Clean Cooking Fuel, Women's Intrahousehold Status, and Son Preference in Rural India

Avinash Kishore* and Dean Spears†

May 15, 2012

Abstract

Despite profoundly negative health consequences of indoor air pollution, most rural Indian households cook using traditional biomass fuel, rather than cleaner cooking fuel. Although many factors contribute to households' continued use of solid fuels, this paper focuses on one: women's intrahousehold status. We use two nationally representative datasets, and implement two complementary empirical strategies. The first strategy demonstrates that observable indicators of low women's status are associated with lower use of clean fuel, despite a broad range of controls. No similar association is found between status and electrification. The second strategy exploits Indian son preference: having a girl first child lowers women's status relative to having a boy first child, and is therefore associated with a three-fourths of a percentage point reduction in the likelihood of using clean fuel. This effect is found throughout the wealth distribution, and is not concentrated among households in states with a high child sex ratio or households where women have some education. Using several other assets as dependent variables – including electrification – no similar effect of having a girl first child was found. To our knowledge this is the first paper applying a causal identification strategy to this implication of women's status.

1 Introduction

Despite profoundly negative health consequences of indoor air pollution, about half of the households in the world cook using solid biomass fuels (Smith, 2002). This situation is even worse in rural India, where nearly

*(Corresponding Author): Harvard Kennedy School: avinash.kishore@gmail.com
†Economics Department, Princeton University

85 percent of households use firewood or dung cakes as the primary source of energy for cooking (NSSO, 2007). Burning these unprocessed biomass fuels in traditional open fire burners (called *"chulhas"*) results in very high levels of indoor air pollution and an estimated 450,000 to 550,000 premature deaths and nearly 500 million cases of illness each year (Smith, 2000). Such indoor air pollution is behind only malnourishment and unsafe drinking water as the major cause of disease and death in India. Women and children suffer disproportionately (ESMAP, 2003) as they spend more time indoors and women do essentially all of the cooking.

A switch to cleaner cooking fuels such as kerosene, liquified petroleum gas (LPG), or biogas would save many lives and reduce suffering from indoor air pollution. However, despite increased availability of cleaner fuel, households' transition from traditional cooking fuel has been slow: according to India's National Sample Survey Organization, LPG use as the primary cooking fuel increased from 1.9 percent of households in 1993-94 to only 8.6 percent in 2004-5; 75 percent of households primarily use firewood for cooking, and 9 percent use dung cakes (NSSO, 1997; NSSO, 2007). Why has the transition to clean fuel use been so slow?

Certainly there are many important factors relevant to households' use of biomass fuels, including price, accessibility, and a low opportunity cost of time spent collecting wood. This paper focuses on the role of one such factor: women's intrahousehold status. Economists have long recognized that households in developing countries are not unified economic actors (*e.g.* Udry, 1996) . In particular, "women and children are among the most deprived in the usual way an Indian household is run" (Bardhan, 2011). Yet it is woman cooks (and their children) who suffer most from indoor air pollution and have the most to gain from a switch to cleaner fuel. Therefore, the central hypothesis of this paper is that higher status of women within a rural Indian household would increase the chances of using clean cooking fuel, the benefits of which would accrue disproportionately to them.

Our paper is not the first to consider the role of women in households' fuel choice. However, to our knowledge this is the first paper to focus on the role of women's status, applying a strategy of causal identification. Hoddinott and Haddad (1995) find a positive relationship between a wife's share of income and the budget share of fuel in household data from Ivory Coast, although the parameter estimate was not statistically significant. In her study of fuel choice in urban Bolivia, Israel (2002) finds female earned income to be associated with a lower probability of firewood use. There are two alternative explanations for the effect she finds: women's opportunity cost of time and their bargaining power in the family. However, she cannot rule out that this is because women with higher earnings have a higher opportunity cost of time, rather than more bargaining power. Duflo et al. (2008) find suggestive evidence for our conjecture in a survey of 2,200 households in rural Orissa, where "households in which women may be more empowered – by virtue of being

members of a savings group – are 2 to 3 percent more likely to use a clean stove..." (73).

This paper uses two complementary empirical strategies. The first, documenting the importance of women's status as a mechanism, studies the association between clean cooking fuel use and indicators of women's intrahousehold status, controlling for observed economic and social factors. The second, more narrowly tailored to verify a causal effect, exploits Indian son preference and the negative impact on women's status of giving birth to a girl rather than a boy; it demonstrates that households where the first-born child is a girl are three-fourths of a percentage point less likely to use clean cooking fuel than households where the first born child is a boy.

A similar empirical strategy was recently used byPham-Kanter (2010), who finds that in the U.S., women who have a first-born son weigh more during the child's teenaged years than women who have a first-born daughter. Remarkably, she, too, attributes this effect to son preference, apparently active in the U.S.(Dahl and Moretti, 2008), although importantly less strongly and openly than in India. Pham-Kanter proposes that having a son rather than a daughter increases U.S. women's intrahousehold bargaining power, allowing them to devote less attention to maintaining low weight.

Section 2 describes the two data sets used in the analysis. Section 3 implements the first empirical strategy, focusing on indicators of women's status. Section 4 implements the second strategy, exploring this implication of son preference. Section 5 concludes.

2 Data

This paper uses the rural sub-samples of two data sets: the India Human Development Survey (IHDS) and the third wave of the National Family Health Survey (NFHS-3), India's version of the Demographic and Health Survey. Both data sets are nationally representative; they were collected only about a year apart. There are two advantages to using these two data sets simultaneously. First, the main results are all roughly replicated to the extent that differing survey designs allow, reducing the likelihood that apparent results are due to sampling error. Second, the data sets have complementary strengths that allow for the two empirical strategies in this paper.

The IHDS surveyed of 41,554 households in 1503 villages and 971 urban neighborhoods across India. IHDS interviews were conducted from late 2004 to late 2005[1] The IHDS was organized in part to resample house-

[1]"IHDS was jointly organized by researchers from the University of Maryland and the National Council of Applied Economic Research (NCAER), New Delhi. Funding for the survey was provided by the National Institutes of Health, grants R01HD041455 and R01HD046166" (Desai et al., 2007). The data and more information are available online at: ihds.umd.edu.

holds from NCAER's 1993-94 Human Development Profile of India (HDPI) survey; this panel structure is used in section 3. The IHDS included a broad range of economic questions (allowing computation, for example, of consumption per capita) and asked a woman in each household a particularly detailed set of questions about intrahousehold gender relations. These permit the first empirical strategy, regressing fuel choice on indicators of women's status.

The NFHS surveyed 109,041 households, including 124,385 women aged 15 to 49, from November 2005 to August 2006.[2]The NFHS, like other DHS surveys, does not include particularly detailed economic questions, but does ask about a range of household assets. The strength of the NFHS is in its detailed demographic questions and large sample size, which permit the second empirical strategy, regressing fuel choice on the sex of the first-born child, exploiting the implications of Indian son preference for women's status.

Table 1 presents summary statistics from these two data sets. Throughout this paper, observations are at the household level, and only rural households are included. These rural households are poor: the median household in the IHDS consumes only 548 Rupees ($55 at PPP) per capita each month. Broadly, the two data sets offer similar statistics. Unfortunately, the largest difference is in clean fuel use: while the NFHS asks about *cooking* fuel specifically, the IHDS asks a single question that combines cooking and other fuel use. Despite this difference in survey operationalization, the rest of this paper finds results using both surveys.

3 Strategy 1: Indicators of Women's Status

The first empirical strategy focuses on indicators of women's status, especially education, but also other indicators of intrahousehold relations included in the IHDS. Are indicators of women's higher intrahousehold status associated with more clean fuel use, controlling for a wide range of social, economic, and geographic covariates?

Gender relations are not easy to quantify and operationalize, especially in a national survey conducted in many states and many languages. In addition to education, we focus on two measures of possibly objective behavior that might be measured with less noise than others. One indicator is husbands beating wives. The survey asks "I would now like to ask you some questions about your community, not your own family. In your community is it usual for husbands to beat their wives in each of the following situations?", and asks about five situations.[3] We use an indicator constructed to be 1 if the woman answers yes to any of these

[2]The survey was implemented by the International Institute for Population Sciences in Mumbai. The data and more information are available online at: www.measuredhs.com.

[3]The five situations are "if she goes out without telling him;" "if her natal family does not give expected money, jewelry, or other items;" "if she neglects the house or the children;" "if she doesn't cook food properly;" "if he suspects her of having relations with other men."

Table 1: Summary statistics

	IHDS 2004-2005	NFHS-3 2005-2006
uses clean fuel %	37.36	15.41
first child female, %	49.20	47.90
household size	5.33	5.39
has electricity %	62.07	61.62
consumption per capita (Rupees per month), mean	719.44	
consumption per capita (Rupees per month), median	548	
scheduled caste, %	24.08	20.04
scheduled tribe, %	9.85	10.43
women's education, %:		
none	58.65	62.81
primary	14.72	15.12
secondary	20.83	19.96
tertiary	5.80	2.11
wife beating, %	43.68	
child bride, %	6.78	
teen bride, %	42.74	
observations	26734	56952

– which occurred in 44 percent of households – and 0 otherwise. We suspect that this indicator reflects a mixture of variation at the community and household level.

The other measures of status reflect the woman's age at marriage: an indicator for having been a child bride (7 percent) and an indicator for having been a teenaged bride (43 percent). Using the IHDS, Desai and Andrist (2010) find that low age at marriage is associated with "gender hierarchies." For example, women who marry at a younger age – controlling for a range of variables including education indicators and household consumption per capita – are more likely to practice *purdah*, less likely to eat with their husbands, and more likely to suffer mobility limitations.

3.1 Econometric Strategy

Section 4 focuses more narrowly on causal identification; here, the purpose of this section is to understand the mechanisms at work, and to verify that women's status appears important. In this empirical strategy, without a special natural experiment, identification depends on the classic "conditional independence assumption": in this case, that the residual correlates of clean cooking fuel use are uncorrelated with our indicators of women's status, conditional on the covariates. As always, this cannot be definitively proven, and its

plausibility depends on the set of control variables available.

We will estimate the linear probability regression[4]

$$\text{clean fuel}_{ij} = \alpha_j + \text{status}_{ij} B_1 + \text{education}^f_{ij} B_2 + \text{education}^m_{ij} B_3 + X_{ij}\theta + \varepsilon_{ij}, \quad (1)$$

where i indexes households, j indexes villages, and X_{ij} is a vector of controls. The *status* vector, wife beating and age at marriage, is unavailable in the NFHS. Sets of education dummies are included for the highest household male and female; these indicators offer finer resolution in the IHDS than in the NFHS. Fixed effects α_j are included at the village level. Village fixed effects would subsume village-level variation in the availability of clean fuel. In the presentation of results, estimates are included stepwise from a simple regression first, then with controls, then with fixed effects.

Women's status is not the only determinant of fuel type; wealth, availability, opportunity cost of time, and social setting all matter. We attempt to control for these factors. A broad range of controls is available from the IHDS. We use, as economic controls, consumption per capita (as a statistically significant quintic polynomial), a count of household assets (as a quadratic), indicators for possessing different types of government ration cards and for using a ration card, having electricity, and the number of hours per day of electricity. As social and demographic controls we include household size, number of married women, an indicator for the woman knowing that smoke from a traditional stove is unhealthy for children, an indicator for knowing a teacher, an indicator for knowing a medical care provider, and a set of eight dummies for religion and caste status: Brahmin, high caste, "other backwards caste," "scheduled caste" or Dalit, "scheduled tribe" or Adivasi, Muslim, Christian, and a combined category for Sikhs and Jains.

The NFHS does not permit as broad a range of controls. We constructed a wealth index from the first principal component of indicators for a set of 14 assets and measures of housing quality, and include it as a quadratic polynomial. We further include household size (as a quadratic) and dummy indicators for being Dalit, Adivasi, or "other backwards caste;" for a small village; for having a nuclear family; and for having a "below poverty line" ration card. In some regressions we included the height-for-age z-score of the oldest child (as a cubic polynomial) as an indicator of health and well-being.

The IHDS was, in part, a panel survey that re-interviewed households from the Human Development Profile of India, a 1993-94 survey of rural households. As a robustness check, we matched the IHDS data to the earlier data, and will additionally present results restricting the sample to households that did not report

[4]We use linear probability model because the coefficients are easier to interpret in a linear probability model. We also ran probit models and the marginal effects are quite similar to the OLS coefficients.

using clean cooking fuel in the HDPI. This will partially mitigate potential bias from high socio-economic status households both including higher-status women and having long used clean cooking fuel. This HDPI sub-sample is less than half the size of the full rural sample; although few rural households used clean cooking fuel in the early 1990s, the IHDS added many households to the original HDPI sample.

3.2 Results

Table 2 presents the estimates from these regressions. In the IHDS results, in all six regressions all three indicators of low status — wife beating, child marriage and teen marriage — have the expected negative coefficient (although not always statistically significantly in the small HDPI sub-sample). The coefficient on wife beating loses statistical significance when village fixed effects are added. For this variable, these fixed-effects might be over-controlling: mechanically, the survey question asked about beatings in the respondent's community, not her household; substantively, we might expect geography to be an important source of variation in women's status.

In both data sets, women's education is statistically significantly associated with clean fuel use, increasingly at higher levels of education. Importantly, this is true for *women's* education in particular. In the IHDS, equivalent indicators for men's education are omitted from the table because none are statistically significant. In the NFHS results, with a larger sample, indicators for men's education are statistically significant. However, the coefficients on women's education are larger. The penultimate rows of the table report results of an F test that jointly tests the null hypothesis that the women's coefficients at each of the three levels of education are equal to the men's coefficients at each respective level; equality is unambiguously rejected. Columns 10, 11, and 12 restrict the sample to those with information on a household child (the sub-sample that will be used in section 4) and include controls for child height-for-age. The results are generally similar.

As a placebo test, we estimated equation1 using the IHDS data with household electrification, rather than clean fuel use, as the dependent variable. Because the benefits of electricity do not particularly accrue to women, we would not expect women's status to importantly influence electrification. However, if the results above are merely confounded effects of high wealth or socioeconomic status, using electrification might produce similar results.

Estimating these regressions with an indicator for having electricity as the dependent variable does not replicate the women's status results. Wife beating has a statistically insignificant *positive* coefficient; the indicators for female education have statistically insignificant *negative* coefficients. The child and teenage marriage coefficients continue to have negative coefficients, but of order of magnitude 10^{-16}.

This first empirical strategy has shown that a set of indicators of women's status are associated with clean fuel use, and that women's education is more strongly associated than men's education, all robust to controlling for a range of economic, demographic, social, and geographic controls. Moreover, women's status is not similarly predictive of electrification. The next section considers another empirical strategy, designed to highlight the causal importance of women's status.

4 Strategy 2: Sex of the First-born Child

Indian households have a well-documented preference for sons over daughters (*e.g.* Sen, 2003). Having a first-born daughter, relative to having a first-born son, can therefore entail loss of intrahousehold status to the mother. In this section, we explore the sex of a household's first child as an exogenous source of variation in women's status.

Figure 1 depicts the basic result: at all wealth levels great enough that any households are using clean cooking fuel, households in which the first child is a boy are more likely to use clean fuel than households in which the first child is a girl. The rest of this section documents this effect in more detail.

4.1 Econometric Strategy

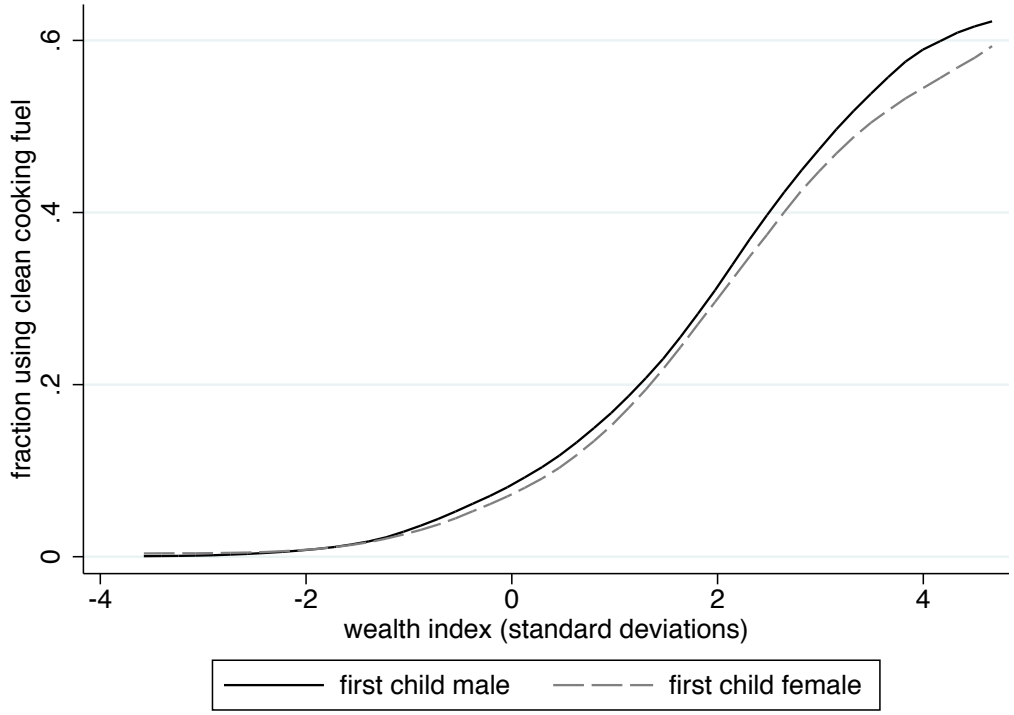
Like in Pham-Kanter (2010) study of maternal weight gain, but unlike in Washington (2008)'s study of U.S. legislators, our empirical strategy will compare the effect of first-born sons to first-born daughters, rather than study the sex composition of all of a household's children. This is because son preference could influence fertility stopping rules while having other effects (Clark, 2000), rendering overall sex composition endogenous. Barring sex-selective abortion, which we discuss below, the sex of the first child is more likely to be outside of households' control (Kishor and Gupta, 2009).

In particular, among households with at least one child, we estimate the linear probability regression

$$\text{clean fuel}_{is} = \alpha_s + \beta \text{female}_{is} + X_{is}\theta + \varepsilon_{is}, \tag{2}$$

where i indexes households, s indexes states, *female* is an indicator that the first-born child was female, and X_{is} is the same vector of controls used in section 3. Because the sex of the first child is like a random experiment, and because Freedman (2008) demonstrates that including covariates can bias experimental estimates, we will present results with and without covariates and state fixed effects.

Figure 1: Clean Fuel Use by Sex of First Child: Local Polynomial Regression, NFHS-3



As evidence towards considering the sex of the first child an exogenous source of variation in women's status, table 3 demonstrates the balance of observable characteristics across households with male and female first children. The only clear imbalance is in number of children, which very probably reflects causal effect: due to son preference, households with a female first child have more children. In the NFHS, households with a boy first child are slightly more likely to be of high caste ($p = 0.09$), but this is not statistically significant after a correction for multiple inference, and in the IHDS the difference is in the opposite direction.

4.1.1 Sex-selective abortion

The most important threat to this empirical strategy would be if sex-selective abortion were more likely to be practiced by richer households, such that the apparent effect of a boy baby on clean fuel use were, in fact, a confounded effect of wealth. In 2002, Arnold et al. reported that "ultrasound typically costs between 500 and 1,000 rupees (about $\$10 - 20$ at the exchange rate of Rs. $45 = 1$USD)."

Importantly, whether this is potentially a problem is *a priori* ambiguous. Sex selection of the first child would only cause endogeneity if it were correlated with omitted determinants of clean fuel use. Clark (2000) finds that in India, "socially and economically disadvantaged couples not only want but also attain a higher

proportion of sons..." so richer families may not be more likely to abort girls. Similarly, separating the effects of relative and absolute wealth using the NFHS-3, Gaudin (2011) documents that "higher absolute wealth is strongly associated with lower son preference, and the effect is 20-40 percent stronger when the household's community-specific wealth score is included in the regression."

In figure 1, we have already presented some evidence against this endogeneity concern: the effect is seen essentially throughout the wealth distribution, and at all wealth levels such that households use clean fuel. For further evidence, we interact $female_{is}$ in equation 2 with the state-level child sex ratio, to demonstrate that the effect on clean cooking fuel is not concentrated where sex-selection is most profound, and with an indicator for the woman having some education, to demonstrate that the effect is not concentrated on educated women who would have the most access to sex-determination technology. Finally, we will present a table of seven placebo regressions, demonstrating that having a female first child is not associated with owning other assets that would also be markers of wealth.

4.2 Results

4.2.1 Boy babies promote clean fuel use

Table 4 confirms the hypothesis behind this empirical strategy: in the NHFS-3, having a female first child, relative to having a male first child, is associated with a household being three-fourths of a percentage point less likely to use clean cooking fuel. This result is robust to including the vector of economic and social controls, as well as state fixed effects. Using the IHDS, reported in column 7, similarly finds a negative effect of a female first child, but does not have a large enough sample to estimate it with precision.

Columns 3, 5, and 6 are included as evidence against endogeneity due to sex-selective abortion. If sex-selective abortion were driving this result, then we would expect it to be strongest in the states with the most sex-selection. However, interacting the indicator for a female child with the state child (0-6) sex ratio[5] in columns 3 and 5 does not change the estimate of the coefficient on $female_{ij}$; indeed the interaction term is positive and statistically insignificant, suggesting that if anything the effect is weaker in absolute value in states with more male children. Similarly, in column 6, the interaction with an indicator that the woman has received some education is positive, implying that the effect is not concentrated among this group with more resources and access to clean fuel and medical services.

[5]These figures are taken from the 2001 Indian census: http://www.censusindia.gov.in/Tables_Published/A-Series/A-Series_links/t_00_004.aspx

4.2.2 Placebo tests

As a final check of the empirical strategy, table 5 replicates the first two columns of table 4, substituting for the dependent variable seven different indicators of asset ownership, including electrification as in section 3. These span the wealth distribution from owning a cot or bed to having a household bank account or radio. In none of these cases is there an effect of having a female first baby.

4.2.3 Instrumental variable

Because we are ultimately interested in the effect of women's status on clean cooking fuel use, not the effect of the sex of the first child *per se*, some readers might have anticipated that sex might be used as an instrumental variable. With such expectations, our second empirical strategy could be interpreted as a reduced form or "intent to treat" estimate. Our data do not permit a single two-stage estimation: it is the NFHS that allows a precise measure of the effect of a female baby, and the IHDS with a rich set of indicators of women's status. However, even if it were possible, we see two reasons why such an instrumentation strategy would not be appropriate.

First, the local average treatment effect – the average effect of changes in women's status due to the sex of the first child – is neither of economic nor policy interest. We do not believe that infants' sex is the most important determinant of women's status, and it is not under the control of policy makers. Second, we have not directly measured women's status; we have observed a range of indicators of it. This is useful to demonstrate the importance of status, but instrumental variable methods are designed to estimate causal effects, and we would not know what would be meant by, or how to interpret, a "causal effect" of an indicator.

However, some readers might nevertheless be interested in the result of an instrumentation strategy. In the spirit of a robustness check, we construct a Wald estimate of an "effect of women's status on clean fuel use" using both data sets. For the numerator, we use the NFHS estimates from the first column of table 4.

For the denominator, we use the IHDS to somewhat arbitrarily construct a binary indicator of high women's status in a household: whether the woman is above the median of a status index, where the status index is the first principal component of her needing permission in three settings, having responsibility to decide what to buy and and what to cook, discussing things with her husband in three situations, and having cash on hand. Having a first girl baby instead of a first boy baby is associated with a 0.0155 lower probability of being above the median of the status index. This has a standard error of 0.0102, so in addition to the other concerns, with a t-statistic of 1.53 this estimate also suffers from weak instrument concerns.

Putting these together, we would find:

$$\hat{\beta}^{Wald} = \frac{\left(\overline{\text{clean fuel}}\middle|\text{ girl baby}\right) - \left(\overline{\text{clean fuel}}\middle|\text{ boy baby}\right)}{\left(\overline{\text{status}}\middle|\text{ girl baby}\right) - \left(\overline{\text{status}}\middle|\text{ boy baby}\right)} = \frac{-0.00757}{-0.01554} = 0.49. \tag{3}$$

Taken literally, this computation would imply that moving a household's woman from the bottom to the top half of this intrahousehold status divide would increase by 49 percentage points its chances of using clean cooking fuel. To reiterate, it is unclear both how to interpret this figure and how to assess its credibility. Fortunately, it is unnecessary to the argument of this paper. In any event, the result of this exercise is not inconsistent with the hypothesized important role for women's intrahousehold status in promoting clean cooking fuel use.

5 Conclusion

Why is the take-up of clean cooking fuel low in rural Indian households? There are surely many reasons, and this paper has documented the importance of one: women's intrahousehold status. The costs of using traditional cooking fuel fall disproportionately on women, and women "are the most deprived in the way Indian households are run."

Our paper is not the first to hypothesize that women's status may influence fuel choice. However, we have focused on documenting its importance, and have implemented a combination of strategies to verify that women's status has a causal effect. In the first strategy, indicators of low women's status were associated with lower use of clean fuel, despite controlling for a broad range of economic and social variables. Women's education was more strongly predictive than men's education. None of these findings were similarly true when electrification, with no special role for women, was used as the dependent variable.

The second strategy argued that, because of widespread Indian son preference, having a girl first child lowers women's status relative to having a boy first child, and is therefore associated with a three-fourths of a percentage point reduction in the likelihood of using clean fuel. This effect is found essentially throughout the wealth distribution, persists despite economic and social controls, and is not concentrated among households in states with a high child sex ratio or households where women have received education. Using many other assets as dependent variables – including electrification – no similar effect of having a girl first child was found.

One limitation of the second strategy is that, while the first strategy has indicated the involvement of women's status, we cannot rule out that part of the effect of having a boy baby on fuel choice is because

the household wishes to invest in *his* health. However, as table 3 reported, parents of first boys and girls are not differentially likely to report that smoke from traditional fuel is harmful to a child's health. More importantly, even this mechanism would reflect a role of gender inequality in household fuel use. Fully separating these two mechanisms is left for future research.

References

Arnold, Fred, Sunita Kishor and T. K. Roy (2002), 'Sex-selective abortions in India', *Population and Development Review* **28**(4), 759–785.

Bardhan, Pranab (2011), 'Challenges for a minimum social democracy in india', *Economic and Political Weekly* .

Clark, Shelley (2000), 'Son preference and sex composition of children: Evidence from India', *Demography* **37**(1), 95–108.

Dahl, Gordon B and Enrico Moretti (2008), 'The demand for sons', *Review of Economic Studies* **75**, 1085–1120.

Desai, Sonalde, Amaresh Dubey, B.L. Joshi, Mitali Sen, Abusaleh Shariff and Reeve Vanneman (2007), India human development survey, computer file ICPSR22626-v2, University of Maryland and National Council of Applied Economic Research, New Delhi.

Desai, Sonalde and Lester Andrist (2010), 'Gender scripts and age at marriage in India', *Demography* **47**(3), 667–687.

Duflo, Esther, Michael Greenstone and Rema Hanna (2008), 'Cooking stoves, indoor air pollution and respiratory health in rural orissa', *Economic and Political Weekly* .

ESMAP (2003), Household energy use in developing countries: A multi-country study, mimeo, World Bank.

Freedman, David A (2008), 'On regression adjustments to experimental data', *Advances in Applied Mathematics* **40**, 180–93.

Gaudin, Sylvestre (2011), 'Son preference in Indian families: Absolute versus relative wealth effects', *Demography* **48**, 343–370.

Hoddinott, J. and L. Haddad (1995), 'Does female income share influence household expenditures? evidence from côte d'ivoire', *Oxford Bulletin of Economics and Statistics* **57**(1), 77–96.

Israel, D. (2002), 'Fuel choice in developing countries: Evidence from bolivia', *Economic Development and Cultural Change* **50**(4), 865–890.

Kishor, Sunita and Kamal Gupta (2009), Gender equality and women's empowerment in India, National Family Health Survey, International Institute for Population Sciences, Mumbai, India.

NSSO (1997), Energy used by indian households, july 1993-june 1994, Report 410/2, National Sample Survey Organization, Ministry of Statistics and Programme Implementation, Government of India.

NSSO (2007), Energy sources of indian households for cooking and lighting, 2004-05, Report 511, National Sample Survey Organization, Ministry of Statistics and Programme Implementation Government of India.

Pham-Kanter, Genevieve (2010), The gender weight gap: Sons, daughters, and maternal weight, working

paper 28997, MPRA.

Sen, Amartya (2003), 'Missing women revisited: Reduction in female mortality has been counterbalanced by sex selective abortions', *British Medical Journal* **327**, 1297–8.

Smith, Kirk R (2000), "'national burden of disease in india from indoor air pollution."', *PNAS* **27**(4).

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

Udry, Christopher (1996), 'Gender, agricultural production, and the theory of the household', *Journal of Political Economy* **104**(5), 1010–1046.

Washington, Ebonya L (2008), 'Female socialization: How daughters affect their legislator fathers' voting on women's issues', *American Economic Review* **98**(1), 311–332.

Table 2: Clean cooking fuel use and indicators of women's status

Panel A: India Human Development Survey

	(1)	(2)	(3)	(4)	(5)	(6)
	full rural sample			traditional fuel in 93-94 data		
wife beating	-0.0484***	-0.0352**	-0.00117	-0.0544**	-0.0400*	-0.0153
	(0.0127)	(0.0129)	(0.0111)	(0.0183)	(0.0180)	(0.0155)
child bride	-0.115***	-0.0537**	-0.0287*	-0.0680*	-0.0221	-0.0132
	(0.0203)	(0.0186)	(0.0137)	(0.0322)	(0.0292)	(0.0215)
teen bride	-0.0691***	-0.0424***	-0.00450	-0.0435**	-0.0337*	-0.00982
	(0.0107)	(0.0106)	(0.00805)	(0.0155)	(0.0145)	(0.0122)
woman's education:						
primary	0.119***	0.0236+	0.0169+	0.0985***	0.00235	0.0246+
	(0.0132)	(0.0128)	(0.00986)	(0.0188)	(0.0175)	(0.0148)
middle	0.215***	0.0407**	0.0236+	0.214***	0.0362+	0.0493*
	(0.0177)	(0.0147)	(0.0126)	(0.0287)	(0.0208)	(0.0191)
completed 10th	0.353***	0.0778***	0.0730***	0.379***	0.103**	0.134***
	(0.0215)	(0.0209)	(0.0211)	(0.0325)	(0.0341)	(0.0357)
further secondary	0.398***	0.0857**	0.0581*	0.376***	0.0826	0.0604
	(0.0263)	(0.0293)	(0.0251)	(0.0471)	(0.0570)	(0.0489)
tertiary	0.479***	0.0709**	0.0671**	0.491***	0.102*	0.111**
	(0.0250)	(0.0249)	(0.0228)	(0.0385)	(0.0398)	(0.0366)
controls		✓	✓		✓	✓
village FEs			✓			✓
n	26734	26734	26734	12442	12442	12442

Panel B: National Family Health Survey

	(7)	(8)	(9)	(10)	(11)	(12)
	full rural sample			sub-sample with child information		
women's education:						
primary	0.0741***	0.00747	0.00848+	0.0573***	0.00202	0.00371
	(0.00645)	(0.00508)	(0.00451)	(0.00758)	(0.00652)	(0.00681)
secondary	0.240***	0.102***	0.0743***	0.189***	0.0793***	0.0534***
	(0.0101)	(0.00712)	(0.00562)	(0.0112)	(0.00887)	(0.00850)
tertiary	0.493***	0.238***	0.170***	0.461***	0.242***	0.162***
	(0.0185)	(0.0162)	(0.0154)	(0.0302)	(0.0248)	(0.0250)
men's education:						
primary	0.0293***	-0.0124***	-0.00974**	0.0253***	-0.00979*	-0.00579
	(0.00397)	(0.00335)	(0.00317)	(0.00493)	(0.00456)	(0.00486)
secondary	0.116***	0.0266***	0.0278***	0.0853***	0.0143**	0.0147**
	(0.00582)	(0.00417)	(0.00386)	(0.00626)	(0.00470)	(0.00485)
tertiary	0.331***	0.147***	0.128***	0.290***	0.137***	0.121***
	(0.0156)	(0.0118)	(0.0103)	(0.0202)	(0.0167)	(0.0162)
controls		✓	✓		✓	✓
village FEs			✓			✓
woman's > man's: F_3	47.77	26.34	14.68	26.23	13.58	4.91
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0021
n	56952	56952	56952	23601	23601	23601

Standard errors clustered by primary sampling unit (village) in parentheses; estimates weighted to reflect sampling strategy.
Two-sided p-values: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Balance of covariates, by sex of first child

	IHDS			NFHS		
	male	female	p-value	male	female	p-value
household size	5.616	5.677	0.131	6.537	6.561	0.570
	(0.038)	(0.041)		(0.040)	(0.041)	
head age	44.594	44.560	0.624	42.794	42.830	0.872
	(0.165)	(0.165)		(0.189)	(0.194)	
has BPL card	0.359	0.357	0.941	0.296	0.296	0.958
	(0.008)	(0.008)		(0.007)	(0.007)	
children in household	1.968	2.136	0.000	1.540	1.604	0.000
	(0.025)	(0.027)		(0.010)	(0.011)	
high caste	0.201	0.207	0.508	0.275	0.264	0.090
	(0.007)	(0.007)		(0.009)	(0.009)	
nuclear family				0.476	0.472	0.556
				(0.006)	(0.007)	
female years of education	4.321	4.375	0.498			
	(0.083)	(0.084)				
knows smoke unhealthy	0.786	0.786	0.802			
	(0.008)	(0.008)				

Table 4: Clean cooking fuel use and sex of the first child

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	National Family Health Survey						IHDS
female child	-0.00757+	-0.00730*	-0.00775*	-0.00663+	-0.00685+	-0.00814+	-0.0000189
	(0.00451)	(0.00360)	(0.00364)	(0.00356)	(0.00360)	(0.00479)	(0.0108)
female			0.000687		0.000475		
× child sex ratio			(0.000987)		(0.000976)		
state child sex ratio			-0.00144				
			(0.000980)				
female						0.00178	
× woman's education						(0.0120)	
woman has some						0.202***	
education						(0.0112)	
controls		✓	✓	✓	✓		
state FEs				✓	✓		
constant	0.115***	0.286***	0.284***	0.274***	0.273***	0.129***	0.372***
	(0.00558)	(0.0127)	(0.0127)	(0.0125)	(0.0125)	(0.00559)	(0.0116)
n	27569	27569	26965	27569	26965	27569	19909

Standard errors clustered by primary sampling unit (village) in parentheses; estimates weighted to reflect sampling strategy. Child sex ratio is $\frac{\text{males}}{\text{females}} \times 100$, by state. Variables interacted with *female child* are de-meaned. Two-sided p-values: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5: Placebo regressions: "effect" of female first child on other assets, NFHS

	(1) clean fuel	(2) electricity	(3) cot or bed	(4) pressure cooker
female child	-0.00757+	-0.00393	0.00445	0.000204
(no controls)	(0.00451)	(0.00763)	(0.00517)	(0.00630)
female child	-0.00730*	-0.00417	0.00419	0.000553
(controls)	(0.00360)	(0.00577)	(0.00494)	(0.00461)
constant	0.115***	0.542***	0.843***	0.250***
(no controls)	(0.00558)	(0.0111)	(0.00594)	(0.00710)

	(5) agricultural land	(6) bicycle	(7) radio	(8) bank account
female child	0.00509	0.00351	0.00376	-0.00537
(no controls)	(0.00768)	(0.00741)	(0.00673)	(0.00718)
female child	0.00594	0.00201	0.00388	-0.00543
(controls)	(0.00686)	(0.00709)	(0.00624)	(0.00638)
constant	0.559***	0.540***	0.274***	0.334***
(no controls)	(0.00852)	(0.00784)	(0.00550)	(0.00664)

Standard errors clustered by primary sampling unit (village) in parentheses; estimates weighted to reflect sampling strategy. Each column uses a different independent variable and reports two regressions, one with and one without control covariates. Two-sided p-values: + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.