Where there is smoke: Solid fuel externalities, gender, and adult respiratory health in India

Aashish Gupta[*]

June 2, 2019

Abstract

Chronic respiratory conditions are a leading cause of death in the world. Using data on lung obstruction from the WHO Survey of Global AGEing and Adult Health (WHO-SAGE 2007-08), this paper studies the determinants of respiratory health in India, home to a third of all deaths from Chronic Obstructive Pulmonary Disease. First, we find that smokers and members of households that use solid fuels (wood, biomass, coal or dung) for cooking have higher lung obstruction. Second, even if a respondent's household uses clean fuels, their lung obstruction is higher if their neighbors use solid fuels. In neighborhoods with high solid fuel use, the lungs of members of households that use clean fuels can be as obstructed as lungs of members of households that use solid fuels. These negative externalities of solid fuel use are robust to additional controls for neighborhood socioeconomic status, falsification tests, tests with placebo measures, and tests using alternative measures of respiratory health as outcomes. Third, the influence of the determinants is patterned by gender. Smoking tobacco is an important influence on lung obstruction among men. Confirming non-linear dose-response relationships, we find that women from households that use solid fuels are the only group not further harmed by neighborhood solid fuel smoke, possibly because of high exposure to pollutants while cooking. The study improves our understanding of behavioral, social, and environmental determinants of respiratory health in India. Importantly, it makes a case for greater public investments to promote the adoption and use of cleaner fuels. (248 words)

Keywords: respiratory health, lung obstruction, solid fuels, negative externalities, gender, India

Where there is smoke:
Solid fuel externalities, gender, and adult respiratory health in India

1 Introduction and background

Chronic Obstructive Pulmonary Disease (COPD), "a lung ailment that is characterized by a persistent blockage of airflow from the lungs" (World Health Organisation, 2015) is among the leading causes of death in the world (GBD Causes of Death Collaborators, 2017). Although less developed countries have a higher burden of disease attributable to chronic respiratory diseases and adult respiratory health is becoming increasingly important for overall population health, empirical research on the determinants of respiratory health at the population level in developing countries is limited (Mannino and Buist, 2007). In high-income countries, smoking tobacco is the principle risk factor for COPD (Currie, 2010; Pauwels and Rabe, 2004). Emerging evidence suggests that smoke from the use of solid fuels (wood, animal dung, coal, charcoal, and crop residues), which are used for cooking and heating, may be an equally important determinant in less developed countries (Salvi and Barnes, 2009, 2010).

Solid fuel smoke can harm the respiratory health of women (Po et al., 2011; Ezzati and Kammen, 2002), as they are the household members primarily responsible for cooking (Hirway and Jose, 2011). Other household members can be harmed because of exposure to indoor air pollution (Venkataramani and Fried, 2011; Hu et al., 2010; Mishra et al., 2005). Researchers have hypothesized that because solid fuel smoke contributes to local air pollution, neighbors can be harmed as well, even if they themselves use clean fuels (Smith, 2002; Torres-Duque et al., 2008; Balakrishnan et al., 2014; Chafe et al., 2014). This points to important negative externalities, or negative effects of solid fuel use on third parties (Pigou, 1924; Fairbrother, 2016). However, existing research has not yet tested or quantified the association between neighborhood externalities of solid fuel use and lung obstruction empirically. Although these negative externalities can extend beyond the neighborhood as well, we investigate local externalities as they are likely to be stronger. Finally, the patterning of the exposures by key social determinants such as gender has not been studied at the population level for developing countries. This paper intends to fill these research gaps, for the important case of India.

Both smoking and cooking are patterned by gender in India, a large section of the Indian population is exposed to solid fuel smoke, and respiratory health is poor. Few women in India smoke (Jha et al., 2002; Mishra et al., 2016), while men do little household work or cooking (Hirway and Jose, 2011; Sharma, 2018). Using data from the only representative survey that measured time use in India, Hirway and Jose (2011) found that while men spent less than an hour per week in India on cooking and related household activities, women spent an average of 25 hours per week. Thus, women who live in households that use solid fuels have high exposure to direct smoke while cooking (Parikh et al., 1999). More than 67.2% of Indian households used solid fuels for cooking (Government of India, 2011a; Gupta et al., 2019), and about one-third of the world's population which relies on solid fuels was estimated to live in India (Bonjour et al., 2013). Finally, in 2016, India was estimated to be home to one-third of all deaths from chronic respiratory conditions (GBD Causes of Death Collaborators, 2017), and almost half of Indian adults had COPD (World Health Organisation, 2013) by the GOLD (2016) criteria. COPD was also among the top three causes of death (Krishnan et al., 2011; Patel et al., 2011). These factors make India an important and urgent context to examine lung obstruction determinants.

The paper informs this literature by considering, for the first time, neighborhood externalities of solid fuel use in an empirical framework, by using population representative survey data on lung obstruction, by examining the gendered patterning of the exposures, and by investigating both dichotomous and continuous measures of lung obstruction as outcomes. We use population level survey data on lung obstruction collected by the World Health Organization's Survey of Global AGEing and adult health (WHO-SAGE) 2007-08 (Kowal et al., 2012).

From a social science perspective, this study contributes evidence from a developing country context to the literature on environment and poulation health, as well as the literature on how social and behavioral determinants mediate this relationship. From a health policy perspective, the study presents evidence of negative externalities, which are a classic argument for public intervention.

We discuss policy and research implications of this study in section 4. The next subsection provides background information on lung obstruction. Section 2 discusses the measures and the empirical strategy, and section 3 presents results.

1.1 Causes, prevalence, and diagnosis of lung obstruction

The lung is the body's organ of gas exchange. The trachea or the windpipe branches into the bronchi inside the lungs, which further divide into smaller branches, called bronchioles. The bronchioles end in microscopic air sacs called alveoli. Long-term exposure to noxious particles and gases produces inflammation in the lungs (Currie, 2010). This inflammation results in the narrowing of the bronchioles, and is accompanied by lung tissue destruction (emphysema), disruption of repair and defense mechanism in the lungs (bronchiolitis), and secretion of mucous and increased cough (chronic bronchitis) (Calverley and Walker, 2003). These changes increase resistance to airflow in the lungs. Persistent respiratory symptoms and lung obstruction characterize Chronic Obstructive Pulmonary Disease (GOLD, 2016). Lung obstruction can result in respiratory or heart failure (Holguin et al., 2005), increases the risk of lung infections and pneumonia (Prescott et al., 1995) while exacerbating the risk of death from co-morbidities such as heart disease and lung cancer (Sin et al., 2006).

Smoking tobacco is a well-established cause of lung obstruction. In a systematic review, Forey et al. (2011) pool odds-ratios from 133 studies and find higher odds of COPD in ever smokers, current smokers, and ex-smokers when compared to non-smokers. While smoking cessation cannot reverse damage, a systematic review found that compared to smoking, quitting was associated with slower increases in lung obstruction (Godtfredsen et al., 2008). Inhaling smoke from solid fuels is the other big risk factor for lung obstruction. Torres-Duque et al. (2008) and Kurmi et al. (2010) review evidence on the impact of solid fuel use on respiratory diseases, including COPD. Kurmi et al. (2010) find higher odds of having COPD among households that use solid fuels. In India, McKay et al. (2012); Moschovis et al. (2015); Dave et al. (2017) have found sex, smoking status, and indoor air pollution to be significant predictors of lung obstruction. However, some studies (Amaral et al., 2017; Sana et al., 2018) have failed to find significantly higher odds of COPD among those who used solid fuels. These studies have not considered solid fuel externalities. The comparison groups in these studies, those not using solid fuels, may also have high lung obstruction because of exposure to neighborhood air pollution.

COPD is diagnosed using a non-invasive physiological test, spirometry, which measures lung obstruction. A spirometer measures the proportion of the volume of air (liters) one can exhale in

1 second to the total volume of air one exhales, after inhaling as deeply as possible (Cotes et al., 2009). This ratio is called Forced Expiratory Volume in 1 second percent (FEV1%). Higher values of FEV1% signify better respiratory health, and imply lesser obstruction of airflow. Existing research on solid fuel use and respiratory health relies on dichotomized COPD as the outcome, based on the Global initiative for Obstructive Lung Disease (GOLD) criteria of diagnosing COPD if FEV1% is less than 70% (GOLD, 2016). The GOLD cut-off over-diagnoses older patients, and under-diagnose younger patients (Pellegrino et al., 2005). While categorization has clinical relevance, such dichotomization does not capture the severity of the disease and reduces statistical power and precision (Weinberg, 1995; Royston et al., 2006). Consequently, we use continuous FEV1% as the outcome, but confirm that the results are robust to using dichotomized COPD.

2 Materials and methods

2.1 Data and measures

The WHO-SAGE (2007-08) is representative of the adult population aged above 18 years in six states of India: Rajasthan, Uttar Pradesh, Assam, Karnataka, Maharashtra and West Bengal. These states had a combined population of 570 million in 2011, representing 47% of India's total population (Government of India, 2011b). The WHO-SAGE is unique among surveys carried out in India in that it conducted spirometry to measure lung obstruction for its 9,551 respondents (World Health Organisation, 2013). The WHO-SAGE also collected information on other individual and household characteristics. Table 1 provides summary statistics for the main measures used in this analysis of WHO-SAGE data.

2.1.1 Outcome

The primary outcome is Forced Expiratory Volume in 1 second percent (FEV1%), the proportion of total air inhaled that a person can exhale in one second. FEV1% is a direct measurement of lung obstruction. The outcome is also called FEV1/FVC (ratio of Forced Expiratory Volume in 1 second to Forced Vital Capacity). We use the continuous measure to preserve the full range of variation in lung function and improve the precision of our estimates. Our results were qualitatively un-

changed when using a dichotomous outcome. Average FEV1% was 70.8% in the WHO-SAGE sample. Mean levels of FEV1% as low as these, very close to the GOLD criteria of FEV1% of 70% to clinically diagnose someone as having COPD, reinforce the impaired state of respiratory health in India. Men and women had similar mean values of lung obstruction (Table 1).

2.1.2 Exposures

The main independent variables that measure exposure to risk factors for lung obstruction are smoking; household solid fuel use, and neighborhood solid fuel use.

Smoking: Smoking status was self-reported. Smoking is coded as a dummy variable (a very small fraction of the sample consisted of former smokers, and second-hand smoke was not observed). About 17% of the SAGE respondents smoked (Table 1). While 31% of the men smoked, only 2% of the women smoked.

Household use of solid fuels: Respondents were asked the fuel their household uses for cooking. Coal/charcoal, wood, agricultural residue/shrubs and animal dung were classified as solid fuels. Kerosene, gas, and electricity were classified as clean fuels. In robustness tests, we examine if the result change by considering only gas and electricity as clean fuels (Table A3). Household use of solid fuels is coded as a dummy variable. About 76% of all households used solid fuels for cooking.

Neighborhood solid fuel use: Neighborhood solid fuel use is measured as the proportion of households, excluding the respondent's household, in a Primary Sampling Unit (PSU) that use solid fuels. A PSU is an urban block or a rural village. Neighborhood solid fuel use is coded as a continuous variable between 0 and 1. The SAGE survey interviewed respondents in 375 different clusters or PSUs, selecting on average 26 households per cluster. Nine out of ten clusters had more than 19 respondents (a histogram of the cluster size is shown in Figure A1). Solid fuel use in the neighborhood is hypothesized to capture the deterioration in air quality in the neighborhood due to solid fuel use. On average, 76% of a respondent's neighbors used solid fuels. Accordingly, the respondent would be assigned a value of 0.76 as a measure of neighborhood solid fuel use.

Although a larger sample from each PSU would be desirable, such specifications have been used productively in research on neighborhood disease exposure, such as sanitation (Geruso and

Spears, 2018) and vaccination (McGovern and Canning, 2015). The neighborhood solid fuel use measure may introduce measurement error due to extreme values (everyone using solid fuels or nobody using solid fuels). This measurement error biases our estimates on neighourhood solid fuel use towards zero, thus making these estimates conservative. An additional concern is the correlation between neighborhood and household solid fuel use. This estimated correlation between these variables was .6969, and the Variation Inflation Factor based on the r-squared in a regression was 1.955. We find similar results when we use district level solid fuel use estimated from the 2011 Indian census instead of neighourhood solid fuel use estimated from the WHO-SAGE 2007-08 (Table A3).

2.1.3   Other explanatory variables

In all regressions, we adjust for age dummies (in years), sex of the respondent, state fixed effects, and a dummy for living in a rural area. This specification of age has the advantage of accounting for non-linear relationships between age and lung obstruction. In the SAGE sample, average age was 40.4 years. Men were slightly older than women, and 47% of the sample was female. Almost 70% of the respondents lived in rural areas (Table 1). Apart from demographic and place controls, we also control for two measures of socioeconomic status, namely, years of education and number of household assets owned. We use dummies for years of education and count of assets, a stronger control than linear controls for wealth and education. Years of education ranged from 0 to 18. Mean years of education were 5.9 years, 7.4 years for men and 4.2 years for women, reflecting an aspect of gender inequality well-known for India (Drèze and Sen, 2013). The WHO SAGE asked respondents the number of chairs, tables, and cars their household owned, whether the household had a cycle, clock, bucket, cot, fridge, fixed line phone, mobile phone, television, computer, radio, livestock, sewing machine, a motorized two-wheeler, bullock cart, and non-homestead agricultural land. Based on these indicators, we created a variable for the count of assets a respondent's household owned. The mean number of assets was 13.2. Using district identifiers in the WHO-SAGE, we matched the respondents with the proportion of households who were using solid fuels in their district in the 2011 census (Government of India, 2011a). Districts in India are a third level of administrative unit, below the federal and the state level, and above the level of rural villages or

urban municipalities. India has 727 districts, and the WHO SAGE was conducted in 139 of them. Mean district level solid fuel use was 74%. We also used estimates of particulate matter pollution (PM 2.5) derived by Van Donkelaar et al. (2016). Since geographic coordinates of SAGE respondents or their clusters were not available, we estimated mean district level exposure in ArcGIS. Mean PM 2.5 exposure was 52.0 μg / m^3, which is relatively very high.

2.2 Empirical strategy

The empirical strategy consists of several steps, including descriptive graphs of cumulative density function of lung obstruction by exposure type, Ordinary Least Squares (OLS) regressions, and multiple robustness checks. The OLS linear regression equation is specified as:

$$respiratory\,health_{ij} = \beta_1 smoke\,tobacco_{ij} + \beta_2 household\,solid\,fuel\,use_{ij} + \beta_3 neighborhood\,solid\,fuel\,use_{ij}$$
$$+ \alpha place_{ij} + \theta demographic_{ij} + \lambda ses_{ij} + \varepsilon_{ij},$$

$$(1)$$

where respiratory health is a measure of obstruction of airflow of respondent i living in primary sampling unit j, measured as FEV1%. We cluster standard errors at the level of the primary sampling unit, to account for the WHO SAGE's multi-stage sample design, and use survey weights throughout. The coefficients of interest are β_1, of whether the respondent smokes, β_2, of whether the respondent's household uses solid fuels, and β_3, of the proportion of other households in a respondent's primary sampling unit who use solid fuels. β_3 is a continuous measure between 0 and 1. Since regression coefficients show the effect of a one-unit change, β_3 indicates the difference between respondents who live in neighborhoods where no other households use solid fuels and respondents who live in neighborhoods where all other households use solid fuels.

$place_{ij}$ controls for state fixed effects and a dummy for rural residence; $demographic_{ij}$ adjusts for sex of the respondent and age dummies, in years; and ses_{ij} controls include dummies for both years of education and count of household assets.

In addition to these controls for SES variables, we conduct further robustness checks, adjusting for additional household and neighborhood SES variables, such as log household expenditure

7

and average neighborhood assets. In placebo tests, we test whether measures similar to the exposures, such as chewing tobacco, drinking alcohol, household ownership of toilets, household electrification, neighborhood open defecation or neighborhood electrification are associated with the outcome. In falsification tests, we test whether other health outcomes, such as Forced Vital Capacity of lungs or adult heights are associated with the exposures. Finally, we test if the exposures are associated with alternative measures of respiratory health, such as self-reported lung disease and Forced Expiratory Volume in 1 second. In placebo tests, it is expected that there will be no association between the placebo variables and the outcome. In falsification tests, it is expected that there will be no association between the falsification outcomes and the exposures. We expect to find an association between the alternative measures of respiratory health and the exposures. These results are available in the supplementary appendix.

3   Results

The results section is organized as follows. We begin by descriptive graphs in section 3.1, first examining the cumulative density function of FEV1% by the three exposures. We then see the extent to which the differences in the exposures can be explained by socioeconomic status. Thereafter, we examine whether household and neighborhood solid fuel use jointly influence lung obstruction. Section 3.2 presents regression results, documenting the association between the outcome and the exposures while adjusting for other explanatory variables, and examining the implications of the gendered patterning of the exposures. Finally, section 3.4 presents the results from robustness checks, including the placebo and falsification tests, and tests with alternative measures of respiratory health as outcomes.

3.1   Descriptive graphs

Figure 1 presents cumulative density function of FEV1% for those exposed to the three risk-factors of lung obstruction (smoking, household solid fuel use, neighborhood solid fuel use) and those not exposed. The cumulative distributions show differences in lung obstruction by exposure at all levels of the distribution of FEV1%. p-values and D-values (absolute maximum distance between the two distributions) from Kolmogorov-Smirnov tests, and p-values and z-statistics from

8

Wilcoxon rank-sum test are also shown to compare the distributions in each of the sub-figures. The p values, which are all less than .001, indicate that the cumulative density functions of those exposed to the risk factors and those not exposed are significantly different. Smokers, members of households that use solid fuels, and residents of neighborhoods in which other members use solid fuels have higher lung obstruction across the distribution, than non-smokers, members of households that use clean fuels and residents of neighborhoods that have lower levels of solid fuel use, respectively. The figure reveals that differences by household and neighborhood solid fuel use exposure are larger than or similar to the differences by smoking.

Figure 2 provides graphical evidence that these differences cannot be explained by differences in socioeconomic status. In fact, measures of socioeconomic status, such as years of education and household assets themselves have a weak relationship with lung obstruction. In Figure 2 the left panel shows differences in years of education, and the right panel shows differences by count of household assets using non-parametric local polynomial regressions with 95% confidence intervals shaded blue. At all levels of education and household assets, respondents whose households or neighbors use solid fuels have higher lung obstruction and lower values of FEV1%.

Figure 3 shows differences in lung obstruction between members of households that use solid fuels and members of households that use clean fuels by neighborhood solid fuel use. The x-axis shows the proportion of households in a primary sampling unit using solid fuels, and the y-axis shows FEV1%. Among those respondents whose households use clean fuels, there is a much steeper decline in FEV1% as neighborhood solid fuel use rises. Among those respondents whose households use solid fuels, as neighborhood solid fuel increases, a modest decline in FEV1% can be seen. That respondents who use cleaner fuels show a steeper decline in FEV1% as neighborhood solid fuel use increases reinforces existing evidence of a downward sloping concave concentration response function. This implies diminishing marginal impacts of increases in pollution (Burnett et al., 2014; Pope III et al., 2015; Gupta and Spears, 2017). The pattern of non-linear decline among respondents whose households use clean fuels also supports these results. Respondents already exposed to direct or indoor smoke from solid fuel use are not much harmed by neighborhood solid fuels.

3.2 Regression results

Table 2 provides results from OLS regressions. The results indicate that average differences by the exposures are statistically significant, in magnitude similar to the ones in the descriptive graphs, and robust to demographic, place, and SES controls. The first 3 models present results for the overall sample, models 4 and 5 restrict the sample to women, and models 6 and 7 consider only men. All regressions adjust for age dummies, sex, state fixed effects and rural residence. All models except model 1 also adjust for socioeconomic status, by using dummies for years of education and number of household assets owned. Models with linear controls for age and measures of socioeconomic status are available in appendix table A1. Appendix table A1 shows that assets and education are not associated with lung obstruction, and that dummies for age and socioeconomic status improve model fit.

Model 1 of table 2 shows that those who smoke, whose household uses solid fuels, and whose neighbors use solid fuels have lower FEV1% and higher lung obstruction. Model 2 adjusts for socioeconomic status but the point estimates or the significance of the coefficients do not change substantially. Compared to a respondent who does not smoke, whose household does not use solid fuels, and whose neighbors do not use solid fuels, a respondent who smokes, or a respondent whose household uses solid fuels, is, on average, able to breathe out 3.7 percentage points less oxygen in the first second. Holding smoking and household solid fuel use constant, a respondent living in a neighborhood in which all her/his neighbors use solid fuels is able to breathe out 6 percentage points less FEV1%. Models 1 and 2 reinforce the results from the descriptive graphs, that household solid fuel use and smoking tobacco are equivalent risks, and neighborhood solid fuel use is an even greater risk for lung obstruction. In other words, a person who is a member of a household that does not use solid fuels, does not smoke, but lives in a neighborhood in which half the households use solid fuels is likely to have the same level of lung obstruction as a person who is a tobacco smoker.

Model 3 shows the interaction between household and neighborhood solid fuel use. Predicted FEV1% for household and neighborhood solid fuel use are available in figure 4. The interaction results and predicted FEV1% closely replicate the results from figure 3: as neighborhood solid fuel rises, FEV1% declines more steeply among members of households that use clean fuels.

3.3 Gender, exposures, and lung obstruction

Women and men differ both in the extent of their exposure to solid fuel smoke and smoking tobacco (Jha et al., 2002; Hirway and Jose, 2011), as well as their physiological responses to the exposures (Aryal et al., 2013), making a case for within-sex comparisons. The next set of models in table 2 present stratified regressions for women (models 4 and 5) and men (models 6 and 7).

Model 4 of table 2 shows that while smoking has a negative association with FEV1% among women, the effect size is much smaller than that of men (model 7), and not significant. This is likely due to selection, lesser intensity of smoking among women, a very small number of women for whom smoking is observed, and possible reporting bias because women may be less willing to say that they smoke. Household solid fuel smoke is significantly and negatively associated with FEV1% among women. While neighborhood solid fuel use has a negative association with FEV1%, this effect isn't statistically significant (*p*-value 0.128). However, model 5, which interacts household and neighborhood solid fuel use for women, shows that neighborhood solid fuel use has a significantly negative association for women whose households don't use solid fuels. On the other hand, women who are members of households that use solid fuels are not further harmed by neighborhood solid fuel use. This result can be more clearly seen through predicted margins, reported in figure 4.

Among men, smoking is associated with higher lung obstruction and lower FEV1%. The association of household use of solid fuels has a similar magnitude as the association of smoking, while the neighborhood solid fuel use coefficient is even higher than that of smoking (model 6). Model 7 reports coefficients from an interaction of household solid fuel use and neighborhood solid fuel use. For men, the interaction is not significant. These patterns, consistent with Burnett et al. (2014) and Pope III et al. (2015), underscore the high levels of exposure to smoke among women exposed to solid fuel smoke from direct cooking.

3.4 Robustness tests

As described previously, we next examine the robustness of the results. The regression tables are shown in the supplementary appendix.

Table A2 adds controls for log monthly expenditure, log monthly expenditure squared, aver-

age neighborhood assets, and average neighborhood assets sequentially to the main specification from Table 2 (model 2), repeated as Model 1 in appendix table A2 for comparison. The log of monthly household expenditure and the square of the log expenditure are not significantly associated with lung obstruction (model 2). Nor does the inclusion of these controls change the magnitude of the associations for the three exposures. Models 3 and 4 add controls for average neighborhood assets (model 3) and average neighborhood education (model 4). As with neighborhood solid fuel use, these variables are constructed as average of other households' average assets, or other respondents' average education. They are also not associated with FEV1%, and do not substantively change the effect sizes of the exposures.

Table A3 tests for robustness in the exposures of household and neighourhood solid fuel use. In Model 1, we test if the main result is valid in the sub-sample of the SAGE in primary sampling units with more than 19 households. These results are similar to the main result (Table 2, Model 1), confirming that the associations are not driven by the small proportion of primary sampling units with fewer respondents. Model 2 uses an instrumental variable approach. We randomly split the respondents in each primary sampling unit into two units. We then instrument the mean neighourhood solid fuel use in the first sub-sample of the primary sampling unit with the mean solid fuel use in the second sub-sample of the primary sampling unit. In this instrumental variable approach, the coefficients for both household and neighourhood solid fuel use are higher than those in the main specification, suggesting that the estimates in the main model are conservative. Models 3 and 4 test for the association between district level solid fuel use and lung obstruction. We estimate district level solid fuel use from the 2011 census, which asked all households the fuel they used for cooking. Model 3 shows that the association of neighourhood and district level solid fuel use with lung obstruction is of similar magnitude. Model 4, which has all the exposures, at the household, neighourhood, and district level, shows that it is the household and neighourhood level exposures that matter. Once household and neighourhood level exposures are accounted for, district level exposures are not significant. This makes sense, because districts are much larger units. Finally, Model 5 separates clean fuels into kerosene and other clean fuels (gas and electricity). The model shows that Kerosene does not appear to be significantly worse than other cleaner fuels.

Table A4 reports placebo tests for smoking tobacco, household solid fuel use, and neighborhood solid fuel use. The associations of chewing tobacco and drinking alcohol with FEV1% are compared to that of smoking tobacco. Owning a toilet and household electrification, measures of household SES, are used as placebos for solid fuel use. Measures for neighborhood electrification and neighborhood open defecation are used as placebos for neighborhood solid fuel use. As before, all regressions control for demographic, place, and other socioeconomic status variables. Panel A shows that there is no association between FEV1% and chewing tobacco, or between FEV1% and drinking alcohol. Chewing tobacco and drinking alcohol cause other health problems, but are not linked to respiratory health, either theoretically or empirically in the WHO-SAGE. Regressions in Panel B find no association of FEV1% with household ownership of toilets and household electrification, and regressions in Panel C find no association of lung obstruction with neighborhood electrification or neighborhood open defecation. The effect sizes for the placebo variables are also close to zero.

In Table A5, we conduct falsification tests with health measures we do not expect to be correlated with solid fuel use. It is well known that exposures such as smoking or solid fuel use harm airflow obstruction but not total lung capacity, which we confirm in model 1. The outcome in model 2 is height of the respondents. While height is influenced by early childhood disease environment, adult exposures to solid fuel use should not influence height, which is confirmed in model 2. We show that word recall, a measure of memory, and walk time, a measure of physical strength, are also not associated with the exposures.

In Table A6, we confirm that the mechanism through which lung obstruction is associated with neighourhood solid fuel use is through ambient air pollution (models 1-3). We also confirm that alternative measures of respiratory health, such as dichotomized COPD and self-reported lung disease are predicted by the exposures. Model 1 examines the relationship between the outcome and the exposures, without controls for state of the respondent and rural residence, along with satellite PM 2.5 exposures. It finds that both neighourhood solid fuel use and satellite PM 2.5 are associated with lung obstruction. Model 2 introduces controls for state dummies and rural residence, as in our main specification. In this model, while neighourhood solid fuel use is significant, the association of satellite PM 2.5 is no longer significant. This could partly because of

the spatial correlation of satellite measures of air pollution, which vary only over large geographic areas, such as states (Van Donkelaar et al., 2016). Still, Model 3, which has PM 2.5 as an outcome, confirms that local solid fuel use is predictive of satellite air pollution, establishing a link between neighourhood solid fuel use, pollution, and health.

Model 4 in Table A6 shows that smoking tobacco, household use of solid fuels, and neighborhood solid fuel use are negatively associated with Forced Expiratory Volume in 1 second, measured in liters. Because the amount of air blown out is also associated with stature, model 4 adjusts for height in meters. Model 5 and 6 use the logistic model with binary outcomes of having COPD (following GOLD 2016 guidelines) and self-reported symptoms of lung disease, following the WHO SAGE guidelines (Arokiasamy et al., 2015). Both models show higher odds of having COPD and self-reported lung diseases among smokers, members of households who use solid fuels, and residents of neighborhoods with high solid fuels use, reinforcing the main finding.

4   Discussion

This paper expands evidence on the determinants of lung obstruction for the important case of India, where one-third of all deaths due to chronic respiratory conditions take place, in three key ways. First, to our knowledge, this is the first study that documents an empirical association between local solid fuel use by other households in one's neighborhood and obstruction in one's own lungs. Multiple robustness checks suggest a causal interpretation of this association. Second, this study improves our understanding of gender as a determinant of respiratory health in India. We find that women who live in households that use solid fuels are not further harmed by neighborhood solid fuel use. They are the only group not further harmed by neighborhood solid fuel use, which points to the already high exposure to indoor smoke they must have while cooking. Gender also influences smoking behavior. Finally, the study establishes that large sections of the Indian population - those who smoke, those who use solid fuels, and those who live near others who use solid fuels - have higher lung obstruction than those who aren't exposed to these determinants of lung obstruction. Higher socioeconomic status does not mitigate the influence of these exposures.

From the point of view of estimating differences in lung obstruction due to the exposures over

a person's life course, the study offers conservative estimates. Some of the respondents who were not exposed at the time of the survey would have been exposed to the exposures earlier in their lives. However, it has some limitations. The lungs of women who cook with solid fuels are more likely to be obstructed (Johnson et al., 2011), but cooking behavior is not observed in the WHO-SAGE. Future health surveys should identify household members responsible for cooking, and if possible, ask about duration of cooking, so that dose-response relationships can be understood. We are able to include only district-level satellite measures of air pollution in our analysis, as ground-level pollution monitoring in India is in a nascent stage (Guttikunda, 2017).

The study is also limited because it uses cross-sectional data. However, from a policy-makers' perspective, this study elevates the importance of promoting clean fuels to improve adult health in India. Even though currently exposed populations may not see immediate or significant improvements in FEV1% (Chapman et al., 2005; Smith-Sivertsen et al., 2009; Kurmi et al., 2012), further obstruction can be prevented. Encouraging the adoption of cleaner fuels remains a substantial challenge (Mobarak et al., 2012; Peel et al., 2015; Jeuland et al., 2015). But the only viable public health strategy for reducing lung obstruction is prevention through the adoption of cleaner fuels. Existing government programs in India to promote cleaner fuels are means-tested, provide modest capital subsidies, involve significant private investments, do not have components for encouraging behavior change, and focus on private harms to women from cooking with solid fuels (Sethi and Deep, 2018; Gupta et al., 2019). There is a case for greater investments in efforts to communicate the private and neighborhood harms of solid fuel use and encouraging households to switch to cleaner fuels.

From an environmental health perspective, there is an urgent need for research that integrates social determinants, environmental exposures, measures of pollution, and respiratory health outcomes. Existing evidence points that women's lower status impedes the adoption of cleaner fuels (Kishore and Spears, 2014; Austin and Mejia, 2017; Gupta et al., 2019). Researching the role of gender in the persistence of solid fuel use in India, and investigating approaches that address this social determinant of adoption of cleaner fuels are an important research concern as well.

References

Amaral, André FS et al.. 2017. "Airflow obstruction and use of solid fuels for cooking or heating: BOLD results." *American journal of respiratory and critical care medicine*(ja).

Arokiasamy, Perianayagam et al.. 2015. "The impact of multimorbidity on adult physical and mental health in low-and middle-income countries: what does the study on global ageing and adult health (SAGE) reveal?" *BMC medicine*, 13(1): 178.

Aryal, Shambhu et al.. 2013. "COPD and gender differences: an update." *Translational Research*, 162(4): 208–218.

Austin, Kelly F and Maria Theresa Mejia. 2017. "Household air pollution as a silent killer: womens status and solid fuel use in developing nations." *Population and Environment*, 39(1): 1–25.

Balakrishnan, Kalpana et al.. 2014. "Addressing the burden of disease attributable to air pollution in India: the need to integrate across household and ambient air pollution exposures." *Environmental health perspectives*, 122(1): A6.

Bonjour, Sophie et al.. 2013. "Solid fuel use for household cooking: country and regional estimates for 1980-2010." *Environmental Health Perspectives (Online)*, 121(7): 784.

Burnett, Richard T et al.. 2014. "An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure." *Environmental health perspectives*, 122(4): 397.

Calverley, PMA and Paul Walker. 2003. "Chronic obstructive pulmonary disease." *The Lancet*, 362(9389): 1053 – 1061.

Chafe, Zoë A et al.. 2014. "Household cooking with solid fuels contributes to ambient PM2. 5 air pollution and the burden of disease." *Environmental health perspectives*, 122(12): 1314.

Chapman, Robert S et al.. 2005. "Improvement in household stoves and risk of chronic obstructive pulmonary disease in Xuanwei, China: retrospective cohort study." *Bmj*, 331(7524): 1050.

Cotes, John E et al.. 2009. *Lung function: physiology, measurement and application in medicine*: John Wiley and Sons.

Currie, Graeme P. 2010. *ABC of COPD*: John Wiley & Sons.

Dave, Mona et al.. 2017. "Household air pollution and lung function in Indian adults: a cross-sectional study." *The International Journal of Tuberculosis and Lung Disease*, 21(6): 702–704.

Drèze, Jean and Amartya Sen. 2013. *An uncertain glory: India and its contradictions*: Princeton University Press.

Ezzati, Majid and Daniel M Kammen. 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): 1057.

Fairbrother, Malcolm. 2016. "Externalities: why environmental sociology should bring them in." *Environmental Sociology*, 2(4): 375–384.

Forey, Barbara A et al.. 2011. "Systematic review with meta-analysis of the epidemiological evidence relating smoking to COPD, chronic bronchitis and emphysema." *BMC pulmonary medicine*, 11(1): 1.

GBD Causes of Death Collaborators. 2017. "Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016." *Lancet*, 390(10100): 1151–1210.

Geruso, Michael and Dean Spears. 2018. "Neighborhood Sanitation and Infant Mortality." *American Economic Journal: Applied Economics*.

Godtfredsen, NS et al.. 2008. "COPD-related morbidity and mortality after smoking cessation: status of the evidence." *European Respiratory Journal*, 32(4): 844–853.

GOLD. 2016. "Global Strategy for the Diagnosis, Management and Prevention of COPD."Technical report, Global Initiative for Chronic Obstructive Lung Disease - GOLD.

Government of India. 2011a. "Census of India 2011: Household Listing Operations." *Office of the Registrar General Census Commissioner, India*.

⸻ 2011b. "Census of India 2011: Houselisting Operations." *Housing and Houselisting Operations*.

Gupta, Aashish and Dean Spears. 2017. "Health externalities of India's expansion of coal plants: Evidence from a national panel of 40,000 households." *Journal of environmental economics and management*, 86: 262–276.

Gupta, Aashish, Sangita Vyas, Payal Hathi, Nazar Khalid, Nikhil Srivastav, Dean Spears, and Diane Coffey. 2019. "Persistence of solid fuel use despite increases in LPG ownership: New survey evidence from rural north India."

Guttikunda, Sarath. 2017. "Its About Time We Got Smarter About Monitoring Our Air Pollution." *The Wire*.

Hirway, Indira and Sunny Jose. 2011. "Understanding women's work using time-use statistics: The case of India." *Feminist Economics*, 17(4): 67–92.

Holguin, Fernando et al.. 2005. "Comorbidity and mortality in COPD-related hospitalizations in the United States, 1979 to 2001." *Chest Journal*, 128(4).

Hu, Guoping et al.. 2010. "Risk of COPD from exposure to biomass smoke: a metaanalysis." *Chest Journal*, 138(1): 20–31.

Jeuland, Marc et al.. 2015. "The economics of household air pollution." *Annu. Rev. Resour. Econ.*, 7(1): 81–108.

Jha, Prabhat, M Kent Ranson, Son N Nguyen, and Derek Yach. 2002. "Estimates of global and regional smoking prevalence in 1995, by age and sex." *American journal of public health*, 92(6): 1002–1006.

Johnson, Priscilla et al.. 2011. "Prevalence of chronic obstructive pulmonary disease in rural women of Tamilnadu: implications for refining disease burden assessments attributable to household biomass combustion." *Global health action*, 4(1): 7226.

Kishore, Avinash and Dean Spears. 2014. "Having a son promotes clean cooking fuel use in urban India: Womens status and son preference." *Economic Development and Cultural Change*, 62(4): 673–699.

Kowal, Paul et al.. 2012. "Data resource profile: the World Health Organization Study on global AGEing and adult health (SAGE)." *International journal of epidemiology*, 41(6): 1639–1649.

Krishnan, Anand et al.. 2011. "Adult mortality surveillance by routine health workers using a short verbal autopsy tool in rural north India." *Journal of Epidemiology & Community Health*.

Kurmi, Om P et al.. 2010. "COPD and chronic bronchitis risk of indoor air pollution from solid fuel: a systematic review and meta-analysis." *Thorax*, 65(3): 221–228.

⸻ 2012. "Indoor air pollution and the lung in low-and medium-income countries." *European Respiratory Journal*, 40(1): 239–254.

Landrigan, Philip J et al.. 2017. "The Lancet Commission on pollution and health." *The Lancet*.

Mannino, David M and A Sonia Buist. 2007. "Global burden of COPD: risk factors, prevalence, and future trends." *The Lancet*, 370(9589): 765–773.

McGovern, Mark E and David Canning. 2015. "Vaccination and all-cause child mortality from 1985 to 2011: global evidence from the Demographic and Health Surveys." *American journal of epidemiology*, 182(9): 791–798.

McKay, Ailsa J et al.. 2012. "Prevalence of COPD in India: a systematic review." *Primary Care Respiratory Journal*, 21: 313–321.

Mishra, Sujata et al.. 2016. "Trends in bidi and cigarette smoking in India from 1998 to 2015, by age, gender and education." *BMJ global health*, 1(1).

Mishra, Vinod, Kirk R Smith, and Robert D Retherford. 2005. "Effects of cooking smoke and environmental tobacco smoke on acute respiratory infections in young Indian children." *Population and Environment*, 26(5): 375–396.

Mobarak, Mushfiq et al.. 2012. "Low demand for nontraditional cookstove technologies." *Proceedings of the National Academy of Sciences*, 109(27): 10815–10820.

Moschovis, Peter P et al.. 2015. "The Association Of Household Air Pollution And Smoking With Risk Of Obstructive Lung Disease Among Adults In India." in *A105. COPD EPIDEMIOLOGY: GLOBAL PERSPECTIVE*: Am Thoracic Soc: A6353–A6353.

Parikh, Jyoti et al.. 1999. "Indoor air pollution: a reflection on gender bias." *Economic and Political Weekly*: 539–544.

Patel, Vikram et al.. 2011. "Chronic diseases and injuries in India." *The Lancet*, 377(9763): 413–428.

Pauwels, Romain A and Klaus F Rabe. 2004. "Burden and clinical features of chronic obstructive pulmonary disease (COPD)." *The Lancet*, 364(9434): 613–620.

Peel, Jennifer L et al.. 2015. "Are Randomized Trials Necessary to Advance Epidemiologic Research on Household Air Pollution?" *Current Epidemiology Reports*, 2(4): 263–270.

Pellegrino, Riccardo et al.. 2005. "Interpretative strategies for lung function tests." *European Respiratory Journal*, 26(5): 948–968.

Pigou, Arthur. 1924. *The economics of welfare*: Routledge.

Po, June YT et al.. 2011. "Respiratory disease associated with solid biomass fuel exposure in rural women and children: systematic review and meta-analysis." *Thorax*, 66(3): 232–239.

Pope III, C Arden et al.. 2015. "Health benefits of air pollution abatement policy: role of the shape of the concentration–response function." *Journal of the Air & Waste Management Association*, 65(5): 516–522.

Prescott, E et al.. 1995. "Chronic mucus hypersecretion in COPD and death from pulmonary infection." *European Respiratory Journal*, 8(8): 1333–1338.

Royston, Patrick et al.. 2006. "Dichotomizing continuous predictors in multiple regression: a bad idea." *Statistics in medicine*, 25(1): 127–141.

Salvi, Sundeep S and Peter J Barnes. 2009. "Chronic obstructive pulmonary disease in nonsmokers." *The lancet*, 374(9691): 733–743.

———— 2010. "Is exposure to biomass smoke the biggest risk factor for COPD globally?" *CHEST Journal*, 138(1): 3–6.

Sana, Adama et al.. 2018. "Chronic obstructive pulmonary disease associated with biomass fuel use in women: a systematic review and meta-analysis." *BMJ Open Respiratory Research*, 5(1): e000246.

Sethi, Nitin and Aroon Deep. 2018. "PMs plan for free gas connections is failing its objective as government had been warned it would." *Scroll.in*.

Sharma, Kanika. 2018. "Living with Pain: Womens Everyday Lives and Health in Rural Bihar." *The Hindu Centre for Public Policy*(Policy Report No. 23).

Sin, Don D et al.. 2006. "Mortality in COPD: role of comorbidities." *European Respiratory Journal*, 28(6): 1245–1257.

Smith, Kirk R. 2002. "Indoor air pollution in developing countries: recommendations for research." *Indoor air*, 12(3): 198–207.

Smith-Sivertsen, Tone et al.. 2009. "Effect of reducing indoor air pollution on women's respiratory symptoms and lung function: the RESPIRE Randomized Trial, Guatemala." *American journal of epidemiology*, 170(2): 211–220.

Torres-Duque, Carlos et al.. 2008. "Biomass fuels and respiratory diseases: a review of the evidence." *Proceedings of the American Thoracic Society*, 5(5): 577–590.

Van Donkelaar, Aaron, Randall V Martin, Michael Brauer, N Christina Hsu, Ralph A Kahn, Robert C Levy, Alexei Lyapustin, Andrew M Sayer, and David M Winker. 2016. "Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites, models, and monitors." *Environmental science & technology*, 50(7): 3762–3772.

Venkataramani, Atheendar S and Brian J Fried. 2011. "Effect of worldwide oil price fluctuations on biomass fuel use and child respiratory health: evidence from Guatemala." *American journal of public health*, 101(9): 1668–1674.

Weinberg, Clarice R. 1995. "How bad is categorization?" *Epidemiology*: 345–347.

World Health Organisation. 2013. *Study on Global AGEing and Adult Health: India*: WHO & IIPS.

———— 2015. "World Health Organisation Fact-Sheet on Chronic Obstructive Pulmonary Disease." *Geneva: World Health Organization*.

Table 1: Descriptive statistics for the sample

	Total		Women		Men	
	mean	s.e.	mean	s.e.	mean	s.e.
Outcome						
Forced Expiratory Volume 1s %	70.8	0.5	71.2	0.7	70.5	0.6
Exposures						
Respondent smokes	0.17	0.01	0.02	0.00	0.31	0.02
Household uses solid fuels	0.76	0.02	0.77	0.02	0.76	0.02
Neighbourhood solid fuel use	0.76	0.02	0.77	0.02	0.76	0.02
Explanatory variables						
Demographic						
Age	40.4	0.3	39.4	0.3	41.3	0.4
Female	0.47					
Socio-economic Status						
Years of education	5.9	0.2	4.2	0.2	7.4	0.2
# of assets owned	13.2	0.2	13.2	0.2	13.3	0.2
Contextual						
Rural	0.70	0.02	0.70	0.03	0.70	0.02
District level solid fuel use	0.74	0.00	0.74	0.00	0.73	0.00
District level PM 2.5 (ug /m3)	52.0	0.3	51.6	0.3	52.5	0.4
N	9551		5777		3774	

Note: All means are weighted using national individual weights, and standard errors are clustered at the level of the primary sampling unit. FEV1% = (FEV1s * 100) / FVC.

Table 2: OLS regression coefficients of the association of smoking, household solid fuel use, and neighborhood solid fuel use with lung obstruction

	overall sample			women		men	
model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
smoke tobacco	-3.403**	-3.764***	-3.812***	-0.307	-0.506	-3.510**	-3.529**
	(1.04)	(1.04)	(1.04)	(3.183)	(3.185)	(1.130)	(1.131)
household uses solid fuels	-3.226*	-3.762**	-7.407*	-3.484*	-10.510*	-4.166*	-5.460
	(1.25)	(1.17)	(3.01)	(1.460)	(4.171)	(1.725)	(3.999)
neighbourhood solid fuel use	-6.082*	-6.088*	-7.299***	-3.399	-8.831***	-5.259*	-6.153+
	(2.50)	(2.45)	(2.06)	(2.231)	(2.480)	(2.457)	(3.453)
hh X neighbourhood solid fuel use			5.79		11.091*		2.011
			(3.71)		(5.149)		(5.191)
female	-0.396	-0.859	-0.849				
	(0.84)	(0.91)	(0.91)				
demographic controls	X	X	X	X	X	X	X
place controls	X	X	X	X	X	X	X
ses controls		X	X	X	X	X	X
constant	72.829***	74.768***	75.497***	71.536***	73.507***	76.119***	76.403***
	(1.96)	(2.62)	(2.71)	(3.284)	(3.490)	(3.151)	(3.355)
observations	9551	9551	9551	5777	5777	3774	3774

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001.
Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (litres) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible). Demographic controls: age dummies and sex (coefficients for female shown). Place controls: rural residence and state fixed effects. SES controls: dummies for years of education and count of assets. Appendix Table 1 compares models with linear controls for age and ses measures with fixed effects for age and ses measures. Appendix Figure 1 shows the predicted margins of the interaction of household and neighbourhood solid fuel use, from models 3, 5, and 7. hh: household.

Figure 1: Cumulative Density Function of lung obstruction by exposure type
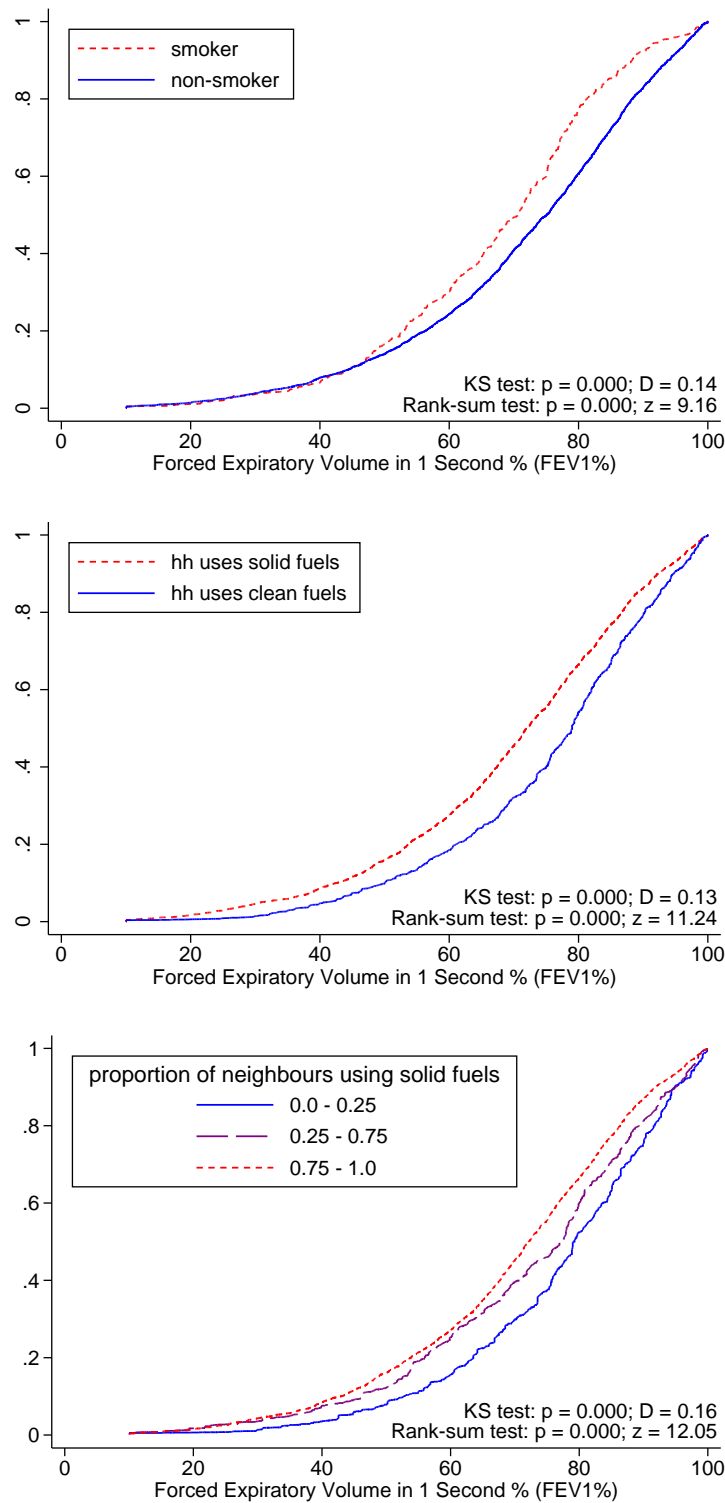
Source: SAGE 2007-08. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (liters) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible). Higher values denote lesser lung obstruction. hh: household

Figure 2: Lung obstruction by household and neighborhood exposure, and two measures of socioeconomic status
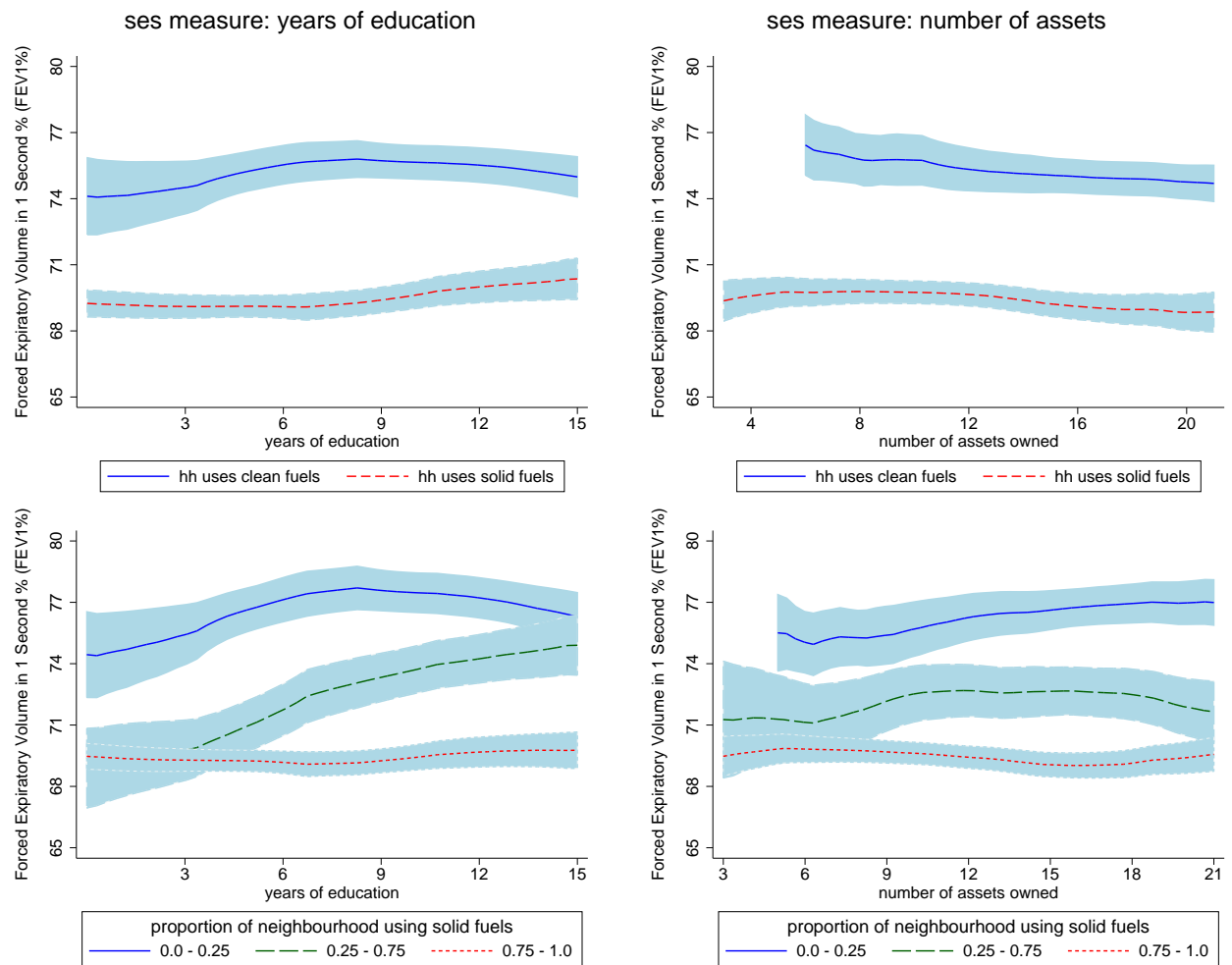
Source: SAGE 2007-08. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (liters) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible). Higher values denote lesser lung obstruction. hh: household

Figure 3: Lung obstruction by neighborhood solid fuel use among households that use solid fuels and households who use clean fuels

Source: SAGE 2007-08. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (liters) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible). Higher values denote lesser lung obstruction. hh: household

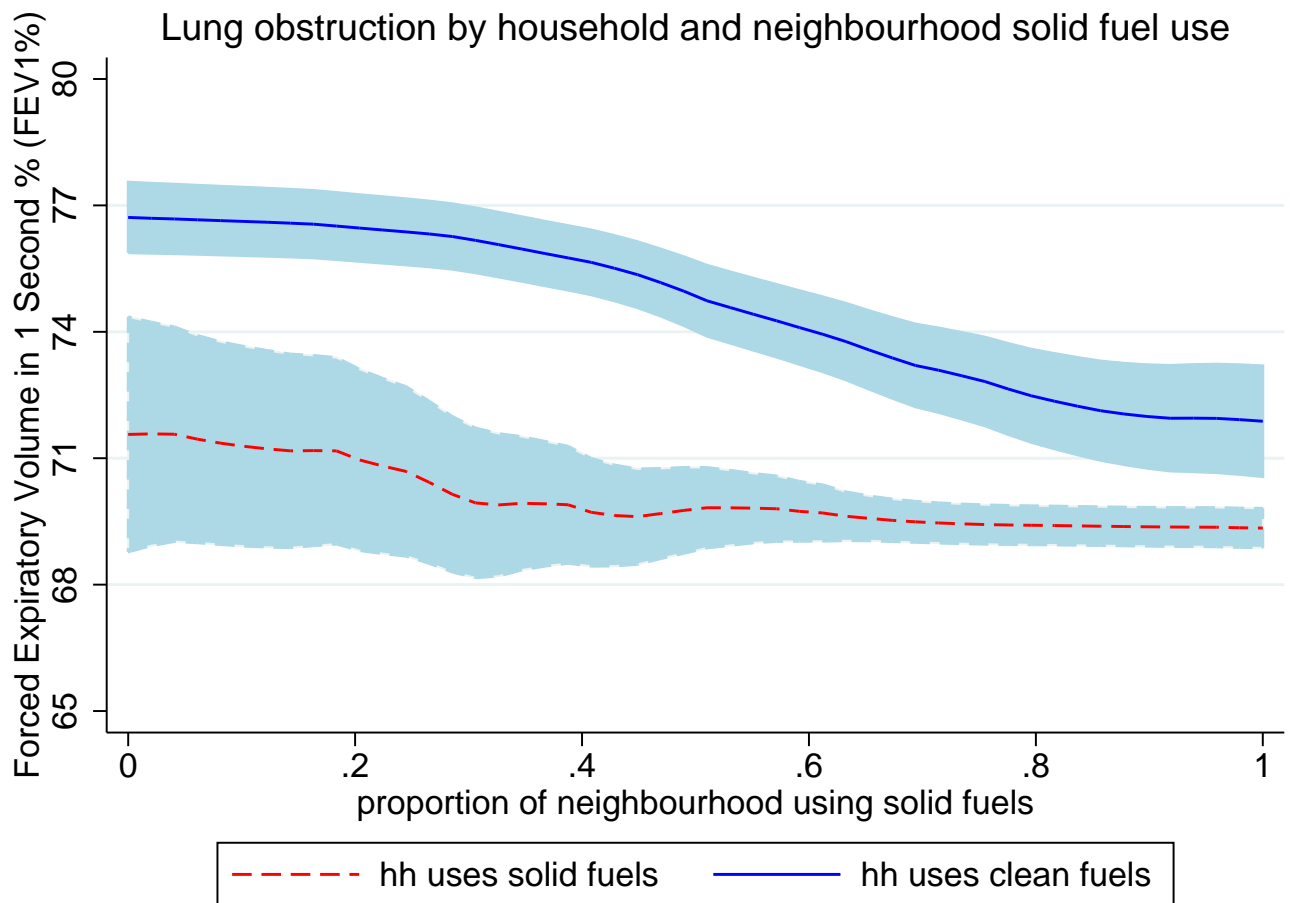Figure 4: Predicted FEV1% from OLS regressions with household and neighbourhood solid fuel interactions
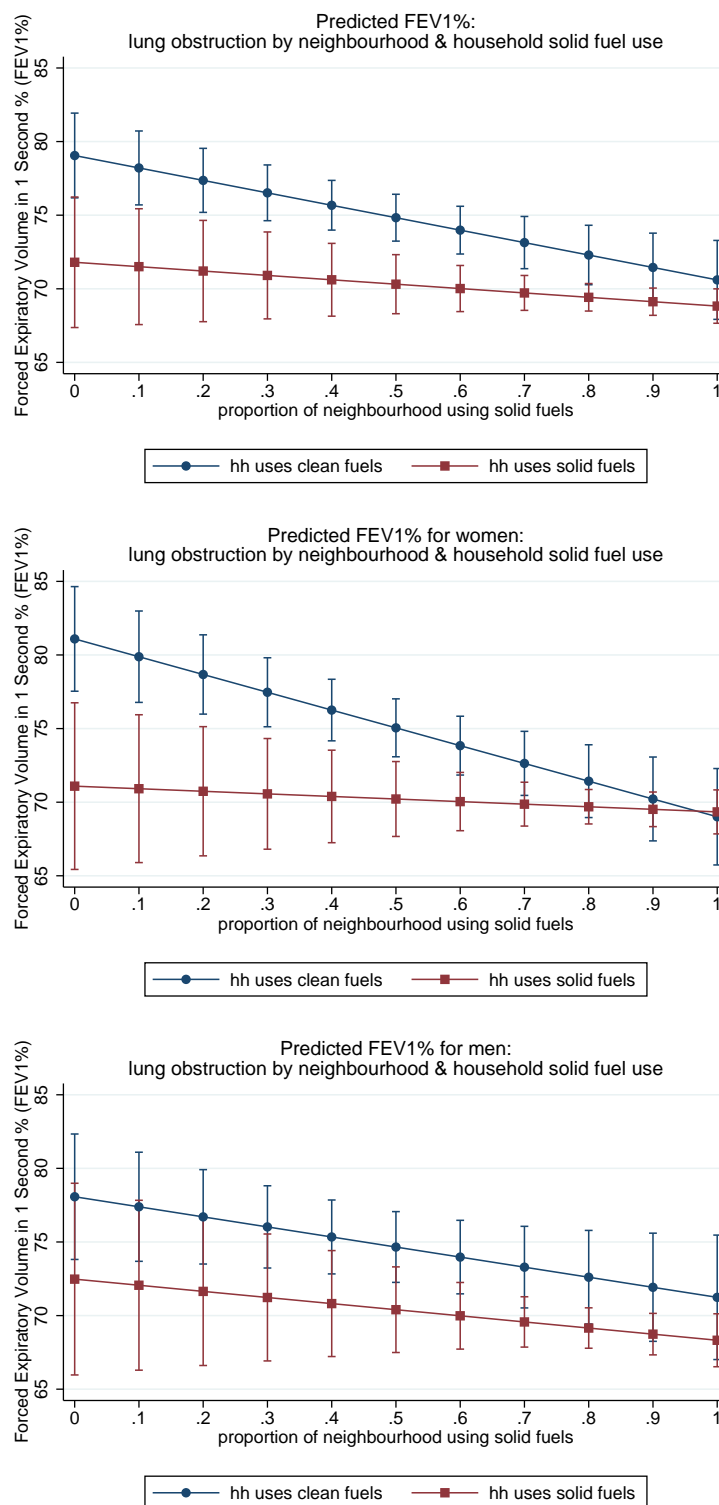

Source: SAGE 2007-08. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (liters) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible). Higher values denote lesser lung obstruction. Graphs show local polynomial regression. hh: household

SUPPLEMENTARY APPENDIX FOR ONLINE PUBLICATION

Appendix table A1 examines the association of smoking, household solid fuel use, and neighborhood solid fuel use with lung obstruction. The models in the table compare the association of the exposures when considered individually (models 1 - 3) and when considered together (models 4 - 6), the magnitudes of association for linear controls of age and socioeconomic status (models 1 - 5), and the magnitude of association for living in a rural area (models 1 - 6). Measures of socioeconomic status are not associated with lung obstruction. Model 6 in A1 is the same as Model 2 in table 2. Operationalizating the measures of socioeconomic status and age as dummies improves model fit, as shown by the r^2. The table shows coefficients for the rural dummy as well, finding that respondents in rural areas are not significantly different from urban respondents once household and neighborhood solid fuel use are adjusted for.

Appendix Tables A2 - A6 examine the robustness of the main result with controls for additional measures of socioeconomic status, placebo tests, falsification tests, mechanisms checks, and tests with alternative measures of respiratory health. The main findings from these tables are discussed in section 3.4.

Table A1: Association between smoking, household use of solid fuels, and neighborhood solid fuel use with lung obstruction: Models with linear controls for age, assets, and education

	(1)	(2)	(3)	(4)	(5)	(6)
smoke tobacco	-3.909***			-3.484**	-3.673***	-3.764***
	(1.069)			(1.066)	(1.062)	(1.040)
household uses solid fuels		-5.241***		-2.968*	-3.491**	-3.762**
		(1.427)		(1.261)	(1.230)	(1.170)
neighourhood solid fuel use			-9.188***	-6.127*	-6.424*	-6.088*
			(2.739)	(2.628)	(2.612)	(2.450)
female	-0.650	0.543	0.584	-0.367	-0.576	-0.859
	(0.936)	(0.828)	(0.828)	(0.877)	(0.939)	(0.908)
rural	-2.691*	-0.533	1.301	1.320	1.384	1.382
	(1.131)	(1.367)	(1.898)	(1.816)	(1.804)	(1.672)
age	-0.124***	-0.134***	-0.134***	-0.124***	-0.127***	
	(0.021)	(0.021)	(0.021)	(0.021)	(0.024)	
assets					-0.070	
					(0.086)	
education					-0.048	
					(0.092)	
state fixed effects	X	X	X	X	X	X
age in years (as dummies)						X
asset count (as dummies)						X
years of education (as dummies)						X
constant	74.108***	76.179***	77.875***	78.248***	80.334***	74.768***
	(2.044)	(1.993)	(2.125)	(2.140)	(2.556)	(2.619)
r-squared	0.048	0.053	0.053	0.060	0.060	0.090
Observations	9551	9551	9551	9551	9551	9551

Source: WHO SAGE 2007-08

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air one can exhale in 1 second to the total volume of air one can exhale after inhaling as deeply as possible.

Table A2: Robustness I: Additional controls for household and neighborhood socioeconomic status

	(1)	(2)	(3)	(4)
smoke tobacco	-3.764***	-3.792***	-3.774***	-3.770***
	(1.04)	(1.04)	(1.04)	(1.04)
household uses solid fuels	-3.762**	-3.577**	-3.352**	-3.363**
	(1.17)	(1.15)	(1.14)	(1.14)
neighbourhood solid fuel use	-6.088*	-6.397*	-7.936**	-7.492*
	(2.45)	(2.59)	(2.78)	(3.18)
log monthly expenditure		1.012+	0.838	0.839
		(0.60)	(0.61)	(0.61)
log monthly expenditure squared		-0.667	0.255	0.284
		(1.21)	(1.32)	(1.31)
average neighbourhood assets			-0.342	-0.385
			(0.24)	(0.27)
average neighbourhood education				0.114
				(0.38)
demographic controls	X	X	X	X
place controls	X	X	X	X
ses controls	X	X	X	X
constant	74.768***	72.409***	71.927***	71.393***
	(2.62)	(10.95)	(10.83)	(10.80)
observations	9551	9551	9551	9551

Source: WHO SAGE 2007-08

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air one can exhale in 1 second to the total volume of air one can exhale after inhaling as deeply as possible.

Table A3: Robustness II: Tests for bias in exposures

specification / sample	psu with n > 19	iv with split psu	district level solid fuel use as an exposure		kerosene separately
	(1)	(2)	(3)	(4)	(5)
smoke tobacco	-3.693**	-3.596***	-3.740***	-3.758***	-3.762***
	(1.136)	(1.069)	(1.049)	(1.041)	(1.040)
household uses solid fuels	-2.777*	-7.324***		-3.748**	-3.710**
	(1.192)	(1.817)		(1.165)	(1.176)
household uses kerosene					0.827
					(3.141)
neighbourhood solid fuel use	-6.223*	-7.595*		-5.460*	-3.710**
	(2.548)	(3.466)		(2.477)	(1.176)
proportion using solid fuels in district			-6.549*	-1.998	
			(3.242)	(2.972)	
demographic controls	X	X	X	X	X
place controls	X	X	X	X	X
ses controls	X	X	X	X	X
constant	74.702***	85.363***	74.101***	76.013***	74.702***
	(2.610)	(4.304)	(3.345)	(3.347)	(2.610)
observations	8519	9361	9551	9551	9551

Source: WHO SAGE 2007-08

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001. Outcomes: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (litres) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible).

Table A4: Placebo tests for smoking, household solid fuel use, and neighborhood solid fuel use

A. Smoking tobacco

	(1)	(2)
chew tobacco	-0.28	
	(0.86)	
drink alcohol		-1.318
		(1.34)
smoke tobacco	-3.772***	-3.522***
	(1.10)	(1.06)
demographic controls	X	X
place controls	X	X
ses controls	X	X
constant	77.408***	77.416***
	(2.99)	(2.95)
observations	9551	9551

B. Household solid fuel use

	(1)	(2)
hh owns toilet	0.742	
	(0.94)	
hh electrified		0.242
		(0.95)
hh uses solid fuels	-5.501***	-5.766***
	(1.05)	(1.07)
demographic controls	X	X
place controls	X	X
ses controls	X	X
constant	82.242***	82.429***
	(3.16)	(3.32)
observations	9551	9551

C. Neighbourhood solid fuel use

	(1)	(2)
neighbourhood open defecation	-0.178	
	(1.95)	
neighbourhood electrification		-0.83
		(2.05)
neighbourhood solid fuel use	-7.309***	-7.852***
	(2.20)	(1.83)
demographic controls	X	X
place controls	X	X
ses controls	X	X
constant	83.252***	84.130***
	(3.34)	(4.00)
observations	9551	9551

Source: WHO SAGE 2007-08

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001. All regressions are weighed using national individual weights. Outcome: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air one can exhale in 1 second to the total volume of air one can exhale after inhaling as deeply as possible. hh: household.

v

Table A5: Falsification tests: Association between other health outcomes and exposures

outcome:	FVC	height	word recall	walk time
	(1)	(2)	(3)	(4)
smoke tobacco	-0.038	0.002	-0.071	-0.061
	(0.058)	(0.004)	(0.105)	(0.047)
household uses solid fuels	-0.050	-0.000	-0.067	0.048
	(0.059)	(0.004)	(0.128)	(0.053)
neighbourhood solid fuel use	-0.008	0.002	-0.300	0.005
	(0.075)	(0.005)	(0.225)	(0.113)
demographic controls	X	X	X	X
place controls	X	X	X	X
ses controls	X	X	X	X
constant	2.765***	1.656***	4.725***	2.994***
	(0.414)	(0.014)	(0.275)	(0.161)
observations	9551	9551	9551	9551

Source: WHO SAGE 2007-08

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001. Outcomes: FVC, or Forced Vital Capacity is the total amount (measured in liters) of air breathed out in the spirometry test. Height is measured in meters. Word recall is based on the delayed recall of ten words. Timed walk measures the time taken (in seconds) to walk 4 meters by

Table A6: Mechanism check and tests with alternative health outcomes

test:	mechanism check: satellite PM 2.5			alternative measures of respiratory health		
specification:	ols			logit		
outcome:	PM2.5	PM2.5	PM2.5	FEV1s	have COPD (FEV1% < 70%)	self-reported lung disease
	(1)	(2)	(3)	(4)	(5)	(6)
smoke tobacco	-2.962**	-3.696***		-0.189***	1.422*	1.526*
	(1.055)	(1.031)		(0.030)	(0.195)	(0.282)
household uses solid fuels	-3.314**	-3.716**		-0.139***	1.467*	1.919*
	(1.175)	(1.164)		(0.037)	(0.220)	(0.505)
neighbourhood solid fuel use	-3.733*	-6.164*	22.132***	-0.162**	1.502*	1.966*
	(1.793)	(2.449)	(6.438)	(0.061)	(0.290)	(0.668)
district level PM 2.5 (ug / m3)	-0.054*	-0.034				
	(0.022)	(0.038)				
demographic controls	X	X		X	X	X
place controls		X		X	X	X
ses controls	X	X		X	X	X
height				X		
constant	86.154***	75.595***	22.783***	0.755		
	(2.730)	(2.905)	(3.516)	(0.565)		
observations	9551	9551	9551	9551	9551	9551

Source: WHO SAGE 2007-08

Note: Robust standard errors, clustered at the level of the primary sampling unit, in parentheses. + p<.1, * p<.05, ** p<.01, *** p<.001.
Outcomes: FEV1% (FEV1s/FVC as a percentage) is the ratio of the volume of air (litres) one can exhale in 1 second to the total volume of air one can exhale (after inhaling as deeply as possible). FEV1s or Forced Expiratory Volume in 1 second (measured in liters) is the total amount of air breathed out in the first second of the spirometry test. Respondents are classified as having COPD if FEV1% was below 70%, based on GOLD criteria. Respondents classified as having lung disease based on self-reported symptoms according to WHO SAGE criteria, of having shortness of breath or experiencing coughing or wheezing for more than ten minutes and coughing sputum or phlegm for most days of a month for the past three months in a year.

Figure A1: Histogram of number of respondents per cluster