
	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)	(0.039)
Squared age	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	0.002
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Cooks and sleeps in the same room	-0.018^{***}	-0.001	-0.029^{***}	-0.044^{***}	-0.011	-0.033
	(0.006)	(0.011)	(0.010)	(0.012)	(0.012)	(0.022)
Kitchen is outside	0.014^{***}	0.013^{**}	0.028^{***}	0.033^{***}	0.023^{***}	0.021^{*}
	(0.003)	(0.005)	(0.005)	(0.006)	(0.006)	(0.011)
Average household education (years)	0.005^{***}	0.002^{***}	0.006^{***}	0.005^{**}	0.007^{***}	-0.001
	(0.000)	(0.001)	(0.002)	(0.002)	(0.002)	(0.006)
Distance to Kabupaten capital	0.010^{***}	0.004	0.014^{***}	0.014^{***}	0.013^{***}	0.006
	(0.003)	(0.005)	(0.003)	(0.004)	(0.004)	(0.009)
Constant	2.739^{***}	29.943^{***}	4.553^{***}	4.681^{***}	4.267^{***}	4.364^{***}
	(0.606)	(4.059)	(0.062)	(0.080)	(0.077)	(0.248)
Observations	60,567	60,567	23,560	12,319	11,241	3,292
R^2	0.504	0.310	0.480	0.257	0.412	0.312
First-stage F statistic			68.356	53.128	68.665	21.983

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All standard errors are clustered at the household-level. The adult sub-sample includes all individuals in the survey that are 9 years and older. The child subsample includes children that are between 9 and 15 years old. Models (1) and (2) were run on the 1997, 2000, and 2007 waves and also include wave fixed effects. Models (3)-(6) were run on the 2007 wave. The reported F-statistic from the first stage is the Kleibergen-Paap Wald F rk statistic, which is robust to the presence of clustering in the data.

while cooking or conducting household chores near the cookstove. The fact that the coefficient on gender is large and significant suggests that we may need to model male and female subsamples separately. Columns 4 and 5 present results for gender subsamples. The results show that the coefficient on firewood is not only statistically more significant but also has a larger magnitude (almost 50 per cent higher) for women than that for men. Column 6 contains results from the regression on the subsample that is between 9 and 14 years old. The results show that the coefficient on firewood on this subsample is -0.180, which is larger in absolute value than for the male and female subsamples. Taken together, these results suggest that cooking with firewood has a large and negative impact on the lung health of all individuals, but in particular, women and children. These results are also consistent with the public health literature on this topic (Smith et al. (2013); Martin et al. (2011)).

Education, height, and housing characteristics (outside kitchen and same room for cooking and sleeping) all have the expected signs. There is no clear a priori reason why urban or rural households should have better respiratory health. While urban areas are more densely populated and have higher levels of production activity that result in higher levels of ambient air pollution, rural residents are more likely to use firewood that potentially worsens respiratory health. Urban residents generally have lower lung capacity than rural residents, although this coefficient is not consistent across specifications. Age has a positive but declining effect on lung capacity, which is consistent with the fact that adults have lungs that are more developed than those of children. However, the lung capacity of adults falls as they get older and their lung health worsens with higher accumulated exposure to indoor and outdoor pollution.

We next consider whether switching cooking fuel affects lung health, as was tentatively suggested in the descriptive analysis above. Although 81.3 per cent of respondents never switched the primary cooking fuel, 11.4 per cent switched to a cleaner fuel while the remaining 7.3 per cent switched to a dirtier fuel at some point during the survey. We begin our analysis with an OLS regression of change in lung capacity measurement on whether the household switched cooking fuel since the previous wave (see Table A.2). These show that switching from a firewood to other fuels improved lung capacity by about 1 per cent, ceteris paribus. Contrary to this, an individuals lung capacity shrinks by about 2 per cent if the household switched the cooking fuel from cleaner fuels to firewood. Both estimates are statistically significant at the 5 per cent level or higher.

However, as noted above the OLS estimates are likely to be biased in this case so instead, we relied on PSM to compare outcomes of switchers to the counterfactual.[11] PSM creates a counterfactual from the predicted probabilities, based on observable characteristics, of being in treatment or control group. We used various sets of covariates for this model but the results were not substantively different. The figures comparing the conditioning variables for matched (using kernel matching) and unmatched samples reveal that the matched samples have much more similar sample means than the unmatched samples (see Online Appendix Figure 1). The average bias between the treated and untreated samples is less than 5 per cent in all the cases, thus passing the balancing test. The probit regressions used to calculate the propensity scores show that the primary determinant of switching are education and urban location (Table A.3). Results in Table 6 show that switching cooking fuel affects lung health even during a short period of a few years. Switchers to a clean fuel between consecutive IFLS waves saw their lung capacity measurement improve by 1.9 per cent while switchers to a dirty fuel saw it worsen by 1.9 per cent. This result is consistent with the descriptive analysis above. We attempted PSM with nearest neighbour matching and radius matching, both of which gave similar results.

Table 6: PSM estimates. Dependent variable – change in the log of lung capacity measurement

	Switched to a clean fuel		Switched to a dirty fuel	
Statistic	Before Matching	Kernel Matching	Before Matching	Kernel Matching
Mean, treated group	0.042	0.042	0.003	0.003
Mean, control group	0.014	0.023	0.021	0.022
Difference	0.028	0.019	-0.018	-0.019
t-stat	4.130	1.800	-2.470	-2.230
Mean standardized bias	13.700	3.500	13.000	1.700
Pseudo R^2	0.025	0.002	0.024	0.001
Observations	10,712	10,712	12,352	12,352
Common support imposed		Yes		Yes

5.3 Robustness checks

In this section, we discuss various robustness checks we conducted to ensure that the results in the previous section are not sensitive to various assumptions and definitions we have used in our

[11] We attempted to instrument the choice of cooking fuel with the main road type in the community as well as the change in road type, but the instrument did not pass the weak-identification test.

analysis. First, we verify that our results are not sensitive to the definition of cooking fuel used in our paper. We consider the effect of defining the dependent variable to include those only using firewood, so excluding the 0.23 per cent of households using charcoal, and adding these into the clean fuel category. The results using this alternative definition are virtually identical to those in Table 5 (see Online Appendix Table 2). We then only keep households that reported to be cooking with firewood, kerosene, liquefied petroleum gas, or electricity (98.9 per cent of our sample). The results are also very similar to the results in Table 5 (see Online Appendix Table 3). These suggest that our findings are not sensitive to changes in our definition of cooking fuel.

Second, since we do not have clear guidance from theory on the determinants of respiratory health, we progressively add covariates to a core model to test whether changing the specification substantively affects the results (see Online Appendix Table 1). Our results show that the magnitude, sign, and significance of the coefficient on the households choice of cooking fuel does not change substantively across specifications. All of the results are negative, significant at the 1 per cent level, and range between 0.06 and 0.13, with the coefficient in the IV model being 0.11. These results suggest that the estimate of the coefficient on fuel choice is robust to the choice of covariates in the model.

Table 7: Determinants of respiratory symptoms (IV probit estimates)

	(1)	(2)	(3)	(4)	(5)	(6)
Symptom →	Cough or difficulty breathing			Stomach-ache		
Sample →	Women	Men	Children	Women	Men	Children
Household cooks with solid fuel	0.509**	−0.071	0.369**	−0.235	−0.214	0.347
	(0.218)	(0.132)	(0.187)	(0.263)	(0.155)	(0.233)
Urban	0.353***	0.059	0.325***	0.077	0.043	0.311***
	(0.094)	(0.058)	(0.085)	(0.117)	(0.069)	(0.106)
Number of health posts in community	−0.003	−0.003**	0.001	−0.004*	−0.006***	−0.006**
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)
Age	0.005***	0.005***	−0.045***	−0.005***	−0.000	0.008
	(0.001)	(0.000)	(0.004)	(0.002)	(0.000)	(0.005)
Cooks and sleeps in the same room	0.084*	0.056**	−0.051	0.039	0.064**	0.041
	(0.046)	(0.027)	(0.042)	(0.052)	(0.031)	(0.051)
Kitchen is outside	−0.053**	−0.042***	−0.074***	−0.038	−0.011	−0.045
	(0.025)	(0.015)	(0.023)	(0.029)	(0.017)	(0.028)
Average household education (years)	0.005	−0.003	0.017***	−0.004	−0.001	0.005
	(0.005)	(0.003)	(0.003)	(0.006)	(0.004)	(0.004)
Distance to Kabupaten capital	−0.013	−0.039**	−0.088***	−0.036	−0.048**	−0.039
	(0.028)	(0.017)	(0.025)	(0.033)	(0.021)	(0.031)
Female			0.004			−0.037
			(0.020)			(0.025)
Constant	−0.950***	−0.383***	−0.407***	−0.477***	−0.803***	−1.361***
	(0.117)	(0.099)	(0.133)	(0.159)	(0.119)	(0.149)
Observations	13,531	39,218	15,587	13,531	39,218	15,587

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We next examine an alternative measure of lung health self-reported respiratory symptoms. This variable is commonly used in the literature (Pitt et al., 2010; Hanna et al. 2012). Table 7 presents results from regressions in which the dependent variable is whether the respondent reported experiencing cough or difficulty breathing symptom as well as a non-respiratory symptom (stomach-ache). We find results consistent with results in Table 5. Women and children are much more likely to report cough symptoms if they live in a household that cooks primarily with firewood. We do not find a similar result with the male sub-sample. We do not find that individuals living in a household that cooks with firewood report higher non-respiratory symptoms (stomach-ache).[12] This result rules out the possibility that the results in columns 1-3 are

[12]We found a similar result for toothache symptoms. This result is available upon request.

due to spurious correlation.

Smoking is of course a potentially important determinant of lung health. Smoking is widely prevalent in Indonesia, particularly among men, and the survey gives information on the number of cigarettes smoked per day. In adding this variable to the model it is important to allow for its endogeneity. We therefore instrument the number of cigarettes smoked by an individual per day with the share of smokers in the community (excluding the individual in question), which we find to be a strongly correlated with our smoking variable. The results in the last column of Table A.1 do indeed show a significant negative impact of smoking on lung health; but even when allowing for this we find that the relationship between solid cooking fuel and lung capacity persists, in fact at a slightly higher level. Finally, following Pitt et al. (2010), we restrict the female sub-sample the wife of the household head and daughters-in-law. We find that the coefficient on firewood for this subsample is -0.153 and is significant at the 5 per cent level, which is larger in magnitude than the coefficient of -0.134 for the full female subsample (see Online Appendix Table 4). This suggests that the impact of cooking with firewood is higher for women who are more likely to be exposed to smoke than others.

6 Conclusion

In this paper, we examined the evidence for whether cooking with firewood causes poorer respiratory health, as has been strongly claimed in the international policy discussion. Using a unique dataset from Indonesia, we used direct measures of lung capacity and addressed the endogeneity of fuel choice of households. We believe this study to be the first to address these two important issues together and therefore offers more robust results. We found that use of firewood is associated with poorer respiratory health measured by lung capacity as well as self-reported respiratory symptoms. We found stronger evidence of this relationship for women and children than for men, and a stronger relationship for women most likely to cook in the household. In addition to these results, we found that switching to a cleaner cooking fuel improves lung capacity whereas switching to a dirtier one lowers it. These latter results corroborate our main findings that suggest a harmful effect of cooking with a solid fuel on lung capacity. One surprising finding of this paper is the extent of downward and upward switching of cooking fuel. As households fortunes change over the lifecycle, they are also likely to switch to cleaner or dirtier cooking fuels. This may have significant impacts on their respiratory health.

The data set and measures used here are superior to most other sources for addressing the question. It is true it did not have information on time spent cooking or on the level of indoor air pollution, which would have aided more accurate analysis of intra-household differences in the health effects of fuel choice. Outdoor air pollution may also be a relevant factor. While having this information could strengthen future findings, it remains the case that this paper finds strong and robust evidence on the large adverse effect of cooking with firewood.

The main policy implication that emerges from this study is that households need to be encouraged to switch from cooking with firewood to cleaner fuels. As well as being beneficial in its own terms, better respiratory health is known to be associated with higher levels of production (Duflo et al. (2008); Strauss and Thomas (2008)). In addition, Dupas (2011) notes that although households in developing countries spend significant resources on curative health measures, they spend relatively little on preventive health measures, of which using cleaner cooking fuels is one. She argues this is because individuals do not always have full information on the costs and benefits of alternative preventive measures and market imperfections often prevent them from making appropriate health investments. These suggest a potential role for public policy. Recent policy debates around indoor air pollution have revolved mostly around improved cooking stoves that burn firewood more efficiently to reduce indoor smoke. Cookstoves are only one possible response; the priority is to encourage switching to cleaner fuels may improve the respiratory health of individuals.

Despite the finding that switching to a cleaner cooking fuel will likely improve respiratory health, this paper is agnostic about policies that can nudge households to switch to cleaner cooking fuels. Steady and reliable availability of kerosene, liquefied petroleum gas, or electricity could encourage households to move away from firewood and charcoal. This would imply higher investment in infrastructure to transport fuel such as roads, railways, and power lines. Availability of alternative sources of cooking fuel, including biogas and solar power should be in the repertoire of policymakers, although it is acknowledged that different locations may need different types of interventions.

References

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton, NJ: Princeton University Press. 8

Bruce, N., Perez-Padilla, R., and Albalak, R. (2002). *The health effects of indoor air pollution exposure in developing countries*, volume 11. Geneva: World Health Organization. 3

Duflo, E., Greenstone, M., and Hanna, R. (2008). Indoor air pollution, health, and economic well-being. *S.A.P.I.E.N.S. Revues.org.*, 1(1):7–16. 2, 4, 14

Dupas, P. (2011). Health behavior in developing countries. *Annual Review of Economics*, 3:425–449. 14

Ezzati, M. and Kammen, D. M. (2001). Indoor air pollution from biomass combustion and acute respiratory infections in kenya: An exposure-response study. *Lancet*, 358(9282):619–624. 2

Ezzati, M. and Kammen, D. M. (2002). The health impacts of exposure to indoor air pollution from solid fuels in developing countries: Knowledge, gaps, and data needs. *Environmental Health Perspectives*, 110(11). 3

Fullerton, D. G., Bruce, N., and Gordon, S. B. (2008). Indoor air pollution from biomass fuel smoke is a major health concern in the developing world. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 102(9):843–851. 3

Gajate-Garrido, G. (2012). The impact of indoor air pollution on the incidence of life threatening respiratory illnesses: Evidence from young children in peru. *The Journal of Development Studies*, pages 1–16. 2, 4

Garantziotis, S. and Schwartz, D. A. (2010). Ecogenomics of respiratory diseases of public health significance. *Annual Review of Public Health*, 31:37–51. 3

Hanna, R., Duflo, E., and Greenstone, M. (2012). Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves. *National Bureau of Economic Research Working Paper Series*, No. 18033. 2, 4

Kurmi, O. P., Devereux, G. S., Smith, W. C. S., Semple, S., Steiner, M. F., Simkhada, P., Hubert Lam, K.-B., and Ayres, J. G. (2012). Reduced lung function due to biomass smoke exposure in young adults in rural nepal. *The European Respiratory Journal*, 41(1):25–30. 4

Lewis, J. J. and Pattanayak, S. K. (2012). Who adopts improved fuels and cookstoves? a systematic review. *Environmental Health Perspectives*, 120(5):637–645. 4

Malla, S. and Timilsina, G. R. (2014). Household cooking fuel choice and adoption of improved cookstoves in developing countries. *World Bank Policy Research Working Paper*, 6903. 6

Martin, William J., I., Glass, R. I., Balbus, J. M., and Collins, F. S. (2011). A major environmental cause of death. *Science*, 334(6053):180–181. 11

Mengersen, K., Morawska, L., Wang, H., Murphy, N., Tayphasavanh, F., Darasavong, K., and Holmes, N. S. (2011). Association between indoor air pollution measurements and respiratory health in women and children in lao pdr. *Indoor Air*, 21(1):25–35. 4

Miller, G. and Mobarak, A. M. (2011). Intra-household externalities and low demand for a new technology: Experimental evidence on improved cookstoves. 2

Nunn, A. J. and Gregg, I. (1989). New regression equations for predicting peak expiratory flow in adults. *British Medical Journal*, 298(6680):1068–1070. 5

Pitt, M., Rosenzweig, M., and Hassan, N. (2010). Short and long-term health effects of burning biomass in the home in low-income countries. 2, 4, 8

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for casual effects. *Biometrika*, 70(1):41–55. 8

Smith, K. R., Frumkin, H., Balakrishnan, K., Butler, C. D., Chafe, Z. A., Fairlie, I., Kinney, P., Kjellstrom, T., Mauzerall, D. L., McKone, T. E., McMichael, A. J., and Schneider, M. (2013). Energy and human health. *Annual review of public health*, 34:159–88. 11

Smith, K. R., Mehta, S., and Maeusezahl-Feuz, M. (2004). *Indoor air pollution from solid fuel use*, pages 1435–1493. World Health Organization, Geneva. 2

Strauss, J. and Thomas, D. (2008). *Health over the life course*, volume 4, pages 3375–3474. Elsevier. 2, 4, 14

Strauss, J., Witolear, F., Sikoki, B., and Wattie, A. M. (2009). The fourth wave of the indonesia family life survey: Overview and field report. *RAND*, WR-675/1-NIA/NICHD. 4

Thomas, D. and Frankenberg, E. (2002). *The measurement and interpretation of health in social surveys*, pages 387–420. World Health Organization, Geneva. 2

WHO (2006). *Fuel for life: Household energy and health*. World Health Organization, Geneva, Switzerland. 3

WHO (2009). *Global health risks*. World Health Organization, Geneva, Switzerland. 2

WHO (2010). Global health observatory data repository. 2

WHO (2012). *World Health Statistics 2012*. World Health Organization, Geneva, Switzerland. 2

World Bank (2011). *Household cookstoves, environment, health, and climate change*. World Bank, Washington, DC. 2

A Appendix

Table A.1: Determinants of the log of lung capacity measurement (IV estimates)

	(1) Two instruments (Second instrument: Distance to market)	(2) Two instruments (Second instrument: Primary schools in community)	(3) Smoking as a second endogenous variable
Household cooks with solid fuel	−0.108***	−0.127***	−0.149***
	(0.041)	(0.039)	(0.055)
Female	−0.299***	−0.299***	−0.630***
	(0.005)	(0.005)	(0.049)
Urban	−0.054***	−0.061***	−0.071***
	(0.017)	(0.016)	(0.023)
Number of health posts in community	0.001*	0.001	−0.000
	(0.000)	(0.000)	(0.001)
Height, meters	0.600***	0.597***	0.248***
	(0.032)	(0.032)	(0.030)
Age	0.030***	0.030***	0.027***
	(0.001)	(0.001)	(0.002)
Squared age	−0.000***	−0.000***	−0.000***
	(0.000)	(0.000)	(0.000)
Cooks and sleeps in the same room	−0.029***	−0.030***	−0.037***
	(0.010)	(0.010)	(0.011)
Kitchen is outside	0.028***	0.029***	0.016***
	(0.005)	(0.005)	(0.006)
Average household education (years)	0.006***	0.005***	0.003
	(0.002)	(0.002)	(0.002)
Distance to Kabupaten capital	0.014***	0.014***	0.017***
	(0.003)	(0.003)	(0.004)
Cigarettes consumed per day			−0.034***
			(0.006)
Constant	4.549***	4.570***	5.546***
	(0.062)	(0.061)	(0.087)
Observations	23,560	23,560	20,236
R^2	0.481	0.476	0.196
First-stage F statistic	46.700	50.135	24.231
Sargan Hansen J-stat p-value	0.128	0.074	0.006

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In column (1), the second instrument for firewood is the distance to the nearest market. In column (2), the second instrument for firewood is the number of primary schools in the community. See notes to Table 5 for other details.

Table A.2: Determinants of change in log of lung capacity measurement between IFLS waves (OLS estimates)

	(1)	(2)	(3)	(4)	(5)
Switched to a dirty fuel	−0.015**	−0.019***	−0.018**	−0.018***	−0.018***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Switched to a clean fuel	0.024***	0.013**	0.014**	0.013**	0.013**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Female		−0.031***	−0.031***	−0.031***	−0.031***
		(0.004)	(0.004)	(0.004)	(0.004)
Age		−0.006***	−0.006***	−0.006***	−0.006***
		(0.000)	(0.000)	(0.000)	(0.000)
Education, years completed		−0.001*	−0.001*	−0.001*	−0.001*
		(0.000)	(0.000)	(0.000)	(0.000)
Change in per capita expenditure since last wave X 100			−0.002	−0.002	−0.001
			(0.001)	(0.001)	(0.001)
Urban				−0.003	−0.003
				(0.004)	(0.004)
Moved since last survey (1993)					−0.000
					(0.001)
Constant	0.009***	0.304***	0.303***	0.303***	0.303***
	(0.003)	(0.008)	(0.008)	(0.008)	(0.008)
Observations	23,415	23,415	23,066	23,066	23,066
R^2	0.002	0.100	0.099	0.099	0.099

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All specifications include Kecamatan and wave fixed-effects. All standard errors are heteroscedasticity-consistent.

Table A.3: Determinants of the log of lung capacity measurement (IV estimates)

	(1) Switched to a clean fuel	(2) Switched to a dirty fuel
Education, years completed	0.035***	−0.049***
	(0.004)	(0.004)
Urban	0.536***	−0.765***
	(0.034)	(0.033)
Female	−0.031	0.039***
	(0.030)	(0.034)
Age, years	0.000	−0.004
	(0.001)	(0.001)
Per cent change in per capita expenditure	0.001***	0.000
	(0.000)	(0.000)
Moved since previous wave	0.063***	0.000
	(0.008)	(0.007)
Constant	0.370	−1.488
	(0.419)	(0.634)
Observations	10,712	12,352

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Also included in the regressions but not reported wre dummies for the sector of activity and the cooking fuel used in the previous wave.

B Online Appendix (Not for publication)

Table B.1: Determinants of change in log of lung capacity measurement between IFLS waves (OLS estimates)

	(1)	(2)	(3)	(4)	(5)	(6)
Household cooks with solid fuel	-0.077^{***}	-0.062^{***}	-0.130^{***}	-0.115^{***}	-0.117^{***}	-0.112^{***}
	(0.019)	(0.016)	(0.037)	(0.039)	(0.039)	(0.041)
Height, meters		0.626^{***}	0.616^{***}	0.614^{***}	0.614^{***}	0.599^{***}
		(0.032)	(0.033)	(0.033)	(0.033)	(0.032)
Age		0.030^{***}	0.030^{***}	0.030^{***}	0.030^{***}	0.030^{***}
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Squared age		-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female		-0.296^{***}	-0.297^{***}	-0.298^{***}	-0.298^{***}	-0.299^{***}
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Urban			-0.049^{***}	-0.048^{***}	-0.048^{***}	-0.055^{***}
			(0.018)	(0.018)	(0.019)	(0.017)
Distance to Kabupaten capital			0.014^{***}	0.014^{***}	0.013^{***}	0.014^{***}
			(0.003)	(0.003)	(0.003)	(0.003)
Number of health posts in community				0.001^{***}	0.001^{**}	0.001^{*}
				(0.000)	(0.000)	(0.000)
Cooks and sleeps in the same room					-0.033^{***}	-0.029^{***}
					(0.010)	(0.010)
Kitchen is outside					0.028^{***}	0.028^{***}
					(0.005)	(0.005)
Average household education (years)						0.006^{***}
						(0.002)
Constant	5.738^{***}	4.523^{***}	4.587^{***}	4.574^{***}	4.570^{***}	4.553^{***}
	(0.009)	(0.048)	(0.059)	(0.061)	(0.060)	(0.062)
Observations	24,429	24,392	24,392	24,095	24,092	23,560
R^2	0.006	0.493	0.480	0.484	0.486	0.480
First-stage F statistic	392.054	392.487	79.829	70.598	69.811	68.356

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

See Table 5 in main text for notes on the regressions.

Table B.2: Determinants of the log of lung capacity measurement (charcoal as a clean fuel)

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator →	OLS	FE	IV	IV	IV	IV
Sample →	Full	Full	Full	Women	Men	Children
Household cooks with firewood	−0.010***	−0.016**	−0.118***	−0.138**	−0.093*	−0.199*
	(0.003)	(0.008)	(0.044)	(0.058)	(0.052)	(0.103)
Female	−0.237***		−0.299***			−0.121***
	(0.003)		(0.005)			(0.009)
Urban	−0.009**	0.041**	−0.058***	−0.078***	−0.035	−0.085**
	(0.004)	(0.016)	(0.018)	(0.024)	(0.022)	(0.041)
Number of health posts in community	0.000	−0.000	0.001*	0.001	0.001	0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Height, meters	1.005***	0.911***	0.600***	0.384***	0.705***	0.490***
	(0.026)	(0.061)	(0.032)	(0.039)	(0.045)	(0.073)
Age	0.024***	0.047***	0.030***	0.025***	0.035***	−0.004
	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)	(0.040)
Squared age	−0.000***	−0.000***	−0.000***	−0.000***	−0.000***	0.003*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Cooks and sleeps in the same room	−0.018***	−0.001	−0.030***	−0.045***	−0.012	−0.037*
	(0.006)	(0.011)	(0.010)	(0.012)	(0.012)	(0.022)
Kitchen is outside	0.014***	0.013**	0.028***	0.033***	0.023***	0.021*
	(0.003)	(0.005)	(0.005)	(0.006)	(0.006)	(0.011)
Average household education (years)	0.005***	0.002***	0.006***	0.005**	0.007***	−0.002
	(0.000)	(0.001)	(0.002)	(0.002)	(0.002)	(0.006)
Distance to Kabupaten capital	0.010***	0.004	0.014***	0.014***	0.013***	0.006
	(0.003)	(0.005)	(0.003)	(0.004)	(0.004)	(0.009)
IFLS wave	0.001**	−0.013***				
	(0.000)	(0.002)				
Constant	2.747***	29.968***	4.557***	4.685***	4.270***	4.547***
	(0.606)	(4.059)	(0.064)	(0.082)	(0.079)	(0.257)
Observations	60,567	60,567	23,560	12,319	11,241	3,303
R^2	0.504	0.310	0.479	0.254	0.410	0.300
First-stage F statistic			54.454	42.801	54.006	16.209

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In this table, charcoal is not in the same category as firewood; it is instead considered to be a clean fuel. See notes to Table 5 for other details.

Figure B.1: Standardized bias across covariates for matched and unmatched samples

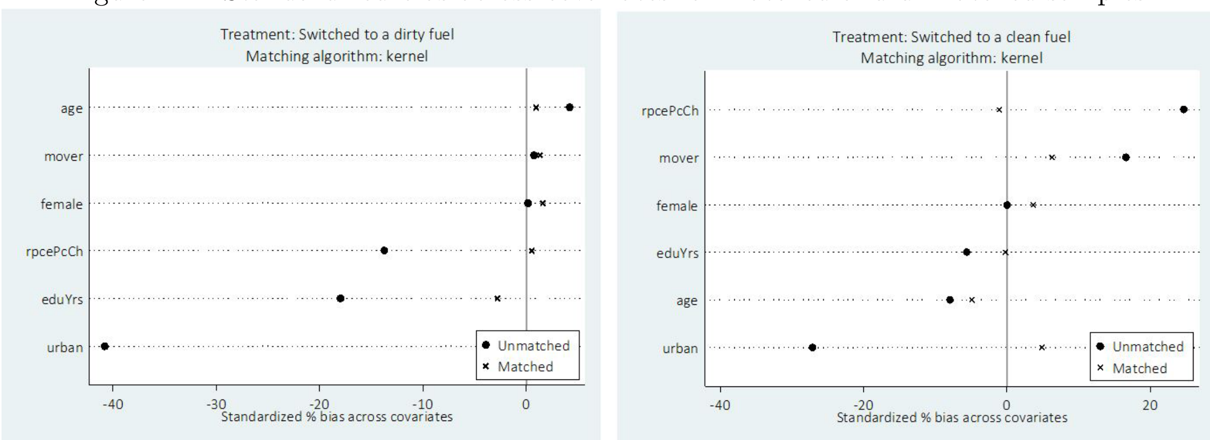

Table B.3: Determinants of the log of lung capacity measurement (restricted categories of cooking fuel)

	(1) OLS	(2) FE	(3) IV	(4) IV	(5) IV	(6) IV
Household cooks with solid fuel	-0.010^{***}	-0.016^{**}	-0.111^{***}	-0.134^{**}	-0.084^{*}	-0.159^{*}
	(0.003)	(0.008)	(0.041)	(0.055)	(0.048)	(0.090)
Female	-0.236^{***}		-0.298^{***}			-0.122^{***}
	(0.003)		(0.005)			(0.009)
Urban	-0.009^{**}	0.041^{**}	-0.056^{***}	-0.078^{***}	-0.032	-0.070^{*}
	(0.004)	(0.017)	(0.017)	(0.023)	(0.020)	(0.037)
Number of health posts in community	0.000	-0.000	0.001^{*}	0.001	0.001	0.001
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Height, meters	1.011^{***}	0.919^{***}	0.607^{***}	0.391^{***}	0.711^{***}	0.499^{***}
	(0.026)	(0.062)	(0.032)	(0.039)	(0.045)	(0.074)
Age	0.024^{***}	0.047^{***}	0.030^{***}	0.026^{***}	0.035^{***}	-0.008
	(0.000)	(0.003)	(0.001)	(0.001)	(0.001)	(0.039)
Squared age	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	-0.000^{***}	0.003^{*}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Cooks and sleeps in the same room	-0.017^{***}	-0.001	-0.028^{***}	-0.043^{***}	-0.010	-0.033
	(0.006)	(0.011)	(0.010)	(0.012)	(0.012)	(0.022)
Kitchen is outside	0.014^{***}	0.013^{**}	0.028^{***}	0.033^{***}	0.024^{***}	0.019^{*}
	(0.003)	(0.005)	(0.005)	(0.006)	(0.006)	(0.011)
Average household education (years)	0.005^{***}	0.002^{***}	0.006^{***}	0.005^{**}	0.007^{***}	-0.000
	(0.000)	(0.001)	(0.002)	(0.002)	(0.002)	(0.006)
Distance to Kabupaten capital	0.010^{***}	0.004	0.014^{***}	0.014^{***}	0.013^{***}	0.006
	(0.003)	(0.005)	(0.003)	(0.004)	(0.004)	(0.009)
IFLS wave	0.001^{**}	-0.013^{***}				
	(0.000)	(0.002)				
Constant	2.722^{***}	29.934^{***}	4.540^{***}	4.669^{***}	4.254^{***}	4.530^{***}
	(0.609)	(4.107)	(0.062)	(0.080)	(0.078)	(0.251)
Observations	60,198	60,198	23,336	12,227	11,109	3,277
R^2	0.505	0.312	0.482	0.258	0.414	0.324
First-stage F statistic			66.688	52.170	66.727	21.364

Notes: t statistics in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. This table only includes the sample of households that cook with firewood, kerosene, gas, electricity. These remaining categories were dropped from the estimation sample: charcoal, do not cook, and other. See notes to Table 5 for other details.