AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Journal of Agricultural and Resource Economics: Preprint ISSN: 1068-5502 (Print); 2327-8285 (Online)
doi: 10.22004/ag.econ.347701
Copyright 2024 the authors

Biomass Cooking Fuel and Schooling Outcomes: Empirical Evidence From Ethiopia

Dalia Fadly, Francisco Fontes, and Miet Maertens

Empirical evidence on the links between biomass fuel and human capital development is limited. This paper estimates the impact of biomass cooking fuel on child schooling using a panel of household survey data (2011-2016) from Ethiopia. Applying an instrumental variable approach and the extended probit model, we find that biomass cooking fuel decreases the probability of school enrollment and increases the probability of school absenteeism. This result is predominantly driven by work-related reasons and is more prominent among boys. Consequently, improved access to cleaner cooking fuels could have positive impacts on important development indicators such as better educational outcomes.

Key words: Biomass Fuel, Health, Smoke, Child Labor, School Absenteeism, Gender, Instrumental Variable

JARE Preprint

Introduction

Globally 2.8 billion people rely on traditional biomass fuels (wood, dung and crop residues) for cooking, lighting and/or heating, with an estimated 4 million premature deaths per year attributable to indoor air pollution (Ezzati and Kammen, 2002; Kelly et al., 2018; International Energy Agency, 2019; Energy Sector Management Assistance Program, 2020). In Sub-Saharan Africa, given that an estimated 80% of the population relies on biomass fuels for their energy needs (International Energy Agency, 2022) and that children younger than 15 years constitute 42% of the population in 2022 (United Nations, 2022), the link between cooking fuel and education is of particular importance. In such a context, this link could imply important synergies between the Sustainable Development Goals (SDGs) of ensuring healthy lives (SDG 3), increased access to quality education (SDG 4) and ensuring universal access to affordable, reliable, and modern energy services (SDG 7.1) (Scharlemann et al., 2020).

The literature focusing on the impact of biomass fuel use for cooking on the well-being of children has found that using biomass fuels for cooking is associated with negative health impacts for children (Bruce et al., 2013;Owili et al.,2017; Kurata, Takahashi, and Hibiki, 2020). Environmental chores, such as fuelwood collection, negatively impacts several schooling outcomes including enrollment, literacy and learning difficulties (Gebru and Bezu, 2014; Scheurlen, 2015; Levison, DeGraff, and Dungumaro, 2018; Choudhuri and Desai, 2021; O'Brien, Do, and Edelson, 2021). With regards to school attendance, however, the results are mixed, with

Dalia Fadly (corresponding author,daliafadly@gmail.com) is a former post-doctoral research fellow at the division of bio-economics, Department of Earth and Environmental Sciences, Katholieke Universiteit of Leuven, Francisco Fontes is a policy analyst at the European Commission, Miet Maertens is head of the Agricultural and Rural Development Economics research group, professor at the division of bio-economics, Department of Earth and Environmental Sciences, Katholieke Universiteit of Leuven.
We appreciate the comments received by the two anonymous reviewers on an earlier version of the manuscript. We acknowledge the funding of The Research Foundation Flanders (FWO), Belgium, grant number 202810/1255221N.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
Review coordinated by Christian A. Langpap.

some authors finding a negative impact of using biomass fuels on school attendance in Vietnam and India (O'Brien, Do, and Edelson, 2021; Biswas and Das, 2022), while others find no effect in Malawi (Kelly et al., 2018).

Given the mixed results in the literature, this paper empirically assesses the relationship between the use of solid biomass fuels (e.g., fuelwood and dung) as a primary cooking fuel on school attendance in Ethiopia and tests whether the impact differs by gender and by reason of absenteeism (work- or sickness-related). We estimate this relationship using household survey data of three waves of the Ethiopia Socioeconomic Surveys (Central Statistical Agency and the World Bank, 2011, 2013, 2016). We find that using a solid biomass cooking fuel decreases the probability of school enrollment and increases the probability of absenteeism for enrolled children. This effect is more pronounced for work-related absenteeism, and for boys.

The paper contributes to the literature in three main ways. First, the study contributes to the limited evidence on the link between household cooking fuel choice and schooling in low- and middle-income countries with quantitative evidence from Ethiopia. Second, the study is one of the first to provide insights on the relative importance of the two main channels that explain absenteeism, namely health-related reasons or time-use through work-related reasons (e.g., environmental chores, agricultural work). Understanding the relative importance of these channels is relevant for policy-makers who need to decide how to best tackle the root causes of absenteeism. Third, we disaggregate the effects by gender to reveal heterogeneous impacts of fuel use. Domestic chores, such as collecting biomass fuel and cooking are likely to be gendered, creating gendered impacts on school attendance.

From a policy perspective, the Ethiopian context is pertinent, as school enrollment and attendance rates remain low[1], the reliance on biomass fuels for cooking is high (92% of households) (Guta, 2014), and the government is committed to addressing both challenges. The focus on the relative importance of health –vs- work-related absenteeism is key, as different types of absenteeism require different policies.

The paper is organized as follows. Section 2 presents a review of the literature on the link between cooking fuel and schooling. Section 3 outlines the data and empirical methodology. Section 4 presents and discusses the empirical results, while Section 5 concludes and suggests some policy implications.

2. Literature review

The literature on the links between household fuel choice and educational outcomes generally considers either the health impacts of biomass fuels or time-use impacts as an explanation for the negative impacts on schooling outcomes. However, the relative importance of the health and time-use channels explaining absenteeism remains, to our knowledge, unaddressed. Moreover, the studies, which explicitly consider the direct link between cooking fuel choice and educational outcomes, are scant and find mixed results. Kelly et al. (2018) find no effect of using cleaner burning biomass-fuel cook stoves on primary school absenteeism in Malawi. Evidence from Vietnam and India shows that using biomass fuels is associated with a large negative impact on school attendance (O'Brien, Do, and Edelson, 2021; Biswas and Das, 2022), with this effect being higher for girls in the Indian context due to traditional household roles (Biswas and Das, 2022).

The literature focusing on environmental chores highlights that collecting cooking fuel is a labor-intensive and time-consuming task that often falls on children and leads to lower school attendance as children reallocate their time to fuel collection (Nankhuni and Findeis, 2004; Palmer and Macgregor, 2009; Köhlin, O.Sills, K. Pattanayak, and Wilfong, 2011; Gebru and Bezu, 2013; DeGraff, Levison, and Dungumaro, 2017;Choudhuri and Desai, 2021). Empirical studies testing this link have found that engaging in environmental chores lowers attendance and enrollment rates

[1] The enrollment rate for primary school (% net) and secondary school (% gross) is 85% and 35% in 2015, respectively (World Bank, 2021)

JARE Preprint

in several countries, including Ghana, South Africa, Malawi, Ethiopia, and Kenya (Nankhuni and Findeis, 2004; Ndiritu and Nyangena, 2011; Porter et al., 2012; Gebru and Bezu, 2013;Choudhouri and Desai, 2021), as well as child literacy in rural Ethiopia (Beyene, Mekonnen, and Gebreegziabher, 2014) and learning difficulties in Tanzania (Levison, DeGraff, and Dungumaro, 2018).

The literature focusing on the health channel finds that children are particularly at risk to the adverse health effects of indoor air pollution in part because they spend more time at home (Bruce et al., 2013; Patel, Patel, and Kumar, 2019; Maji, Mehrabi, and Kandlikar,2021). Evidence suggests that indoor air pollution linked to fuel use contributes to lower birth weight, chronic malnutrition, and respiratory infections in children, all of which could affect schooling outcomes and school absenteeism (Glewwe and Miguel, 2007; Kurata, Takahashi, and Hibiki, 2020). Beyond the effects on physical health, exposure to indoor smoke has been shown to impair cognitive ability of children, which negatively affects learning, short-term memory recall and visuospatial processing (Smith et al., 2011).

There is thus strong evidence that both time-use and health risks emanating from the use of biomass cooking fuel can negatively affect schooling outcomes. However, studies so far have not investigated the relative importance of these two channels and whether impacts differ by child gender. In this paper, we seek to quantify the association between the use of biomass cooking fuel and child school absenteeism in Ethiopia. We focus on the relative importance of work-related absenteeism, which captures fuel collection time, and health-related absenteeism, that proxy the impacts from indoor air pollution.

This question is especially relevant in the Ethiopian context where the heavy reliance on biomass energy represents a challenge for socio-economic development. Only 48% of the population had access to electricity in 2019 (Ministry of Water, Irrigation and Energy, 2019; World Bank, 2021) and 98% of households in rural areas largely depend on traditional biomass fuels for cooking and heating (International Energy Agency, 2019). The increasing demand for fuelwood due to population growth combined with a lack of access to biomass energy substitutes exerts considerable pressure on forests and vegetation stocks (Benti et al., 2021). With regards to education, Ethiopia has below-average educational completion rate compared to its sub-Saharan African peers, with completion rates of 61%, 29% and 8% for primary, secondary, and tertiary education, respectively (World Bank, 2021). Dropout rates are high, which is explained by several factors, including the need to engage in paid or unpaid work, economic shocks, inadequate learning resources, and conflict (Federal Ministry of Education, 2021).

JAR Preprint

3. Data and Methodology

3.1 Data

We use data from the Ethiopia socio-economic survey (ESS) carried out by the Central Statistics Agency of Ethiopia (CSA) and the World Bank. The ESS uses a nationally representative sample[2] of households living in rural and urban areas[3], where urban areas include both small and large towns. The sample is drawn from a total of 433 enumeration areas (EA) which are selected in the first stage based on probability and are proportional to the size of the total EA in each region. In the second stage, households were selected from each EA using stratified random sampling.

[2] The sample is not representative for small regions, including Afar, Benshangul Gumuz, Dire Dawa, Gambella, Harari, and Somali regions.

[3] In 2011, the sample covers rural and small town areas only (333 EA) while in subsequent years (2013 and 2015) both rural and urban areas are covered, including 433 EA. Out of the 433 EAs 290 were rural, 43 were small town EAs, and 100 were EAs from major urban areas.

Table 1: Variables Description and Statistics of Pooled Sample

		Pooled Sample	
Variable	**Variable Description**	**Mean**	**Std. Dev.**
Child Characteristics (N=28,537)			
Enrollment	A binary variable=1 if the child is currently attending school	0.580	0.493
Absenteeism	A binary variable=1 if the child has been absent from school for more than a week in the past month	0.066	0.248
Absent sick	A binary variable=1 if the child has been absent from school for more than a week in the past month for being sick	0.027	0.163
Absent work	A binary variable=1 if the child has been absent from school for more than a week in the past month for work reasons	0.038	0.192
Average enroll	The share of children in the woreda who are enrolled in school	0.256	0.053
Age	The age of the child (years)	11.488	4.169
Gender	The gender of the child, where 1=female, 0=male	0.494	0.500
Biological child	A binary variable=1 if the head of the household is the father of the child	0.832	0.374
Household characteristics (N=13,724)			
Biomass	A binary variable=1 if the household uses primarily biomass fuel for cooking	0.914	0.280
Age of head	The age of the head of the household (years)	46.141	12.880
Gender head	A binary variable=1 if the head of the household is a female	0.223	0.416
Size	The number of household members permanently living together	6.325	2.242
Educ mother	A binary variable=1 if the mother has at least completed primary schooling	0.235	0.424
Educ father	A binary variable=1 if the father has at least completed primary schooling	0.403	0.490
Parent agriculture	A binary variable=1 if at least one of the parents works in agriculture	0.781	0.414
Income [a]	Real net income of the household (100 Ethiopian Birr) in the last 12 months	95.548	671.021
TLU [b]	Tropical livestock unit, total	2.635	6.809
Forest share	Share of the forest cover in the woreda (%)	3.374	11.559
Center distance	Distance to the nearest population center with +20,000 residents (KM)	36.558	32.467
School distance	Distance to the nearest primary school (KM)	0.872	4.086

Notes: [a] Total net income of the household is calculated as the total income earned from all sources such as from agriculture activities, self-employment, wages, and government transfers, minus all expenses incurred in generating that income, such as cost of seeds, fertilizers, rental of land/equipment, transport, and hired labor. We used the consumer price index with 2010 as a base year to calculate the real net income.
[b] TLU is calculated using the following conversion factors 0.01 for chicken,0.1 for goat and sheep,0.7 for cows, buffaloes and calves,0.2 for pigs (Food and Agriculture Organization, 2018).

JARE Preprint

The dataset covers all regions, including Addis Ababa and consists of three waves. The first wave of the dataset (2011-2012) focuses exclusively on households in rural areas and small towns and includes 3,969 households. In the second wave (2013-2014), 3,776 of the original 3,969 households were re-interviewed and an additional 1,486 households were added to the sample. In the third wave (2015-2016), the sample reached 4,954 households, out of which 3,699 households were present in the (2011-2012) wave. The analysis in this paper is carried out at the individual-level (children) and consists of children with ages ranging from 5 to 19 years old[4] using the pooled sample. As a result, the final sample across the three waves consists of 28,537 child-year observations and 13,724 household-year observations distributed over 4,435 households. Using the ESS data, we construct variables related to child school absenteeism, demographic characteristics, household fuel use, and socio-economic characteristics, summarized in Table 1.

The two main variables of interest in the analysis are school absenteeism and type of cooking fuel. School absenteeism is specified as a binary variable at the child-year level that takes the value 1 if the child has been absent for more than a week in the past month of the school year due to work- or sickness-related reasons, and takes the value 0 otherwise. It should be noted that since school absenteeism is conditional on school enrollment, the outcome variable is observed only if the child is enrolled in school. While the one-month recall period is sufficiently brief for respondents to recall, the framing of the question means that we are not able to accurately measure the severity of absenteeism. This is because the time span implied by the phrasing "more than a week"' could encompasses anything above 5 school days. The variable cooking fuel is a binary variable at the household-year level that equals 1 if a household reported using biomass fuels, including fuelwood, dung or crop residues, as the main cooking fuel, and 0 otherwise. Alternative cooking fuel options such as kerosene, charcoal, biogas or electricity are purchased from the market rather than collected.

Absenteeism is a relatively rare event, with only 6.6% of observations over the different survey waves having been absent for over one week in the preceding month (2.7% and 3.8% for absenteeism that is related to sickness or to work, respectively as shown in table 1). Yet, the sample size at the child-year level is sufficiently large to detect with a minimum degree of statistical precision how a given variable affects this rare event. This is one of the key reasons why we opted for using multiple waves of data and focused on individual-level, rather than household-level data, as it ensures a sufficiently large sample, allowing us to more effectively capture the determinants of absenteeism.

As shown in Table 1, 91% of the children in the sample live in households that reported using biomass fuel (fuelwood, dung and crop residues) as their primary cooking fuel and only 58% of children were enrolled in school during the period (2011-2016). We observe in table (2) a statistical difference in the average rate of absenteeism, both for overall absenteeism and for sickness, between children in households using versus not using biomass fuels. Absenteeism among children from households using biomass fuel is higher, especially for work-related reasons. There is also a statistically different mean rate of school enrollment across the two groups of households as the average rate of enrollment among non-biomass users is higher (77%) compared to households using biomass fuel (56%). The two groups of households are also different in several ways.

As highlighted in Table 2, the parents of children in households who use biomass fuel are characterized by lower average levels of formal education, are more likely to be employed in the agricultural sector and more likely to be male-headed, similarly in households with non-enrolled children. These households also exhibit lower incomes, larger household sizes and higher levels of livestock ownership. Finally, children in households that use biomass fuels reside in areas with greater forest cover. In table A1 (Appendix), we show the summary statistics of households across

[4] Although formal schooling in Ethiopia is until 18 years, the sample includes those who are beyond this age but still attending school. These households represent a very small percentage of the sample and are likely to be students who either enrolled later or have failed and are repeating the grade.

JAAE Preprint

Table 2: Descriptive Statistics and Comparison of Means by the Type of Cooking Fuel and Enrollment Status

	Biomass Fuel users		Non-Biomass fuel users		Non-enrolled children		Enrolled children	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Enrollment	0.563	0.496	0.769	0.422***	0.000	0.000	1.000	0.000
Absenteeism	0.069	0.254	0.039	0.193***	0.000	0.000	0.066	0.248
Absent sick	0.027	0.162	0.029	0.168	0.000	0.000	0.027	0.163
Absent work	0.042	0.201	0.010	0.097***	0.000	0.000	0.038	0.192
Average enrollment	0.255	0.052	0.273	0.056***	0.245	0.049	0.265	0.054***
Age	11.385	4.141	12.589	4.307***	11.065	4.839	11.793	3.577***
Gender	0.488	0.500	0.551	0.498***	0.497	0.500	0.491	0.500
Biological child	0.844	0.363	0.702	0.458***	0.796	0.403	0.858	0.349***
Biomass	1.000	0.000	0.000	0.000	0.953	0.212	0.887	0.317***
Age of head	46.319	12.856	44.238	12.981***	46.250	13.621	46.062	12316
Gender head	0.214	0.410	0.319	0.466***	0.208	0.406	0.233	0.423**
Size	6.386	2.225	5.667	2.314***	6.293	2.323	6.348 (	2.182**
Educ mother	0.208	0.406	0.533	0.499***	0.139	0.346	0.305	0.460***
Educ father	0.380	0.485	0.639	0.480***	0.287	0.452	0.486	0.500***
Parent agriculture	0.818	0.386	0.388	0.487***	0.780	0.414	0.782	0.413
Income	89.026	680.236	165.146	558.770***	87.185	677.537	101.593	666.226*
TLU	2.811	7.068	0.755	2.007***	2.899	6.839	2.444	6.780***
Forest share	3.556	11.925	1.430	6.149***	3.147	11.120	3.538	11.864***
Center distance	38.229	32.165	18.726	30.255***	39.465	33.158	34.456	31.795***
School distance	0.869	3.960	0.911	5.240	0.948	4.216	0.818	3.988***

Note: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level. To test if the difference in means between groups is significant or not, we used the *Z*-test for binary variables and the *t*-test for continuous variables.

JARE Preprint

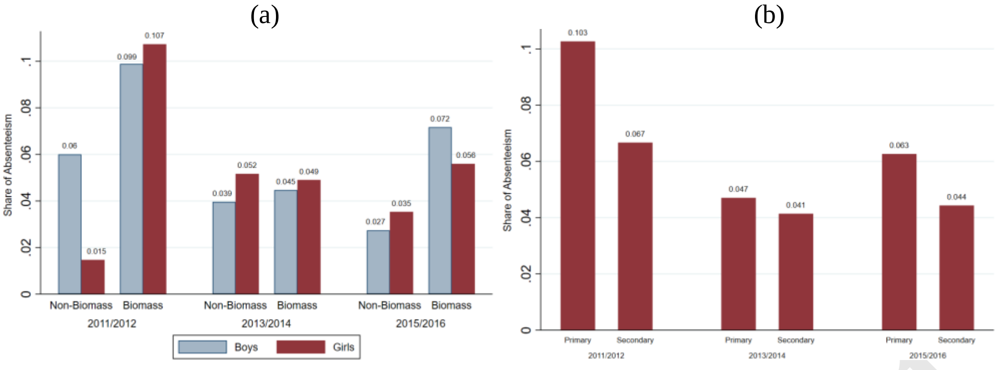

Figure (1): School Absenteeism Trend by Cooking Fuel (Panel a) and by Grade Level (Panel b)

survey years (2011-2016), indicating a declining trend in biomass fuel usage for cooking, although many households still rely on it. While 86 % of households in Addis Ababa use clean fuels for cooking, non-biomass fuel users are also present in other cities. For example, in Diredwa and Somali, 16% and 14% of households, respectively, do not use biomass for cooking. However, this share drops to just 3.3% in Afar and 5.3% in Southern Nations, Nationalities and Peoples' Region (SNNP).

Over time, absenteeism from school has decreased, real income of parents and their education have risen, and a higher share of women have become heads of their household. During this period overall school enrollment also increased. Since, we only observe absenteeism for enrolled children we also wanted to investigate whether enrolled and non-enrolled children seemed to differ in their covariates and this seems to be supported by the results in Table 2. Specifically, enrolled children tend to live in Woredas with higher overall enrollment rates, have more educated parents, and live in households that are richer and that are less likely to use biomass fuels less. These systematic differences suggest that there may be selection bias, an issue we will return to in the methodology section.

Since the literature often emphasizes the gendered nature of absenteeism, we also show some of the trends related to absenteeism by type of cooking fuel and gender in Figure (1)-panel (a). We note that, while in some years female absenteeism is higher, but that this is not always the case in our sample and there is no clear discernible pattern. Finally, panel (b) indicates that absenteeism rates in primary school is higher than in secondary school, but both follow trends that are relatively similar.

3.2 Econometric Estimation

In order to estimate the relationship between biomass cooking fuel and the probability of a child being absent from school, two key empirical challenges need to be addressed. First, there is a concern about endogeneity, where unobservable factors (such as preferences, cultural beliefs) could bias the choice towards biomass cooking fuel and be correlated with the outcome variable. For instance, households who have certain unobservable characteristics (e.g., value more education, more aware of health impacts of biomass fuel) may bias the choice of cooking fuel and could affect education. To minimize the problem of endogeneity, we use an instrumental variable (IV) approach to estimate the effect. In order to be valid, the instrument must satisfy several conditions, namely: (1) it must be correlated with the variable capturing the use of biomass fuel for cooking (relevance condition), (2) it must not be correlated with the source of heterogeneity (exogeneity condition), and (3) it must not have any direct impact on schooling outcomes (exclusion restriction).

JARE Preprint

As an instrument for the cooking fuel variable, similar to Biswas and Das (2022), we use exogenous variation in the share of forest cover in a given woreda[1] that can be hypothesized to be correlated with biomass fuel use but not correlated to child schooling outcomes through channels other than biomass fuel use. The relevance and validity of the instrument are tested using various tests, which are reported in the tables and discussed, in the next section. Moreover, while no test can fully prove the validity of a given instrument, we use a falsification test that has widely been used in the literature (Di Falco, Veronesi, and Yesuf, 2011; Fontes, 2020; Daidone and Fontes, 2023) which tests whether the instrument satisfies some key admissibility criteria. The logic of this falsification test is that, if a variable is a valid instrument, then it should affect the use of biomass fuel for cooking but should not affect absenteeism among households that do not use biomass fuel.

The second empirical issue that needs to be addressed is related to sample selection biases arising from the fact that not all school-age children in the sample are enrolled in school in a given year. As a result, the sample data (children who were enrolled at school) are a non-random subset of the population of children, even after controlling for explanatory variables. While it is impossible to know from the outset whether enrolled and non-enrolled children differ in terms of their unobservable characteristics, the fact that the means for certain key variables are consistently different suggests that these two groups may differ from one another. In the presence of sample selection, estimates from econometric models could be biased (Marchenko and Genton, 2012). Therefore, we apply an approach that includes all children in the pooled sample in the first step (selection equation to estimate school enrollment denoted by the variable $enroll$) and in the second step, estimate school absenteeism (using enrolled children only).

Given the above discussion, it is important to acknowledge that while there are established links between cooking fuel and education, some confounding factors such as poverty, income levels, infrastructure, and cultural preferences play a significant role in determining a child's access to schooling as well as the choice of cooking fuel. While it is impossible to fully take into account such unobservable factors, the fact that our empirical approach focuses on the use of an IV and that it controls for sample-selection is likely to reduce the likelihood that these factors drive the results that we will find later.

With regards to the choice of model, we have opted for the Probit model nested within the Extended Regression Model (ERM) to explore how the choice of cooking fuel impacts school absenteeism, given that our dependent variable is binary. We use an extended probit regression model (*eprobit*) that simultaneously tackles challenges such as endogenous covariates, non-random distribution of the treatment variable, and sample selection (Angrist, 2001). This type of model thus allows for observations that are correlated within panels or within groups and Heckman-type sample selection (Wooldridge, 2014). The eprobit model is a system of three equations (equations 1-3), explained below.

Equation (1) is the outcome equation where we estimate school absenteeism ($Absent$) as a function of cooking fuel use ($cookfuel$) and control for several child, parent, and household socioeconomic and demographic characteristics. The baseline probit regression model is expressed as follows:

$$(1) \qquad Outcome\ equation{:}\quad Pr\left(Absent_{ijt}=1 \middle| X_{ijt},\gamma_{jt}\right)= \Phi(\beta_0+\beta_1 cookfuel_{jt} \\ +\beta_k X_{ijt}+\beta_\omega \gamma_{jt}+\lambda\,\mathrm{t}+\varepsilon_{1ijt})$$

Where subscripts *t*, *j*, and *i* denote the year, the household, and the individual child, respectively; $\Phi\,(.)$ represents the standard normal distribution function; β_1 is the parameter of interest that denotes the estimated effect of biomass cooking fuel *(cookfuel)*; β_k $d$enotes the vector of coefficients for the control variables at the individual level (X_{it}); β_ω refers to the coefficient of household characteristics of individual i, λ denotes the coefficient for time trend (i.e., a variable capturing the survey year), and ε_{ijt} is the idiosyncratic error term. Moreover, a woreda fixed effect

[1] Districts of Ethiopia are called woredas.

JARE Preprint

is added to control for potentially omitted variables that are unobserved in the data set and which helps to control for spatial differences in local policies, infrastructure and resource endowment not captured by the other variables. The model estimates the coefficients β that maximize the likelihood of observing the actual outcomes given the predictor variables. In our set of results, we cluster the standard errors at the household level as the sampling strategy was based on household as a unit of observation.

However, since the outcome variable in equation (1) is only observed for enrolled children, if this is directly estimated using OLS/Probit without taking into account sample selection (i.e., the fact that absenteeism is unobserved for children not enrolled and this being correlated with certain unobservable characteristics), this might lead to a bias in the estimated coefficient. Since there are several unobserved factors that may affect both the enrollment decision and absenteeism (e.g., importance parents give to schooling/education, skill of the child, etc…), it is important to control for this in the empirical approach by applying the selection equation expressed in equation (2) below.

$$(2) \quad Selection\ equation\text{: } \Pr\left(Enroll_{ijt} = 1 \middle| X_{ijt}, \gamma_{jt}\right) = \Phi(\beta_0 + \beta_1 cookfuel_{jt} + \beta_k X_{ijt} + \beta_\omega \gamma_{jt} + \lambda\, \mathrm{t} + \varepsilon_{2ijt})$$

Importantly, in order to estimate equations (1) and (2), it is also necessary to include at least one variable in the selection equation that is not in the outcome equation for the model. In this paper, we include the average rate of school enrollment in a given woreda in equation (2). We argue that the spillover effect of schooling decision among households due to peer effects or social interaction is a good predictor of school enrollment, but not necessarily of absenteeism (Lalive and Cattaneo, 2009). Finally, if we want to use an instrumental variable, we also need to estimate equation (3) which estimates the probability of using biomass cooking fuel (the endogenous variable) as a function of several controls and an instrument (i.e., the share of forest cover in the woreda).

$$(3) \quad IV\ equation\text{: } \Pr\left(Cookfuel_{jt} = 1 \middle| \gamma_{jt}\right) = \Phi(\beta_0 + \beta_1 forestshare_{wt} + \beta_\omega \gamma_{jt} + \varepsilon_{3t})$$

When estimating equations (1)-(3), above, we run several regressions. First, we focus on the overall sample without differentiating the effects by gender or type of absenteeism. Second, we estimate separate models for the two types of absenteeism (health[2]- and work-related[3] absenteeism). Third, given that the literature often finds that the burden of environmental chores in developing countries falls predominantly on girls and women (Biswas and Das, 2022), we estimate the models by gender to test for gendered impact of cooking fuel on schooling outcome.

We use several time-varying and time-invariant control variables that are commonly used in the literature. We account for child characteristics, such as gender, age and whether the child is the child of the household head as these are variables that have been shown to be significant factors in explaining participation in environmental chores, absenteeism and schooling outcomes (O'Brien, Do, and Edelson, 2021; Ndiritu and Nyangena, 2011; Assaad, Levison, and Zibani, 2010).

At the household-level, we include several variables, including household size, household income, parental education and sector of employment, distance from school, and characteristics of the household head. Household size is included as tasks are likely to be divided over more members in large households, which could impact absenteeism. In a similar fashion, we control for the household's total real income[4] over the past year, as household socioeconomic has been found to influence schooling outcomes. With regards to parental variables, we include both parental education and the sector of employment. We argue that more educated parents are more

[2] We use "being absent from school because of sickness" as the outcome variable.
[3] We use "being absent from school for work related reasons" as the outcome variable.
[4] Nominal income from different survey rounds is converted to real income in 2010 price levels using a consumer price index.

JAAE Preprint

likely to value the education of the child and discourage absenteeism (Haile and Haile, 2012;Mani et al., 2013) and that their employment status and sector of employment is likely to affect both the demand for household labour and schooling outcomes (Glick and Sahn, 2000; Assaad, Levison, and Zibani, 2010; Maertens and Verhofstadt, 2013). With regards to the characteristics of the household head, the rationale is that the age and gender of the household head to capture the household's experience, access to resources and vulnerability (as female-headed households may face higher vulnerability) all of which could affect schooling outcomes (Huisman and Smits, 2009). To address the degree of remoteness of a household, we include variables that capture the distance to the nearest primary school and distance to the nearest center with more than 20,000 residents. Beyond individual and household-level characteristics, we also control for time-invariant woreda-specific factors by including a woreda-specific binary variables and we also capture trends in school absenteeism over time by including the survey round as a covariate.

4. Results and discussion

4.1 Results

Table 3 shows the regression results on school absenteeism in general (i.e., pooling together work- and health-related school absenteeism), with the baseline probit model in column 1. The results of the extended probit regression model controlling for sample selection only can be found in columns 1a (outcome equation) and 1b (selection equation). The results that control for both sample selection and endogeneity are in columns 2a (outcome equation), 2b (selection equation) and 2c (endogeneity using forest cover in the Woreda as an instrument). As shown in Table 3, when estimating the model using a simple probit, we find no significant effect of biomass fuel. However, when estimating the *eprobit* model, we find a negative and statistically significant correlation between the error terms of the outcome and biomass use equation, which strongly suggests that the use of biomass fuel is endogenous binary variable. This provides a strong motivation for using the *eprobit* model.

The findings of our preferred specification (columns 2a-2c) reveal that the usage of solid biomass fuel as the main cooking fuel is associated with an increase in the probability of school absenteeism, with the effect being significant at 1% level. The estimated marginal effect of biomass fuel on absenteeism is significant as absenteeism is likely to increase by 20 percentage points in households that use biomass fuels for cooking. Moreover, we find that using biomass fuel for cooking lowers the probability of being enrolled in school.

In order to show that our instrument is relevant we carry out several tests. Column 2c in Table 3 shows a strong relationship between the instrument and the use of biomass fuels, which is confirmed by the Kleibergen–Paap Wald *F*-statistic (Kleibergen and Paap, 2006). The first-stage *F*-statistic on the excluded instruments is also above the conventional threshold in the literature, which suggests that the instrument is strong (Staiger and Stock,1997). As further evidence of the admissibility of the instrument, we also show the results of the falsification tests in Table A2 of the appendix. These results confirm that the IV (the share of forest cover in the woreda) is a statistically significant determinant of the use of biomass cooking fuel (first stage results of columns 1) but is not significant determinants of school absenteeism among children in households that do not use biomass fuel (columns 2). This provides further supporting evidence that the instrument does not influence school absenteeism through mechanisms other that the use of biomass fuel.

Regarding the control variables, most of them exhibit the expected sign. The coefficient on the time trends shows that school absenteeism has decreased and school enrollment has increased

JARE Preprint

Table 3: Determinants of School Absenteeism

	Probit model	eprobit with sample selection		eprobit with sample selection and IV		
	Eq. 1	Eq. 1	Eq. 2	Eq. 1	Eq. 2	Eq. 3
	(1) Absenteeism	(1a) Absenteeism	(1b) Enrollment	(2a) Absenteeism	(2b) Enrollment	(2c) Biomass fuel
Biomass fuel	0.091	0.184*	-0.432***	3.917***	-2.092*	
	(0.090)	(0.104)	(0.051)	(0.332)	(1.222)	
Age	-0.041	-0.237*	0.759***	-0.116	0.702***	-0.001
	(0.029)	(0.118)	(0.014)	(0.071)	(0.090)	(0.002)
Age squared	0.002	0.009*	-0.031***	0.005*	-0.029***	-0.000
	(0.001)	(0.005)	(0.001)	(0.003)	(0.003)	(0.000)
Gender	-0.004	-0.002	-0.011	0.024	-0.021	-0.007**
	(0.032)	(0.031)	(0.019)	(0.020)	(0.020)	(0.003)
Biological child	0.083	-0.008	0.330***	-0.137***	0.366***	0.036***
	(0.055)	(0.079)	(0.030)	(0.048)	(0.031)	(0.006)
Age of household head	-0.002	-0.002	0.001	-0.004***	0.002*	0.001***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
Gender of head	-0.059	-0.078	0.090***	-0.065*	0.094***	0.007
	(0.051)	(0.051)	(0.031)	(0.039)	(0.032)	(0.007)
Household Size	-0.033***	-0.032***	0.000	-0.018**	0.001	0.001
	(0.011)	(0.011)	(0.006)	(0.009)	(0.006)	(0.001)
Mother education	-0.091*	-0.180**	0.410***	-0.075	0.305***	-.043***
	(0.055)	(0.075)	(0.032)	(0.060)	(0.108)	(0.007)
Father education	-0.046	-0.144*	0.405***	-0.045	0.322***	-0.032***
	(0.047)	(0.072)	(0.027)	(0.054)	(0.091)	(0.005)
Job in agriculture	0.210***	0.141**	0.207***	0.214***	0.318***	0.075***
	(0.060)	(0.074)	(0.029)	(0.063)	(0.077)	(0.008)
Income, real net	-0.011	-0.009	-0.003	-0.005	-0.003	-0.000
	(0.008)	(0.008)	(0.005)	(0.007)	(0.005)	(0.001)
TLU	0.001	0.001	-0.003	-0.003	-0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)
Distance to center	0.000	0.001	-0.003***	-0.004***	-0.001	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.000)

(continued on next page…)

JARE Preprint

Table 3. Cont. from previous page

	Probit model	eprobit with sample selection			eprobit with sample selection and IV	
	Eq. 1	**Eq. 1**	**Eq. 2**		**Eq. 1**	**Eq. 1**
	(1) **Absenteeism**	**(1a)** **Absenteeism**	**(1b)** **Enrollment**		**(1)** **Absenteeism**	**(1a)** **Absenteeism**
Distance to school	0.001 (0.001)	0.008 (0.004)	-0.001 (0.001)	0.000 (0.001)	-0.001 (0.001)	0.000 (0.000)
Time trend	-0.130*** (0.027)	-0.132*** (0.026)	0.053*** (0.012)	-0.068*** (0.021)	0.049*** (0.012)	
Average school enrollment			4.960*** (0.366)		4.590*** (0.671)	
Forest share in the woreda						0.001*** (0.000)
Observations	16120	28537	28537	28537	28537	28537
Woreda FE	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap Wald Fstat	NA	NA	NA	24.63	NA	NA
Cragg-Donald Wald F stat.	NA	NA	NA	25.43	NA	NA
Corr(e.biomass,e.absent)	NA	NA	NA	-0.856**	NA	NA
Marginal effect	0.011	0.031	NA	0.201**	NA	NA

Notes: FE stands for fixed effects. Robust standard errors are clustered at the household level. Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level. The values reported are the regression coefficients. Column 1 refers to the results of equation (1) using the probit model estimations. Columns (1a) and (1b) denote the results of the extended probit regression model (ERM) controlling for sample selection bias (i.e., equations 1 and 2). Columns (2a-2c) are the results using ERM controlling for endogeneity and sample selection bias. The use of cooking fuel is instrumented by the share of forest cover in the woreda. The Kleibergen-Paap Wald *F*-stats refer to the test of instrument relevance, are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Staiger and stock (1997). The Cragg-Donald Wald *F- statistic* rejects weak- identification test.

during the 2011-2016 period. We find no significant absenteeism by gender[5], but find that children in larger households have a lower probability of being absent from school, possibly because chores (including fuelwood collection) are distributed across larger number of household members (Baiyegunhi and Hassan, 2014). We also find that children of parents who have attended formal education are less likely to be absent and more likely to be enrolled, in line with the literature (Beyene, Mekonnen, and Gebreegziabher, 2014; Gebru and Bezu, 2014). Children whose parents work in agriculture are more likely to be absent, which we attribute to a substitution effect between schooling and demand for on-farm labor, which is plausible given the labor-intensive nature of smallholder agriculture in Ethiopia[6]. Finally, we find that, in general, absenteeism is lower when the head of the household is older and the household is female-headed, although these results are not significant across all specifications.

[5] A result also found in Gebru and Bezu (2014) and O'Brien, Do, and Edelson (2021)

[6] Huisman and Smits (2009) argue that this substitution will be particularly evident in communities where schooling is not compulsory.

JARE Preprint

With regards to household welfare status, we note that the coefficients associated with these variables are generally insignificant. Although the coefficient on income is negative for school enrollment and absenteeism, it is insignificant. The evidence in the literature on the effect of household welfare status on absenteeism and enrollment is generally mixed[7]. In our context, the lack of statistical evidence could be explained by the fact that the cost of primary and secondary schooling in Ethiopia is low and thus income may not be the main constraint.

Table 4 shows the estimated eprobit results for health-related school absenteeism (columns 1 and 2) and work-related absenteeism (columns 3 and 4). The relationship between biomass fuel use and absenteeism due to sickness is inconclusive, with the coefficient on biomass fuel being statistically significant and positive when controlling for sample selection and endogeneity. On the other hand, the coefficient of biomass fuel in the work-related absenteeism is consistently positive and significant, suggesting a positive effect of using biomass fuel for cooking on the likelihood of work-related absenteeism. This finding is in line with other qualitative studies in the literature where a negative effect of fuelwood use on child schooling was reported (O'Brien, Do, and Edelson, 2021; Nankhuni and Findeis, 2004), either due to fatigue or due to the need for children to spend time collecting fuelwood. The marginal effect of using biomass fuel for cooking on absenteeism for work-related reasons is higher (between 7.9 and 21.5 percentage points) than the marginal effects for biomass fuel on absenteeism due to sickness.

With regards to the other covariates, we find significant and positive effect of gender on health-related absenteeism (column 2), but that the effect of gender on work-related absenteeism is insignificant and negative[8]. In line with expectations and other studies, being a biological child of the household head decreases the likelihood of absenteeism for work, compared to being adopted or a foster child (Fafchamps and Wahba, 2006). It could be driven by a preference by the head of the household to keep his/her own children at school, while putting a disproportionate burden of domestic chores on other children in the household. In addition to this, we also find that the relationship between the age of children and school absenteeism due to work-related reasons is non-linear, with the work falling predominantly on older children as they tend to be stronger and thus preferred for environmental chores that require carrying heavier loads of biomass fuel (Ndiritu and Nyangena, 2011) , and thus are more likely to be absent from school.

Finally, the results of the regressions for the sub-samples of boys and girls are shown in Table 5, where columns (1-2) show the result for girls and columns (3-4) are for boys. The estimated results suggest a gendered impact of using biomass fuel for cooking on school absenteeism. More importantly, the marginal effects of biomass fuel tend to be larger for boys, which is likely explained by the fact that work-related absenteeism is more important and that boys (especially older boys) are generally responsible for a higher share of the non-chore workload (e.g., helping on farm, etc…). While our results are still consistent with the general finding that girls are affected, we find that the effect is higher for boys, which is different from what is found in some other studies that report that girls are generally given more responsibilities related to household chores compared to boys who are expected to contribute to the household income (Huisman and Smits, 2009).

[7] We also note that other papers focusing on the effect of income on schooling outcomes tend to find mixed results. Cogneau and Jedwab (2007) find that a 10% rise in household income leads to a 0.023-0.030 increase in the likelihood to attend school for 5-17 year-old children in the Ivory coast. Chaudhury, Christiaensen, and Asadullah (2006), on the other hand, only find a very modest effect of income on enrollment (Chaudhury, Christiaensen, and Asadullah, 2006) and also Iddrisu, Danquah, and Quartey (2017) pointing that household welfare does not necessarily influence primary school enrollment in Ghana.

[8] A possible explanation for this is that there may be a gender bias in health-related expenditure as found in Orazem and King (2007) which could increase the probability of girls to fall sick with respect to boys.

IAAE Preprint

Table 4: Determinants of School absenteeism by reason

	Sick related reason		Work related reasons	
Variables	**eprobit with selection (1)**	**eprobit with selection+ IV (2)**	**eprobit with selection (3)**	**eprobit with selection+ IV (4)**
Biomass fuel	-0.147	3.555***	0.480***	3.758***
	(0.109)	(0.549)	(0.124)	(0.550)
Age	-0.028	-0.012	-0.355***	-0.222**
	(0.099)	(0.060)	(0.096)	(0.094)
Age squared	0.000	0.000	0.015***	0.010***
	(0.004)	(0.002)	(0.004)	(0.004)
Gender	0.066	0.063**	-0.042	-0.004
	(0.043)	(0.029)	(0.035)	(0.027)
Biological child	0.091	-0.076	-0.107	-0.190***
	(0.074)	(0.060)	(0.076)	(0.052)
Age of household head	-0.001	-0.003**	-0.002	-0.004***
	(0.002)	(0.001)	(0.002)	(0.001)
Gender of head	0.015	-0.015	-0.125**	-0.101**
	(0.064)	(0.046)	(0.056)	(0.047)
Household Size	-0.055***	-0.033**	-0.006	-0.005
	(0.013)	(0.013)	(0.012)	(0.009)
Mother education	-0.012	-0.153***	-0.286***	-0.031
	(0.073)	(0.053)	(0.074)	(0.091)
Father education	-0.028	-0.098*	-0.211***	-0.029
	(0.073)	(0.050)	(0.065)	(0.076)
Job in agriculture	0.197***	-0.160*	0.080	-0.211***
	(0.073)	(0.088)	(0.081)	(0.079)
Income, real net	-0.014	-0.009	-0.002	-0.001
	(0.009)	(0.008)	(0.009)	(0.007)
TLU	-0.008	-0.009	0.003	-0.002
	(0.011)	(0.007)	(0.002)	(0.002)

(continued on next page…)

JARE Preprint

Table 3. Continued from previous page

	Sick related reason		Work related reasons	
Variables	**eprobit with selection (1)**	**eprobit with selection+ IV (2)**	**eprobit with selection (3)**	**eprobit with selection+ IV (4)**
Distance to center	-0.000	-0.005***	0.002**	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)
Distance to school	0.001	0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.002)	(0.001)
Time trend	-0.137***	-0.079***	-0.101***	-0.064***
	(0.031)	(0.028)	(0.029)	(0.024)
Observations	28537	28537	28537	28537
Woreda FE	Yes	Yes	Yes	Yes
Kleibergen-Paap Wald *F*-stat	NA	21.992	NA	21.992
Cragg-Donald Wald *F*-stat.	NA	25.43	NA	25.43
Corr(e.biomass,e.absent)		-0.815***		-0.773***
Marginal effect	-0.0086	0.0963**	0.079**	0.215**

Notes: FE stands for fixed effects. Robust standard errors are clustered at the household level. Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level.The values reported are the regression coefficients. The outcome variable in columns (1) and (2) is absenteeism due to sickness and is estimated using eprobit controlling for selection bias (column 1) and eprobit with selection bias and endogeneity (column 2), respectively.Columns (3) and (4) apply the same approach as in (1) and (2) using absenteeism due to work-related reasons as the dependent variable.

4.2 Robustness checks

To test the robustness of the results to model specification and sample selection, we run three robustness checks. First, we apply a linear probability model (LPM with the IV) instead of the eprobit model and test whether the inclusion of household fixed effects alter our main conclusions. Table S1 in the online supplementary material shows that the marginal effect of biomass cooking fuel on absenteeism is positive and significant at least at the 10% level in all cases when we control for endogeneity (column 1), include household-level fixed effects (column 2) or individual-level fixed effects (column 3), although the magnitude of the effects are quite different depending on the estimated model. Overall, the marginal effects of the LPM are comparable to the eprobit results.

Second, we test whether the results could be driven by the inclusion of urban households in the sample. It could be argued that since the sample includes households living in rural and urban areas, the characteristics of households in urban areas[9], their preferences and the availability of alternative fuels could be biasing our results. In Table S2, we run the eprobit model with IV and different estimations of the LPM using the rural sample only and we find that the results are unchanged even when we control for household and individual level fixed effects. Third, given

[9] The urban sample constitutes 20% of the total sample with 73% of the children are enrolled in school. Absenteeism rate is 4%. Regarding the use of biomass cooking fuel, 33.6% of household use clean fuel for cooking and 66.4% are using biomass cooking fuel.

JAREprint

Table 5: Determinants of School Absenteeism by Gender

	Girls Absenteeism		Boys Absenteeism	
Variables	(1)	(2)	(3)	(4)
Biomass fuel	0.143	3.517***	0.229*	4.349***
	(0.136)	(0.539)	(0.135)	(0.325)
Age	-0.198	-0.116	-0.257**	-0.112
	(0.175)	(0.111)	(0.127)	(0.074)
Age squared	0.008	0.005	0.010**	0.005*
	(0.007)	(0.005)	(0.005)	(0.003)
Biological child	-0.084	-0.239***	0.077	-0.028
	(0.104)	(0.070)	(0.099)	(0.054)
Age of household head	-0.004*	-0.005***	-0.001	-0.003**
	(0.002)	(0.002)	(0.002)	(0.001)
Gender of head	-0.066	-0.071	-0.090	-0.055
	(0.066)	(0.049)	(0.066)	(0.047)
Household Size	-0.019	-0.010	-0.042***	-0.021*
	(0.014)	(0.011)	(0.014)	(0.011)
Mother education	-0.188*	0.045	-0.185**	0.095
	(0.101)	(0.087)	(0.089)	(0.067)
Father education	-0.067	0.065	-0.198**	0.042
	(0.101)	(0.071)	(0.082)	(0.062)
Job in agriculture	0.093	-0.182**	0.211**	-0.271***
	(0.095)	(0.082)	(0.097)	(0.082)
Income, real net	-0.008	-0.004	-0.011	-0.004
	(0.011)	(0.009)	(0.010)	(0.008)
TLU	0.002	-0.002	0.001	-0.005
	(0.003)	(0.002)	(0.002)	(0.003)
Distance to center	0.002	-0.005***	0.001	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Distance to primary school	0.003*	0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Time trend	-0.173***	-0.102***	-0.107***	-0.047**
	(0.035)	(0.036)	(0.032)	(0.021)
Observations	14087	14087	14450	14450
Woreda FE	Yes	Yes	Yes	Yes
Kleibergen-Paap Wald *F*-stat	NA	15.27	NA	16.63
Cragg-Donald Wald *F*-Stat.	NA	10.79	NA	11.96
Corr (e.biomass, e.absent)	0.023	-0.805	0.042	-0.895**
Marginal effect		0.187**		0.213**

Notes: FE stands for fixed effects. Robust standard errors are clustered at the household level. Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level. Columns (1)-(2) denote the results for the sub-sample of females whereas columns (3)-(4) denote the results for the sub-sample of males. Columns (2) and (4) denote the results from the ERM controlling for endogeneity and sample selection using the share of forest cover in the woreda as an instrument.

JARE Preprint

that many clean fuel users are in Addis Ababa, this group may be driving our results. To address this, Table S3 replicates the models from Table 3, excluding households from Addis Ababa. The results remain consistent, controlling for selection bias and endogeneity, and confirm a negative and significant marginal effect of biomass fuel on school absenteeism.

4.3 Discussion

We find that the use of biomass fuel for cooking is associated with an increase in the predicted probability of school absenteeism using different model specifications and controlling for sample selection bias and endogeneity. The results shed light on the importance of the type of primary cooking fuel and its effects on schooling absenteeism and indicate that the impact of cooking fuel on absenteeism is primarily driven by work reasons. The analysis carried out for sub-samples of boys and girls reveals that both genders are affected by the use of biomass fuel, however, the effects seem to be more pronounced for boys, especially for older ones. According to the results, boys who are older than 12.8 years old get more work responsibilities, which affects their school attendance.

Our overall finding of the link between cooking fuel and school absenteeism is consistent with the findings in O'Brien, Do, and Edelson (2021) and Kelly et al. (2018), but our paper provides new evidence on the link between fuel use and education by identifying the main channels through which biomass fuel use affects school absenteeism and highlighting differential impacts by gender.

However, it is important to also highlight the limitations of our study. First, it should be noted that our estimates resulting from the empirical models are likely to be underestimations of the actual effect of biomass fuel use on school absenteeism. This is due to the survey design and the way in which the question on school absenteeism is framed. The question to which households responded, addressed if the child has been absent for more than a week in the past month only. The time frame of the last month as well as the reference to absenteeism of at least a week likely result in an underreporting of the incidence of school absenteeism. This underreporting might be correlated with firewood collection, which is typically not done for long time periods but recurrently, and this could therefore bias our econometric results downward. Second, we are not able to reveal exactly what kind of work explains absenteeism and this is not trivial from a policy perspective. If, for example, most of the absenteeism is explained by on-farm work, then the nature of the policies to best tackle this issue would probably be seasonal. This would be completely different if the absenteeism was driven by work in other, less seasonal, sectors. Third, given the limitations of the dataset, we are not able to test the effects on actual measures of learning achievements, such as test scores or repetition rates. While absenteeism is an important outcome, it would be crucial to understand the linkages between dirty fuels and more long-term learning outcomes.

JARE Preprint

5. Conclusion

This paper presents empirical evidence on the effects of cooking fuel choice on school absenteeism, based on an extended probit regression model (eprobit) and an instrumental variable estimation using panel household survey data from Ethiopia for the period 2011-2016. To better understand which channel of effect, work or sickness, is most important and which children are most affected by the cooking fuel choice, we investigate the heterogeneity of impact by reason of absenteeism and by gender. Overall, we find that the use of polluting cooking fuel significantly increases the likelihood of school absenteeism among children. Contrary to initial expectations, this effect is mainly driven through absenteeism for work rather than through absenteeism for sickness and the adverse health effects of using dirty fuel. We find the strongest and most consistent adverse effects of cooking fuel choice for boys, who are stronger and might therefore be more likely to miss school for environmental chores such as firewood collection.

We need to note that these findings are specific to Ethiopia, and that spillover effects of biomass cooking fuel might vary in other settings. Yet, the paper demonstrates the existence of important spillovers of adopting modern cooking fuels, with a focus on educational outcomes, an issue that remains important for socio-economic and rural development. Our results suggest that, beyond the well-researched health effects of improving access to cleaner cooking fuels, government programs and policies focusing on cleaner cooking fuels could have positive implications for schooling outcomes. Policies that can maximize these energy-education synergies are most relevant.

The results entail three relevant policy implications. First, the finding that work-related absenteeism is driving the link between using biomass fuel for cooking and school absenteeism and that older children are more likely to be affected, has important implications for policy targeting. One possible policy implication relates to the targeting of interventions and a stronger focus on households with older children when targeting either fuel-related interventions or labour-saving interventions, as these could yield larger effects on educational outcomes. A second potential policy implication would be to focus on policies that increase the cost of school absenteeism, such as, for instance, cash transfers or vouchers that are conditional on school attendance. Third, our paper highlights the potential intergenerational implications of poor schooling outcomes. Specifically, one consistent finding throughout our paper is that parental education, especially mothers' education is a much stronger predictor of absenteeism than several other variables (e.g., income), especially in the case of work-related absenteeism where parents' education is more likely to matter. This implies not only that impacts of absenteeism could extend to future generations (since today's children are tomorrow's parents) but also that interventions that improve the education of parents (e.g., adult literacy campaigns) could be an effective way to reduce absenteeism.

Beyond the policy implications, the results point to a potentially important inter-sectoral spillover effect between energy, education and natural resource degradation. Specifically, it highlights that beyond the well-researched health and environmental benefits of a transition towards clean and sustainable cooking fuels, such a transition could also have spillover effects in terms of education outcomes and that these are quantitatively large. Together, this highlights that access to clean cooking fuels has the potential to contribute to other sustainable development goals (SDGs), including education, health and gender equality.

Finally, the paper identifies several potential avenues for future research. First, given the limitations related to our data, we are unable to quantitatively test what types of programmes are most effective at reducing absenteeism, which is of critical importance for Governments seeking to address the issue. Second, while our results seem to suggest that absenteeism is primarily driven by work-related reasons, it would be interesting to estimate the impacts on other measures of learning achievement, such as test scores or repetition rates. This is important as literature has shown that both exposure to smoke and fatigue can impair learning. It might well be that the largest effect of biomass cooking fuels may manifest itself more strongly through learning outcomes, rather than absenteeism. This is, unfortunately, something we are not able to test with the current data and something that remains a potentially fruitful avenue for future research. Being able to get more clarity on these issues is critical to enable Governments to design more effective interventions as well as make the most of the synergies that exist between the different SDGs.

[First submitted November 2023; accepted for publication October 2024.]

JARE Preprint

References

Angrist, J. D. 2001. "Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice". *Journal of Business & Economic Statistics*, 19(1), 2-28. doi: 10.1198/07350010152472571.

Assaad, R., Levison, D., and Zibani, N. 2010. "The effect of domestic work on girls' schooling: Evidence from Egypt". *Feminist Economics*, 16(1), 79-128. doi:10.1080/13545700903382729.

Benti, N. E., Gurmesa, G. S., Argaw, T., Aneseyee, A. B., Gunta, S., Kassahun, G. B., Aga, G.S and Asfaw, A. A. 2021. "The current status, challenges and prospects of using biomass energy in Ethiopia". *Biotechnology for Biofuels,* 14(1), 1-24. doi:10.1186/s13068-021-02060-3.

Beyene, A. D., Mekonnen, A., and Gebreegziabher, Z. 2014. "Natural resource collection and children's literacy: empirical evidence from panel data in rural Ethiopia". *Environment for Development* Discussion Paper, (14-18). Available online at: https://media.rff.org/documents/EfD-DP-14-18.pdf

Biswas, S., and Das, U. 2022. "Adding fuel to human capital: Exploring the educational effects of cooking fuel choice from rural India". *Energy Economics*, 105, 105744. doi:10.1016/j.eneco.2021.105744.

Bruce, N. G., Dherani, M. K., Das, J. K., Balakrishnan, K., Adair-Rohani, H., Bhutta, Z. A., and Pope, D. 2013. "Control of household air pollution for child survival: estimates for intervention impacts". *BMC Public Health*, 13(3), 1-13. doi: 10.1186/1471-2458-13-S3-S8.

Central Statistical Agency of Ethiopia (CSA). Ethiopia Socioeconomic Survey,Wave 3 (ESS3) 2015-2016. Public Use Dataset. Ref: ETH_2015_ESS_v03_M. Downloaded on March 2020.

Central Statistics Agency of Ethiopia (CSA). Ethiopia Socioeconomic Survey. Wave 2 (ESS2) 2013-2014. Public Use Dataset. Ref. ETH_2013_ESS_v02_M. Dataset downloaded on March 2020.

Central Statistics Agency of Ethiopia (CSA). Ethiopia Socioeconomic Survey. Wave 1 (ESS1) 2011-2012. Public Use Dataset. Ref. ETH_2011_ESS_v01_M. Dataset downloaded on March 2020.

Chaudhury, N.,Christiaensen, L and Asadullah, M.N. 2006. "Schools, Household, Risk, and Gender: Determinants of Child Schooling in Ethiopia". CSAE Working Paper Series 2006-06, Centre for the Study of African Economies, University of Oxford. Available online at: https://ideas.repec.org/p/csa/wpaper/2006-06.html

Choudhuri, P., and Desai, S. 2021. "Lack of access to clean fuel and piped water and children's educational outcomes in rural India". *World Development,* 145, 105535. doi: 10.1016/j.worlddev.2021.105535.

Cogneau, D., and Jedwab, R. 2007. "Household income and investments in child health and education in Ivory Coast". Mimeo: Paris School of Economics. Available online: https://www.openbase.in.th/files/seminar_jan14_Jedwab.pdf

Daidone, S., and Fontes, F. P. 2023. "The role of social protection in mitigating the effects of rainfall shocks. Evidence from Ethiopia". *Journal of Productivity Analysis,* 60(3), 315-332. doi: 10.1007/s11123-023-00688-x.

DeGraff, D. S., Levison, D., and Dungumaro, E. W. 2017. "Environmental Chores, Household Time Use, and Gender in Rural Tanzania". In Connellly,R. and Kongar, E., ed. *Gender and Time Use in a Global Context*. New York: Palgrave Macmillan, pp. 407-434. doi: 10.1057/978-1-137-56837-3

Di Falco, S., Veronesi, M., and Yesuf, M. 2011. "Does adaptation to climate change provide food security? A micro-perspective from Ethiopia". *American Journal of Agricultural Economics,* 93(3), 829-846.doi: 10.1093/ajae/aar006.

Energy Sector Management Assistance Program (ESMAP). 2020. The State of Access to Modern Energy Cooking Services. Washington, DC: World Bank. Available online at: http://documents.worldbank.org/curated/en/937141600195758792/The-State-of-Access-to-Modern-Energy-Cooking-Services. [Accessed January 2024]

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), 1057-1068. doi:10.1289/ehp.021101057.

Fafchamps, M., and Wahba, J. 2006. "Child labor, urban proximity, and household composition". *Journal of Development Economics*, 79, 374-397. doi: 10.1016/j.jdeveco.2006.01.005.

Federal Ministry of Education. 2021. "Education Sector Development Programme VI (ESDP VI) 2020/21-2024/25". Addis Ababa: The Federal Democratic Republic of Ethiopia.

Fontes, F. P. 2020. "Soil and Water Conservation technology adoption and labour allocation: Evidence from Ethiopia". *World Development*, 127, 104754. doi:10.1016/j.worlddev.2019.104754.

Food and Agriculture Organization. 2018. "Guidelines on methods for estimating livestock production and productivity. Rome, FAO.

Gebru, B., and Bezu, S. 2014. "Environmental resource collection: implications for children's schooling in Tigray, northern Ethiopia". *Environment and Development Economics*, 19(2), 182-200. doi: 10.1017/S1355770X13000454.

Glewwe, P., and Miguel, E. A. 2007. "The impact of child health and nutrition on education in less developed countries". *Handbook of Development Economics*, chapter 4, 3561-3606. doi: 10.1016/S1573-4471(07)04056-9

Glick, P. and Sahn, D. 2000. "Schooling of Girls and Boys in a West African Country: The Effects of Parental Education, Income, and Household Structure". *Economics of Education Review*, 19 (1), 63–87. doi: 10.1016/S0272-7757(99)00029-1.

Guta, D. D. 2014. "Effect of fuelwood scarcity and socio-economic factors on household bio-based energy use and energy substitution in rural Ethiopia". *Energy Policy*, *75*, 217-227. doi: 10.1016/j.enpol.2014.09.017.

Haile, G., and Haile, B. 2012. "Child labour and child schooling in rural Ethiopia: nature and trade-off. *Education Economics"*, 20(4), 365-385. doi:10.1080/09645292.2011.623376.

Huisman, J., and Smits, J. 2009. "Effects of household-and district-level factors on primary school enrollment in 30 developing countries". *World Development*, 37(1), 179-193.doi: 10.1016/j.worlddev.2008.01.007.

International Energy Agency. 2019. "Africa Energy Outlook: special Report". IEA, Paris

International Energy Agency. 2022. "Africa Energy Outlook 2022". World Energy Outlook Special Report. IEA, Paris.

Iddrisu, A. M., Danquah, M., and Quartey, P. 2017. "Analysis of school enrollment in Ghana: A sequential approach". *Review of Development Economics*, 21(4), 1158-1177. doi: 10.1111/rode.12302.

Kelly, C. A., Crampin, A. C., Mortimer, K., Dube, A., Malava, J., Johnston, D., Unterhalter, E. and Glynn, J. R. 2018. "From kitchen to classroom: Assessing the impact of cleaner burning biomass-fuelled cookstoves on primary school attendance in Karonga district, northern Malawi". *PloS one*, 13(4).doi: 10.1371/journal.pone.0193376

Köhlin, G., O. Sills, E., K. Pattanayak, S., and Wilfong, C. 2011. "Energy, Gender, and Development: What Are the Linkages? Where Is the Evidence?". *Social Development Working Paper 125*. Washington, DC: World Bank. Available online at https://documents1.worldbank.org/curated/en/430841468156268921/pdf/644100WP0Energ000public00BOX361536B.pdf

Kurata, M., Takahashi, K., and Hibiki, A. 2020. "Gender differences in associations of household and ambient air pollution with child health: evidence from household and satellite-based data in Bangladesh". *World Development*, 128, 104779. doi: 10.1016/j.worlddev.2019.104779

Lalive, R. and Cattaneo, M. A. 2009. "Social interactions and schooling decisions". *The Review of Economics and Statistics*, 91(3), 457-477. doi: 10.1162/rest.91.3.457

Levison, D., DeGraff, D. S., and Dungumaro, E. W. 2018. "Implications of environmental chores for schooling: Children's time fetching water and firewood in Tanzania". *The European Journal of Development Research*, 30(2), 217-234. doi:10.1057/s41287-017-0079-2.

Maertens, M., and Verhofstadt, E. 2013. "Horticultural exports, female wage employment and primary school enrolment: Theory and evidence from Senegal". *Food Policy*, 43, 118-131.doi: 10.1016/j.foodpol.2013.07.006.

Mani, S., Hoddinott, J., and Strauss, J. 2013. "Determinants of schooling: Empirical evidence from rural Ethiopia". *Journal of African Economies*, 22(5), 693-731. doi: 10.1093/jae/ejt007.

Maji, P., Mehrabi, Z., and Kandlikar, M. 2021. "Incomplete transitions to clean household energy reinforce gender inequality by lowering women's respiratory health and household labour productivity". *World Development*, 139, 105309. doi:10.1016/j.worlddev.2020.105309.

Marchenko, Y. V., and Genton, M. G. 2012. "A Heckman selection-*t* model". *Journal of the American Statistical Association*, 107(497), 304-317. doi: 10.1080/01621459.2012.656011

Mekonnen, A. and Köhlin, G. 2009. "Determinants of Household Fuel Choice in Major Cities in Ethiopia", Environment for Development Discussion paper series no. 08-18. Available online at https://www.efdinitiative.org/sites/default/files/efd-dp-08-18.pdf

Mekonnen, A., and Köhlin, G. 2008. "Biomass Fuel Consumption and Dung Use as Manure: Evidence from Rural Households in the Amhara Region". Environment for Development Discussion Paper series no. 08-17. Available online at https://media.rff.org/documents/EfD-DP-08-17.pdf

Ministry of Water, Irrigation and Energy. 2019. "National Electrification Program 2.0. Integrated Planning for Universal Access". Addis Ababa, Ethiopia.

Nankhuni, F. J., and Findeis, J. L. 2004. "Natural resource-collection work and children's schooling in Malawi". *Agricultural Economics*, 31(2-3), 123-134. doi:10.1016/j.agecon.2004.09.022.

Ndiritu, S. W., and Nyangena, W. 2011. "Environmental goods collection and children's schooling: Evidence from Kenya". *Regional Environmental Change*, 11(3), 531-542. doi:10.1007/s10113-010-0168-7.

O'Brien, J., Do, P., and Edelson, M. 2021. "The effects of fuelwood on children's schooling in rural Vietnam". *Journal of Asian Econo*mics, 101266. doi:10.1016/j.asieco.2020.101266.

Orazem, P. and King, E. M. 2007. "Schooling in developing countries: The roles of supply, demand, and government policy", in T. Schultz and J. Strauss, eds. *Handbook of Development Economics*. North Holland, pp. 3475–559. doi:10.1016/S1573-4471(07)04055-7.

Owili, P. O., Muga, M. A., Pan, W. C., and Kuo, H. W. 2017. "Cooking fuel and risk of under-five mortality in 23 Sub-Saharan African countries: a population-based study". *International Journal of Environmental Health*, 27(3), 191-204. doi:10.1080/09603123.2017.1332347.

Palmer, C., and Macgregor, J. 2009. "Fuelwood scarcity, energy substitution, and rural livelihoods in Namibia". *Environment and Development Economics*, 14, 693–715. doi:10.1017/S1355770X08005007

Patel, S. K., Patel, S., and Kumar, A. 2019. "Effects of cooking fuel sources on the respiratory health of children: evidence from the Annual Health Survey, Uttar Pradesh, India". *Public Health*, 169, 59-68. doi:10.1016/j.puhe.2019.01.003.

Porter, G., Hampshire, K., Abane, A., Munthali, A., Robson, E., Mashiri, M., Tanle. A., Maponya, G., and Dube, S. 2012. "Child porterage and Africa's transport gap: Evidence from Ghana, Malawi and South Africa". *World Development*, 40(10), 2136-2154. doi:10.1016/j.worlddev.2012.05.004.

Scharlemann, J. P., Brock, R. C., Balfour, N., Brown, C., Burgess, N. D., Guth, M. K., Ingram, D.J., Lane,R.,Martin,J.G.C,Wicander,S. and Kapos, V. 2020. "Towards understanding interactions between Sustainable Development Goals: the role of environment–human linkages". *Sustainability Science*, 15(6), 1573-1584. doi:10.1007/s11625-020-00799-6.

Smith, K. R., McCracken, J. P., Weber, M. W., Hubbard, A., Jenny, A., Thompson, L. M., Balmes, J.,Diaz, A., Arana, B. and Bruce, N. 2011. "Effect of reduction in household air pollution on childhood pneumonia in Guatemala (RESPIRE): a randomised controlled trial". *The Lancet*, 378(9804), 1717-1726. doi:10.1016/S0140-6736(11)60921-5.

Staiger, D., and Stock, James, H. 1997. "Instrumental variables regression with weak instruments". *Econometrica*, 65(3), 557-586. doi: 10.2307/2171753.

United Nations (2022). United Nations Department of Economic and Social Affairs, Population Division: Revision of World Population Prospects (2022).

Wolde-Ghiorgis, W. 2002. "Renewable energy for rural development in Ethiopia: the case for new energy policies and institutional reform". *Energy Policy*, 30(11-12), 1095-1105. doi:10.1016/S0301-4215(02)00061-7.

Wooldridge, J. M. 2014. "Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables". *Journal of Econometrics*, 182(1), 226-234. doi:10.1016/j.jeconom.2014.04.020.

World Bank. 2021. World Bank Development Indicators (WDI) databank. Available online at https://databank.worldbank.org/source/world-development-indicators [Accessed January 2024]

JARE Preprint

Appendices

Table A1: Summary Statistics of Households by Survey Year

	2011		2013		2015	
Variables	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**
Cooking with Biomass fuel	0.974	0.159	0.858	0.349	0.861	0.346
Average school enrollment	0.255	0.055	0.262	0.054	0.252	0.049
Age of child (years)	10.950	3.470	11.727	3.467	11.602	3.330
Gender of child (1=female)	0.499	0.363	0.516	0.369	0.499	0.371
Biological child (binary)	0.799	0.376	0.742	0.402	0.772	0.384
Age of household head (years)	44.775	14.472	45.192	14.598	46.447	14.048
Gender of head (female=1)	0.234	0.423	0.286	0.452	0.275	0.447
Household Size	5.563	2.123	5.236	2.164	5.431	2.127
Mother education (binary)	0.178	0.369	0.267	0.422	0.306	0.442
Father education (binary)	0.329	0.453	0.402	0.464	0.446	0.475
Job in Agriculture (binary)	0.795	0.365	0.715	0.406	0.738	0.413
Income, real net (1000 ETB)	64.612	200.652	83.891	352.231	117.857	1119.86
TLU	2.278	3.231	2.066	3.076	2.065	7.716
Forest share in the woreda (%)	3.984	12.944	3.268	11.418	3.104	11.070
Distance to nearest center (KM)	40.047	32.177	33.641	33.294	33.721	32.824
Distance to primary school (KM)	0.950	4.416	0.813	4.263	0.791	3.183
Observations	2938		4007		3854	

Notes: The sample of households is balanced on covariates. Households which include school aged children regardless of their enrollment status are 2141 in 2011, 3083 in 2013 and 3082 in 2015. The mean values are the average values over all households. For individual characteristics (e.g., age and gender of child), we first took the mean value within a household for a given wave, and then calculated the mean across households.

JARE Preprint

Table A2: Falsification Test of the Instrument

	(1)	(2) Non-biomass users	(3)
Forest share	0.001***	0.000	
	(0.000)	(0.001)	
Age	-0.000	0.001	-0.009*
	(0.002)	(0.007)	(0.005)
Age squared	-0.000	-0.000	0.001**
	(0.000)	(0.000)	(0.000)
Gender	-0.007**	0.008	0.006
	(0.003)	(0.009)	(0.006)
Biological child	0.036***	0.019	-0.028*
	(0.006)	(0.012)	(0.016)
Age of household head	0.001***	-0.000	-0.001**
	(0.000)	(0.000)	(0.000)
Gender of head	0.008	0.011	-0.019**
	(0.007)	(0.011)	(0.010)
Household Size	0.001	-0.003	-0.004**
	(0.001)	(0.002)	(0.002)
Mother education	-0.042***	-0.006	0.016
	(0.007)	(0.013)	(0.012)
Father education	-0.030***	-0.012	0.015
	(0.005)	(0.014)	(0.011)
Job in agriculture	0.076***	0.018	-0.072**
	(0.008)	(0.014)	(0.035)
Income, real net	-0.000	0.002	-0.001
	(0.001)	(0.002)	(0.002)
TLU	0.001	0.002	-0.001
	(0.001)	(0.003)	(0.001)
Distance to center	0.001***	0.000	-0.001**
	(0.000)	(0.001)	(0.000)
Distance to school	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Time trend	-0.014***	-0.012	-0.004
	(0.002)	(0.008)	(0.006)
Biomass fuel			0.749***
			(0.271)
Observations	28,592	1,886	16,564
Woreda FE	Yes	Yes	Yes

Notes: Robust standard errors are clustered at the household level. Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level. Column (1) is the first stage result (i.e., biomass fuel is the dependent variable) and column (2) is the falsification test using the sample of non-biomass fuel users and absenteeism as the dependent variable. Column (3) is the second stage regression of the IV approach using the share of forest cover in the woreda as an IV and absenteeism as the dependent variable. While a falsification test does no prove beyond doubt the validity of an instrument, we would expect a valid instrument to be highly relevant (i.e., significant in columns (1)) but have no effect on the non-adopters (i.e., insignificant in columns (2)).

JARE Preprint