Negotiating clean energy: Women's Bargaining Power and impact of PMUY

Massimo Filippini[1,3] and
Keshav Sureka*[1]

[1]Center of Economic Research (CER-ETH), ETH Zürich, Switzerland
[3]Università della Svizzera italiana, Switzerland

Last revision: September 8, 2025
(Preliminary: Please do not cite)

Abstract

Access to clean cooking fuel remains a persistent challenge for billions globally, with women bearing a disproportionate share of the associated health burden. This paper provides an empirical analysis of India's Pradhan Mantri Ujjwala Yojana (PMUY)—the world's largest initiative to promote LPG adoption among poor households. Using a difference-in-difference approach with propensity score matching on nationally representative survey data, we quantify the impact of PMUY on clean fuel uptake and women's health outcomes, focusing especially on tuberculosis and hemoglobin levels. The study also examines the central role of women's bargaining power within households and the presence of cheap alternative fuels as key factors that modulate the effectiveness of the program. We find a statistically significant increase in LPG adoption among eligible households, with program effects amplified in those where women have greater agency. Furthermore, women in treated households experienced notable reductions in tuberculosis incidence and improvements in hemoglobin metrics relative to matched controls. These results underscore the importance of targeting intra-household dynamics and local fuel availability in designing and implementing successful clean cooking interventions.

*Corresponding author at Center of Economic Research (CER-ETH), ETH Zürich, Zürichbergstrasse 18, 8032 Zürich, Switzerland. <ksureka@ethz.ch>.

1 Introduction

In 2015, approximately 2.7 billion people around the world lived without access to clean cooking fuels and technology, according to the IEA report (IEA (2015)). Out of this, about 25% resided in India, making up 56% of the Indian households (NFHS (2017)). Alternatives to clean cooking fuel, wood, dung, crop waste, etc., lead to high levels of indoor air pollution (IAP) of the home, causing chronic heart and lung diseases. Each year, around 3.2 million people die prematurely from a disease attributed to IAP, of which India alone is home to one million (WHO (2014)). Early studies (Bailis et al. (2005)) indicate that indoor air pollution will lead to 9.8 million premature deaths worldwide by 2030. Since women spend considerably more time doing housework than men (Duflo (2012)), the largest burden of these deaths falls on women living in low- and middle-income countries.

In light of these facts, the United Nations has established "access to affordable, reliable, sustainable, and modern energy for all" as one of the 15 Sustainable Development Goals (SDG 7). In response, many developing countries have initiated programs and initiatives to promote the transition to cleaner cooking fuels. For example, Indonesia has launched a large-scale program aimed at reducing the use of kerosene and encouraging the adoption of Liquefied Petroleum Gas (LPG), which is considered a cleaner energy source compared to traditional fuels such as wood, dung, and crop waste used in cooking. The idea was to phase out the kerosene subsidy and instead invest in building LPG infrastructure. However, cooking behavior is formed over long periods and is difficult to change (Lindgren (2020)). From a policymaker's point of view, it is important to evaluate the impact and factors affecting the efficiency of such programs that aim to transition to cleaner cooking fuel and address IAP.

This paper offers an empirical evaluation of the world's largest government-led LPG transition initiative, Pradhan Mantri Ujjwala Yojana (PMUY), focusing on both the adoption of LPG and its subsequent effects on women's health. Particular attention is paid to the role of women's agency in facilitating the adoption of LPG within households. To this end, we first present an illustrative model of intra-household decision making, delineating the barriers that impede the transition to LPG. We specify which obstacles the program is designed to address and identify those that persist despite intervention.

The empirical analysis proceeds in two main parts. First, we employ a difference-in-difference approach with matching techniques to estimate the impact of PMUY on household-level LPG adoption. Additionally, a triple difference method is used to assess whether program effects are magnified in households where women possess greater bargaining power. In the second part, we examine the influence of the scheme on the health outcomes of individual women in the targeted group, focusing specifically on tuberculosis incidence and hemoglobin levels. Again, we utilize a difference-in-difference analysis with matching to gauge these health effects. To control for broader external factors that can affect general health, we also compare the outcomes for women with those for men within the same households.

To identify the impact of the program on the adoption of LPG and on women's health, we use the fact that the scheme was provided to only a specific group of households. The primary objective of the scheme was to improve access to clean fuel (LPG) for poor households by providing deposit-free LPG connections to adult women from poor households. Under the scheme, the household will pay zero upfront cost of installation. The central government would provide financial support for about 50% of the installation cost for eligible households, and the other 50% would be paid out of the subsidy on additional fuel purchases.

The empirical strategy employs a difference-in-difference analysis using repeated cross sections where households with "below poverty line (BPL)" cards are the group eligible for treatment and others as the control group to estimate the impact of policy. To address concerns about the parallel trend assumption, we use propensity score matching to match households in the treated group with households in the control group based on the criteria for BPL card eligibility. We cannot directly test the parallel trend assumption because we have data only for two periods after the eligibility criteria were defined. We used nationally representative data from the National Family Health Survey (NFHS) 2015 and 2019 to collect information on household cooking methods and whether they have BPL cards. We also collect information about the hemoglobin levels of men and women in households from this data set. The timing of the data sets is appropriate for our analysis because the scheme was launched in 2016, and the target number of connections was completed in 2019. To estimate the program's impact on adopting LPG, we use a binary variable indicating LPG as the primary cooking fuel as our dependent variable. We used hemoglobin levels and tuberculosis cases of men and women as dependent variables to estimate the health impact of the program.

Our analysis reveals that households exposed to the PMUY program experienced a statistically significant increase in LPG adoption compared to the control group. The triple difference analysis further indicates that the effect of the program was especially pronounced in households where women had greater bargaining power, as measured by two indicators: female household leadership and cases where the wife earns more than the husband. In contrast, the impact of the program was less substantial in rural areas relative to urban settings, reflecting the widespread availability of cheaper alternative cooking fuels such as crop and animal waste in rural locations. Supporting this interpretation, we also find a diminished program effect among households owning cattle or agricultural land. These heterogeneous effects are consistent with the barriers identified in our intra-household bargaining framework.

Regarding health outcomes, we find that the women in the households exposed to treatment exhibited larger reductions in tuberculosis incidence and hemoglobin levels compared to their counterparts in the control group [1] compared to the control group. These findings are consistent with the hypothesis that women, as the primary users of cooking fuel and spending more time indoors, should have a greater health benefits of the program than men.

This paper contributes to several strands of the literature. First, it contributes to the literature evaluating the impact of the PMUY on LPG use in India in two key ways. First, it is a national-level analysis with household and individual-level data. Second, we aim to establish the causal impact of the scheme. Several papers evaluating the scheme focus on specific regions (Swain and Mishra (2020), Gill-Wiehl et al. (2022)) or provide descriptive analysis of the impact of the scheme (Dabadge (2018), Ranjan and Singh (2020)). In terms of this contribution, the paper closest to ours is Asharaf and Tol (2024). They also use the NFHS data set of the two years to make a causal estimation of the impact. They focus on the difference in impact among different social classes and geographical regions. We complement their analysis by highlighting the role of intra-household and economic barriers on the impact of adoption.

Second, our paper contributes to the limited but growing literature on evaluating the health impact of programs aimed at reducing indoor air pollution through incentives for adopting clean energy sources. To our knowledge, we know only four studies evaluating directly the impact of policy efforts, such as subsidies on clean energy, on health outcomes. Nandwani and Jain (2024) report that health and respiratory ailments declined for the households exposed to the PMUY program by 7 and 13 %. Calzada and Sanz (2018) analyzed Peru's gas cookstove adoption subsidy (FISE) program but found no impact on health outcomes. Gould et al. (2023) analyzed the gas stove to electric cooking transition program in Ecuador and found that the program reduced all-cause and respiratory-related hospitalizations nationwide. Verma and Imelda (2023) analyzed the health impact of substituting kerosene subsidy with LPG subsidy in Indonesia and found that this increased the lung capacity of women.

Third, our paper relates to the extensive body of literature highlighting the non-unitary nature of household decision-making (see Chiappori and Mazzocco (2017) for a review). We highlight how such a non-unitary nature (even in the collective decision-making context) of decision-making can adversely affect women more than men. Thus, our paper relates to the literature documenting worse socio-economic, health, and education outcomes for women than men in the *same household* in low and middle-income countries arising from patriarchal norms (Sen (2017), Klasen and Wink (2002), Chatterjee (1990), Muralidharan and Prakash

[1] Generally speaking, higher hemoglobin levels indicate better health. However, exposure to pollutants tends to increase hemoglobin concentration as a compensatory mechanism for decreased quality of red blood cells associated with reaction with the pollutants (Aitchison and Russell (1988)). Thus, smokers, for example, maintain higher hemoglobin levels than non-smokers (Malenica et al. (2017)). We use a non-linear IV approach to test the causal connection between LPG use and hemoglobin levels (see appendix).

(2017)). The traditional gender roles lead to poor nutritional intake, lower school enrollments, and generally reduced economic opportunities for women, giving rise to a lower quality of life for women than men, even within the same household. We highlight another channel through which such gender roles can lead to worse health outcomes for women than men. We find that households headed by a male member or households where the husband earns more than the wife were less likely to adopt LPG as primary cooking fuel. This is indicative of the non-unitary nature of the household norms that yield worse health outcomes for women than men.

Finally, our paper contributes to the literature examining the barriers to LPG adoption and consumption in India. Afridi et al. (2021) uses random assignment of villages to information treatment to find that an increase in health and subsidy awareness in the district of Indore in India led to 13% rise in LPG consumption. Afridi et al. (2024) find that a rise in the over-the-counter price of LPG leads to a significant decline in its consumption, especially among low-income households. In our paper, we find that the impact of the policy was lowest on households with relatively easy access to alternative cooking fuels like dung and wood. For example, households with livestock or agricultural land had significantly less impact.

The structure of the paper is as follows. We begin by providing an overview of the PMUY program and its implementation. Next, we introduce a theoretical model of intra-household decision-making, which serves to elucidate the specific constraints addressed by the program and underscores the influence of women's bargaining power on LPG adoption. We then detail the data sources and the process for constructing the counterfactual. Section 6 outlines the methodology employed in both components of the empirical analysis, as well as potential threats to identification. Section 7 presents the empirical results and their interpretation, followed by a discussion of the key findings

2 Program details and background

In 2005, about 66% of households in India used solid fuels like wood, animal dung, and coal for cooking. It improved modestly by 10% over the next 10 years to around 56% in 2015. The partially burned particles and toxic fumes from these fuels harm human health and primarily contribute to household air pollution. Transitioning to cleaner fuels like LPG could improve the situation; however, it required a high upfront installation cost of about 3000 rupees. This was a considerable amount given that in 2011, about 60% of the population lived on less than 35 Rs per day in rural areas and 66 Rs per day in urban areas NSSO (2013). In light of these facts, the central government of India launched a program to facilitate the transition to clean cooking fuel in India.

PMUY or *Pradhan Mantri Ujjawala Yojana* is a central government program launched in 2016 in India with the aim to "safeguard the health of women and children by providing them with a clean cooking fuel" - Liquified Petroleum Gas (LPG). The initial target of the scheme was to provide 50 million deposit-free connections, which was revised to 80 million gas connections, which was achieved in September 2019 (MoPN (2023)). The all-India coverage of LPG increased from 61.90% in 2016 to 95% in April 2019 (CAG (2019)). The program's second phase, PMUY-II, was launched in August 2021 with a target to release an additional 10 million connections, which was fulfilled in January 2022. An additional 6 million connections were released under the second phase until December 2022. In this paper, we evaluate the effect of the first phase of PMUY.

2.1 Financial Assistance

Under the program, the government would bear 1600 rupees per connection towards security deposits for LPG cylinders, pressure Regulators, installation charges, etc., as one-time financial assistance to adult women of BPL households. It was prescribed that the oil marketing companies (OMCs) [2] would provide an option to the beneficiaries to opt for a loan, if they so desire, to cover the cost of the cooking stove and first refill. The OMCs would recover the EMI of the loan amount from the subsidy amount due to the beneficiaries on refills (CAG (2019)). The main idea behind the program was to remove the one-time high cost of transitioning to a new cooking method so that households can be nudged towards using LPG.

S. No.	Particulars	Amount in Rupee
1	Security Deposit (14.2 Kg LPG Cylinder)	1250
2	Security Deposit (Pressure Regulator)	150
3	Suraksha Hose	100
4	Domestic Gas Consumer Card Booklet	25
5	Installation, Administrative charges	75
A	Total (1+2+3+4+5)	1600 (Assistance from GOI)
6	Cost of Stove	990
7	Cost of Refill (indicative) for 14.2 kg cylinder	517
B	Total (6+7)	1507 (Optional loan facility by OMCs)
C	Grand Total (A+B)	3107
D	Budgetary support from Central Govt.	1600
E	Finance (or Loan) to Beneficiary by OMCs	1507

Table 1: Break up of financial assistance and details of loan amount

[2] Indian Oil Corporation Limited (IOCL), Bharat Petroleum Corporation Limited (BPCL) and Hindustan Petroleum Corporation Limited (HPCL)

2.2 Targeted Beneficiaries

The targeted beneficiaries of the program's first phase were women from "Below Poverty Line (BPL)" households appearing in the Socio-Economic Caste Census (SECC) 2011. Under the scheme, the connection was to be provided under the name of an adult female household member. This was a conscious decision by policymakers to empower women in poor households. The SECC used a three-step method detailed in the appendix to identify the BPL households. The method used seven deprivation criteria and poverty caps to identify the BPL households. However, the use of only seven indicators caused bunching in scoring. In other words, those seven items do not provide enough variation in deprivation counts to match poverty caps precisely. Moreover, slight changes in poverty estimation methods lead to different households being identified as BPL to the extent of 25% Alkire and Seth (2013). We use this less-than-precise identification of BPL households to construct our control group using a matching method.

3 A model of Intra-household fuel decision

In this section, we present an intra-household decision-making model to illustrate the barriers to the adoption of LPG. We use the collective models of intra-household decision-making in a static framework studied in detail by Bourguignon et al. (2009); Browning and Chiappori (1998); Chiappori and Ekeland (2006, 2009).

We assume that there are two members in the household h, w, and they collectively optimize a weighted sum of the utility function. Thus, the outcomes are Pareto efficient. The weight parameter is denoted by μ, which depends on two parameters, α and z. α captures the share of contribution by w to household income Y. z denotes other exogenous variables that influence the weighting parameter. To maintain the focus of the model on cooking fuel usage, we assume that the utility of the members is increasing, monotonic, and concave in food consumption F^i and health status H^i. Food is produced according to a production function f that depends on the fuel used and other raw materials denoted by X priced at P_X. Two types of fuel can be used: LPG at price P_{LPG} or other fossil fuel OF at price P_{OF}. The process of food production also produces indoor air pollution and toxins denoted by T. To use LPG, a household has to spend an amount B as installation cost. Finally, to capture the influence of gender norms, we let the indoor air pollution T affect utility only for w such that $U'_T \leq 0, U''_T \geq 0$, and the health status of w is also influenced adversely by T. Also, we assume that $\frac{\partial T}{\partial OF} \geq \frac{\partial T}{\partial LPG}$ for all levels of LPG and OF.

Thus, the household maximizes a weighted sum of the utility of household members subject to food production constraint, the indoor air pollution constraint, the health constraint of household members, and the budget constraint. The decision variable involves the division of food produced $F = \{F_w, F_h\}$, the decision to install LPG, $a \in \{1, 0\}$, and the optimal fuel input combination between LPG and $Nlpg$, and finally the raw material input X. This problem can be written as follows.

$$\begin{aligned}
\underset{F,a,LPG,Nlpg,X}{max} \quad U^h(F^h, H^h) + \mu(\alpha, z) \cdot U^w(F^w, H^w, T) \\
s.t. \\
F^h + F^w = f(a \cdot LPG, OF, X) \\
T = T(a \cdot LPG, OF, X);\ H^w = H^w(T) \\
P_{LPG} \cdot LPG + P_{OF} \cdot OF + P_X \cdot X \leq Y - a \cdot B
\end{aligned} \tag{1}$$

In the above setup, three barriers to the use of LPG can be identified. In the empirical exercise, we observe the role of all these barriers, which is consistent with our model.

- Installation cost: If B is high relative to income, the LPG setup will not be installed. This is the barrier that the government attempts to remove through the program. Through gov

- Opportunity cost: The Second barrier associated with transition is the presence of alternative fuels like wood, dung-cake, crop husk, etc. If these alternative fuels are readily available at low prices, then the LPG use will be low. This means we will observe low program impact in rural areas and for households with land or farm animals.

$$\frac{\partial LPG}{\partial P_{Nlpg}} \geq 0 \tag{2}$$

- Intra-household decision: The Third barrier is associated with the bargaining power between men and women in the household. If the bargaining power of women is low in the household, then the use of LPG will also be reduced.

$$\frac{\partial LPG}{\partial \mu} \leq 0 \tag{3}$$

4 Data

We employ two waves of *National Family Health Survey* (NFHS) conducted in 2015 and 2019 in this paper. NFHS is a large-scale survey conducted in a representative sample of households throughout India. Both waves include information on about six hundred thousand households. The survey is organized to provide essential data on health and family welfare, focusing more on women's health and factors that can impact women's health, like the primary cooking fuel used in the household, the location of the kitchen, etc. It is important to note that the survey does not provide information on the amount of LPG or other fuels used. It only provides which cooking fuel was primarily used by the household. Thus, we observe the following $Dlpg_i$. That is, we observe the household to be primarily using LPG as cooking fuel only if its proportional use by the household exceeds a certain threshold γ_i. This is important because, as we explain later, it influences the interpretation of the coefficient that we estimate.

$$Dlpg_i = \begin{cases} 1 & if \ \frac{LPG_i}{Nlpg_i} \geq \gamma_i \\ 0 & if \ \frac{LPG_i}{Nlpg_i} < \gamma_i \end{cases} \tag{4}$$

The survey also collects information on whether the household had BPL cards and information on some of the deprivation criteria used in SECC to classify BPL households[3]. In addition to household-level variables, the survey provides information on household members, primarily women. We use the household member information to gain information on hemoglobin levels, tuberculosis, age, and smoking habits, among other variables. The two variables considered to reflect the bargaining power within a household are the gender of the head of the household and the indicator on who earns more. Note that in the first part of the empirical exercise, i.e. effect of policy on LPG adoption and heterogeneity analysis, the household is the unit of analysis. Whereas in the second part of the empirical exercise of the impact of policy on health outcome, the individual is the unit of analysis.

4.1 Constructing counterfactual

The program was intended for a specific group of households, i.e., the Below Poverty Line households (BPL), which created two groups, where one was intended for the treatment and the group was not. Along with repeated cross-section data before and after policy implementation, this creates a natural setting for applying the standard difference-in-difference approach to estimate the causal effect of the policy. However, the targeted beneficiaries of the program were not randomly assigned but were deliberately chosen to be the poorer households, as described in the targeted beneficiaries section above. This means that households in the BPL and non-BPL groups were systematically different in ways that could influence the cooking fuel choice. This means that we have to be careful with constructing a counterfactual. In the literature, to address these concerns, we can find two approaches based on weighting the importance of observations. The first approach uses matching methods to find similar households in the control group and then employs only the matched sample with the respective weight for further econometric analysis. This method has been used by Asharaf and Tol (2024) to estimate the impact of policy on the adoption of LPG. The second approach suggested by Sant'Anna and Zhao (2020) is to first estimate inverse probability

[3] we provide details on these variables in the appendix

weights using non-linear regression for all observations and then use these weights in further econometric analysis. In this paper, we use the matching approach with weights in the main part of the analysis and the second approach as a robustness check.

4.2 Matching

For the first approach, we utilize propensity score matching to match households from the BPL group, i.e., the group intended for treatment, to the non-BPL group. To do this, we exploit the fact that the NFHS survey provides information on whether the household had a BPL card and five out of the seven deprivation/inclusion criteria used in SECC to classify BPL. These criteria had influenced about 97% of BPL households in the SECC (see appendix for details on the criteria). Using a logit model and these variables, we construct propensity scores for each household to be categorized as BPL. The model is as follows. z is a linear function of variables x provided in table 2 below. The distribution of propensity scores of the matched control group and targeted beneficiary is given in Figure 4.2 below.

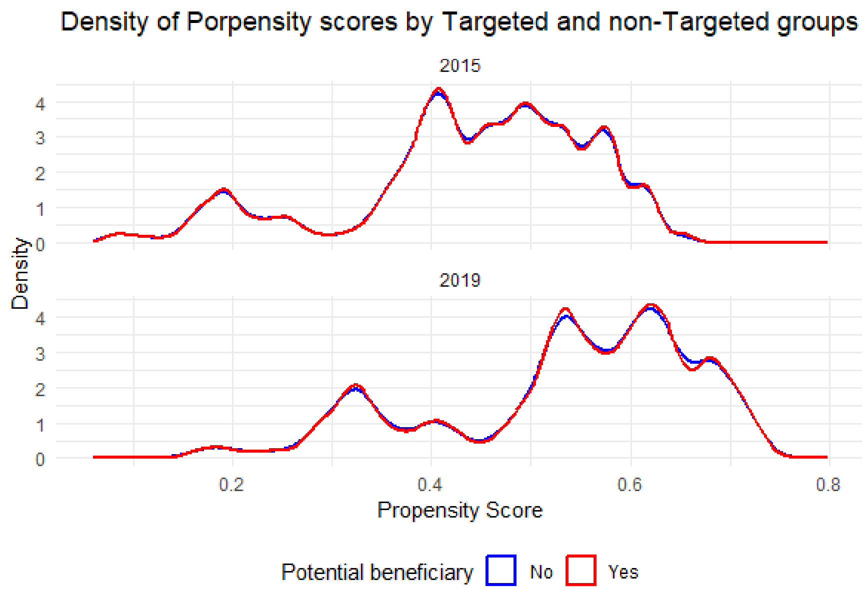
$$Pr(BPL_i = 1|x_i) = [1 + e^{z_i}]^{-1} \tag{5}$$

$$z_i = x_i\beta \tag{6}$$

SECC criteria	NFHS variable (x_i)	% of BPL household in SECC
D1 D1 D1	Number of in household Rooms Room material Wall material	13.28%
D3	Sex of the head of the household	3.86%
D5	Caste	21.56%
D6	No Educated adult	23.52%
D7	Acres of land	30.04%
Automatic Exclusion Criteria		
	Car	
	Refrigerator	

Table 2: Variables used in constructing propensity scores.

After calculating the propensity score, we find matches for BPL households in the same district from the non-BPL households using the nearest-neighbor matching method with replacements. The idea is that the household in the non-BPL group most similar to the household in the BPL group should be included in the analysis. Thus, multiple BPL households can match with one non-BPL household, and many non-BPL households will not have any match. The number of BPL units that match each non-BPL unit determines the weight



of that non-BPL unit. The weights are then normalized such that the sum of the weights of the control unit equals the number of control units matched. We do this matching exercise separately for two years and find the following number of matches (almost all BPL households were matched to some non-BPL households). [4] In Table 3, we present information regarding the number of households considered in both waves (All non-BPL) and the subset of households included after matching for econometric analysis. For example, in 2015, we identified 218397 treated households and 104218 untreated households through the matching procedure.

In the first part of the empirical exercise, we consider socioeconomic variables at the household level as well as the individual level, such as the sex of the head of the household, the presence of an educated adult in the household, whether any member in the household has a bank account, the wealth index of the household, an indicator of who earns more in the household, among others. Table 4 provides summary statistics about the treated households (BPL) and the untreated households (Matched Non-BPL) considered in the empirical analysis. We also included descriptive statistics on the complete non-BPL sample (All non-BPL) before matching for comparison purposes.

In the second part of our empirical analysis, we examine individual data from women living in both treated and untreated households following matching. We present descriptive statistics on various socioeconomic characteristics of these women in Table 5. It is important to note that since multiple women can reside in a single household, the number of observations included in the analysis using individual data is greater than the number of observations in the household-based analysis.

[4]The matching exercise was performed in R software using the MatchIt package.

Household	Year 2015	Year 2019
Targeted(BPL)	218,397	293,766
All non-Tar(non-BPL)	356,018	310,028
Mathced non-Tar(non-BPL)	104,218	115,347

Table 3: Number of targeted and matched control households

5 Methodological approaches

After matching, we employ a difference-in-difference method with the repeated cross-section data to estimate the impact of policy on the adoption of LPG and on individual health indicators. In the following subsections, we first present in detail the econometric models used to estimate the effect of policy and the effect of other barriers identified in the theory model, and then we present the effect of policy on health indicators. For each econometric model, we highlight identification challenges and detail how we address these challenges.

5.1 Adoption effect analysis

To estimate the impact of policy on LPG adoption, we use the following specification with the household as the unit of analysis.

$$LPG_{itj} = \alpha_0 + \gamma_0 Tar_{itj} \times Post_{itj} + \gamma_1 Tar_{itj} + \gamma_2 Post_{itj} + \beta_1 H_{itj} + \beta_2 S_{itj} + \alpha_j + \epsilon_{itj} \quad (7)$$

In the above equation, LPG_{itj} is a binary variable taking value 1 if the household i in period t in the district j reported LPG as their primary cooking method. BPL takes the value 1 if the corresponding household had a BPL card, and $Post$ takes the value 1 if the household is recorded in the post-treatment period and 0 otherwise. H is the set of household control variables, and S is the set of social position of the household (caste and religion). α_j are district fixed effects and ϵ_{itj} is random error. γ_0 captures the policy impact.

Challenges to identification: It is important in this specification to rule out the impact of other policies on the adoption. Note that there were no other policies that directly incentivized the adoption of LPG in Indian households (Asharaf and Tol (2024)). However, there were programs over the same period targeted toward low-income (not specifically "BPL") households that could impact LPG use in the treated and control groups differently through channels such as wealth or income effects. To mitigate bias from such effects, we include a variety of household controls indicative of the resources of the household. Furthermore, district fixed effects account for the different time-invariant policies that were in place at the district level before the start of the PMUY scheme. We admit that, although crucial, these controls cannot themselves rule out the effect of different incentives that could have been offered over this period to the targeted and non-targeted groups.

Table 4: Household Summary Statistics by BPL Status

Variable	Matched Non BPL	BPL	All Non BPL
LPG	40.13	35.18	51.23
Livestock	49.23	57.98	45.10
Own Land	44.76	49.29	42.90
Rural	76.30	82.41	65.67
Female Head	16.56	17.54	14.87
Edu Adult	75.65	72.28	80.61
Sep Kitcen	68.75	68.65	73.49
Bank Acc	90.94	92.59	92.50
Children	0.51	0.46	0.48
Wealth Index			
1	29.17	31.26	16.36
2	23.76	26.88	17.86
3	20.32	21.39	19.22
4	16.58	14.32	21.34
5	10.17	6.15	25.21
Caste			
Scheduled Caste	22.23	21.90	17.70
Scheduled Tribe	24.54	25.89	15.48
Other Backward Class	38.59	38.15	39.35
None of them	14.29	13.48	26.78
Did not report	0.35	0.59	0.70
Relationship			
One adult	9.14	6.57	6.96
Two adults, opposite sex	37.28	28.41	31.14
Two adults, same-sex	2.46	2.48	2.26
Three+ related adults	50.73	62.25	59.18
Others	0.39	0.28	0.46
Religion			
Hindu	79.00	78.37	75.88
Muslim	8.90	9.24	10.03
Christian	7.95	8.28	8.05
Sikh	0.88	0.90	3.30
Buddhist/ Neo-Buddhist	1.46	1.40	1.43
Jain	0.10	0.05	0.22
Jewish	0.00	0.00	0.00
Parsi/Zoroastrian	0.03	0.02	0.01
No religion	0.07	0.10	0.04
Others	1.61	1.64	1.03

Notes: All the values are percentages of the households except "Children," which shows an average number of under 5 years old children per household. The matched Non-BPL column shows weighted proportion using the weights generated from the propensity score matching exercise.

Variables	matched non BPL	BPL	all non BPL	Variables	matched non BPL	BPL	all non BPL
Household controls				**Individual control**			
LPG	38.79	35.24	50.85	Hemoglobin	116.16	116.02	117.06
Livestock	54.49	61.8	49.93	BMI	21.8	21.58	22.32
Own_land	46.8	50.53	45.08	Age	29.82	30.07	30.14
Rural	76.88	81.76	66.41	Don't Smoke	91.25	90.65	92.46
Female head	15.24	16.08	14.1	Children	1.86	1.9	1.76
sep kitchen	70.1	70.52	74.81				
Wealth Index				**Education**			
1	26.89	27.61	14.33	Incomplete Primary	9.7	10.28	7.71
2	24.19	27.07	17.57	Complete Primary	3.35	3.12	3.61
3	20.89	22.52	19.49	Incomplete Secondary	42.06	42.91	43.39
4	17.23	15.76	22.16	Complete Secondary	5.2	4.68	7.1
5	10.8	7.04	26.46	Higher	10.8	7.61	16.74
				No edcuation	28.89	31.41	21.46
Caste				**Nutritional intake**			
Scheduled Caste	22.59	22.22	17.98	Milk	63.17	62.19	69.99
Scheduled Tribe	24.29	25.14	14.89	Lentils	89.48	89.44	89.7
Other Backward Class	38.84	38.72	40.6	Vegetables	88.64	88.37	87.46
None of them	13.99	13.42	25.92	Fruits	43.09	40.52	49.06
Did not report	0.29	0.51	0.6	Chicken	34.23	34.97	28.45
Religion				Egg	41.68	43.03	35.65
Hindu	78.43	78.1	75.54	Fish	32.25	32.79	28.93
Muslim	10.01	10.43	11.27				
Christian	7.5	7.57	7.34				
Sikh	0.97	0.94	3.31	No. of households	176,147	412,976	529,712
Buddhist	1.39	1.27	1.31	No. of women	243,334	593,304	756,301
Jain	0.09	0.04	0.2				
Jewish	0	0	0				
Zoroastrian	0.02	0.02	0.01				
No religion	0.06	0.08	0.03				
Others	1.54	1.53	0.98				
Relationship							
One adult	3.49	2.13	2.41				
Two adults, opposite sex	27.61	17.93	21.12				
Two adults, same-sex	2.19	2.22	1.85				
Three+ related adults	66.37	77.43	74.26				
Others	0.34	0.29	0.36				

Table 5: Summary statistics for health impact analysis

To show the robustness of the estimates toward such effects, we identified two programs, the beneficiaries of which could be identified with reasonable plausibility in the survey. 1. Housing Scheme: Aimed to provide subsidized loans to make brick and cement houses 2. Skilling Scheme: to provide training to unskilled youth from poor households for gainful employment.

To rule out the effect of the housing scheme, we run the regression specified in (7) on only those households that did not have brick and cement houses in both years (that is, the group of households that were plausibly not impacted by the housing scheme). The idea is that if it were true that the housing scheme inflates the coefficient through the wealth effect in our original regression, then the estimated effect of the LPG scheme should be lower in the households that were not affected by the housing scheme over this period. We report this result in the appendix (Table 14), and we find that the coefficient is slightly larger. This indicates the housing scheme does not inflate our original estimates.

To rule out the influence of the skilling scheme, we split the sample into two groups, where one group contains all the households that did not have any educated adult members, and the other group contains the rest. The idea is that the skilling scheme targeted uneducated youths; thus, if the scheme had any impact on the adoption of LPG through the income effect, then the two groups should display different rates of adoption of LPG. We report this result in the appendix, and we find no difference in the impact of the LPG scheme on the adoption of LPG in these two groups.

To estimate the effect of the program for the households where women had greater bargaining power and households that had cheaper access to the alternative fuel, we use the following triple difference equation, where all the notations are the same as before, except for extra interaction terms involving x_{itj}. The term x_{itj} represents variables for which we estimate heterogeneous effects given by γ_0.

$$\begin{aligned} LPG_{itj} = \alpha_0 + \gamma_0 Tar_{itj} \times Post_{itj} \times x_{itj} + \gamma_1 Tar_{itj} \times Post_{itj} + \gamma_2 Post_{itj} \times x_{itj} + \\ \gamma_3 Tar_{itj} \times x_{itj} + \gamma_4 Tar_{itj} + \gamma_5 Post_{itj} + \gamma_6 x_{itj} + \beta_1 H_{itj} + \beta_2 S_{itj} + \alpha_j + \epsilon_{itj} \end{aligned} \tag{8}$$

x_{itj} represents one variable out of five- resides in a rural area, owns livestock, owns agricultural land, has a female head or wife earns more than husband. All these are binary variables that take value 1 if the household has the associated characteristics. The interpretation of γ_0 therefore changes depending on the variable that x_{itj} represents, reflecting the impact of policy on households with that particular characteristic. For example, if x_{itj} is equal to rural γ_0 estimates the effect of policy on households in rural areas.

5.2 Health Effect analysis

To estimate the impact of policy on health, we use the following specification with the individual as the unit of analysis.

$$HI_{itj} = \alpha_0 + \gamma_0 Tar_{itj} \times Post_{itj} + \gamma_1 Tar_{itj} + \gamma_2 Post_{itj} + \beta_1 H_{itj} + \beta_2 I_{itj} + \beta_3 N_{itj} + \alpha_j + \epsilon_{itj} \quad (9)$$

In the above equation, HI_{itj} is the *Health indicator* for individual i in district j at time t. The control variables are nutritional intake N_{itj} and individual I_{itj}, and household H_{itj} controls. α_j are the district fixed effects. We use two indicators of health in two separate regressions with HI_{itj} =*Tuberculosis*, a binary indicator of having TB, and HI_{itj} = *Hemoglobin* levels in grams per liter, a continuous variable. We estimate the above model for women members in the matched sample of the household. To rule out the effect of general polices that affect health indicators, we also estimate the same model for male members, excluding the nutritional intake controls.

The advantage of using TB as a health indicator is that there is clear evidence in the medical literature that exposure to indoor air pollution is associated with increased risk of TB [5]. However, TB cases are self-reported and subject to systematic errors if there is a lack of awareness about the disease. This can bias our estimate. Thus, we also look at another health indicator hemoglobin levels of individuals. These are measures by professionals in the survey, and hence have less probability of systematic errors. However, the disadvantage is that there is no clear evidence in the medical literature about the relation between indoor air pollution (IAP) and hemoglobin levels.

The biological argument is that exposure to carbon monoxide leads to a reaction with hemoglobin that reduces the oxygen-carrying capacity of red blood cells. This triggers a compensatory mechanism in the body to produce more red blood cells in order to make up for the reduced efficiency (Aitchison and Russell (1988)). However, exposure to particulate matter can also cause inflammation that can reduce the hemoglobin levels directly (Odo et al. (2023)). Some empirical studies that examine the link between hemoglobin levels and indoor air pollution (Deng et al. (2024), Neufeld et al. (2004), Kwag et al. (2021), Honda et al. (2017)) point in different directions and are not conclusive. The fundamental issue in these studies is that individuals using dirty cooking fuels also lack other resources that affect their overall health status. Thus, to establish the relation between the use of LPG and hemoglobin, we use a non-linear [6] IV approach suggested by Wooldridge (2010). We find that non-LPG users maintain higher hemoglobin levels indicative of the presence of a compensatory trigger mechanism associated with increased exposure to carbon monoxide.

Challenges to identification: Here also it is important to rule out the other factors that could have been operational at the same time as the policy and influenced the health indicators. Estimating impact for male members of the household is very useful towards this

[5] See Obore et al. (2020) for a systematic review and meta-analysis.

[6] We use a non-linear IV approach because the endogenous variable LPG is binary. The method and results are reported in the Appendix

end. If we do not find an impact for male members, then it rules out many confounding factors that could have differently affected health indicators for targeted and non-targeted household health. For example, general cleanliness campaigns in poor neighborhoods would have influenced the health indicators for both males and females. Thus, an absence of effect for male rules out this and other similar exogenous effects.

Table 6: Policy Impact on Adoption of LPG

	Dependent variable: LPG		
VARIABLES	(1)	(2)	(3)
Tar × Post	0.026***	0.023***	0.023***
	(0.003)	(0.003)	(0.003)
Tar	-0.074***	-0.015***	-0.015***
	(0.003)	(0.002)	(0.002)
Post	0.179***	0.176***	0.176***
	(0.005)	(0.004)	(0.004)
Constant	0.323***	0.070***	0.073***
	(0.003)	(0.021)	(0.021)
Observations	730,857	730,857	730,857
R-squared	0.278	0.548	0.549
Household controls		Yes	Yes
Caste and Religion			Yes

Standard errors clustered at district level in parentheses
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

6 Empirical results

In this section, we present the empirical results of our estimates. We start by reporting the policy's effect on the adoption of LPG. We then present the important results on the role of women's agency on the effectiveness of the policy. We also present the role of the cost of alternative fuels on the effect of policy. Further, in the second section, we report the policy's impact on health.

6.1 LPG adoption

Impact of Policy

Table 6 reports the regression results for equation 7. Each column reports results with an increasing set of controls. The key coefficient is reported against $Tar \times Post$ row. We see that the coefficient is around 0.023 and is significant at a 1% significance level, with standard errors clustered at the district level. This indicates that the LPG users in targeted groups saw an increase of about 2.3 percentage points, more than the control group over the policy period. In the pre-policy period, 22% of the households had LPG connections. Thus, the policy led to just over 10% more growth in LPG usage for the BPL (targeted) group. It is important to highlight one caveat here - the question asked in the survey was related to "primary cooking fuel". Thus, even though some households may have started using LPG for preparing only some meals, they would not report LPG as their primary cooking fuel. Thus, the estimate presented here is likely a lower bound on the actual impact of the policy. Note that applying the Inverse probability weight approach, we obtain similar results as illustrated in the appendix table 12.

Table 7: heterogeneous Policy impact: Intra-household bargaining

	Dependent Variable: LPG	
	(4)	(5)
VARIABLES	x=Female head	x=Earn.more
Tar × Post × x	0.018**	0.065**
	(0.007)	(0.030)
Tar × Post	0.020***	0.020***
	(0.003)	(0.007)
Tar × x	-0.008	-0.015
	(0.005)	(0.022)
Post × x	0.009	-0.038
	(0.007)	(0.030)
Tar	-0.014***	-0.011**
	(0.002)	(0.005)
Post	0.175***	0.177***
	(0.004)	(0.007)
x	0.004	0.024
	(0.005)	(0.022)
Constant	0.075***	0.112*
	(0.021)	(0.068)
Observations	730,857	100,713
R-squared	0.549	0.548
Household controls	Yes	Yes
Caste and Religion	Yes	Yes

Standard errors clustered at district level
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

Intra-household bargaining

The regression results associated with the intra-household barriers are reported in Table 7 [7]. We use two different measures to capture the intra-household bargaining power of women in the household. The first column uses the gender of the head of the household member, and the second column uses the information on who, between husband and wife, earns more in the household. Note that the second column has fewer observations because information

[7] In the Table, we present regression results for equation 8. For brevity, we only display the coefficients for the interaction-related terms; controls are included in all the regressions.

Table 8: Heterogeneous policy impact: Cost of alternative

	Dependent Variable: LPG		
	(1)	(2)	(3)
VARIABLES	x=Livestock	x=Own land	x=Rural
Tar × Post × x	-0.048***	-0.019***	-0.032***
	(0.006)	(0.006)	(0.007)
Tar × Post	0.050***	0.032***	0.046***
	(0.004)	(0.004)	(0.006)
Tar × x	0.032***	0.010***	0.018***
	(0.003)	(0.004)	(0.005)
Post × x	0.006	-0.000	0.061***
	(0.006)	(0.006)	(0.008)
Tar	-0.033***	-0.020***	-0.028***
	(0.003)	(0.002)	(0.004)
Post	0.174***	0.176***	0.130***
	(0.005)	(0.005)	(0.007)
x	-0.090***	-0.028***	-0.140***
	(0.004)	(0.004)	(0.006)
Constant	0.077***	0.073***	0.099***
	(0.021)	(0.021)	(0.021)
Observations	730,857	730,857	730,857
R-squared	0.549	0.549	0.549
Household controls	Yes	Yes	Yes
Caste and Religion	Yes	Yes	Yes

Standard errors clustered at district level in parentheses
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

related to earnings was not available for all the households. Both indicator suggests that women would have more bargaining power in the household where the household head is a woman or the wife earns more than the husband. In both cases, we see that the policy had a higher impact on the households where women have higher bargaining power. These findings are consistent with the theoretical model of intra-household decision making that we presented earlier.

Cost of alternative

The regression results associated with the cheaper availability of alternatives are presented in Table 8. Different columns associate with different indicators of cheaper alternatives. The first column reports the impact on the household that had livestock. The households with livestock have cheaper and easier access to biofuels such as cow dung. This is a cheaper alternative to LPG. This means that households with livestock are less likely to switch to LPG after the policy relative to households without livestock. The first row reports this difference in the impact of the policy for the household with livestock compared to the household without livestock. We see in the table that the household with livestock had a negligible policy effect of .2 percentage points, which is 4.8 percentage points below the effect for households without livestock. Similarly, households with own agricultural land (column 2) will have cheaper access to wood and crop waste to use as cooking fuel. We can see a significantly lower impact for such a household by 1.9 percentage point. Finally, in the last column, access to substitute is generally easier in rural areas because of more farms and agricultural lands, and consequently, we see lower impact in such areas [8].

Table 9: Impact of policy on hemoglobin levels

	Female Haemoglobin				Male Haemoglobin	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Post × Tar	-0.209*	-0.232*	-0.232*	-0.233*	0.014	0.070
	(0.126)	(0.127)	(0.127)	(0.127)	(0.336)	(0.332)
Post	-1.087***	-1.126***	-1.372***	-1.420***	-0.393	-0.731**
	(0.215)	(0.214)	(0.215)	(0.214)	(0.362)	(0.359)
Tar	0.005	0.142	0.177*	0.177*	-0.567**	-0.123
	(0.091)	(0.092)	(0.092)	(0.092)	(0.233)	(0.231)
Constant	115.960***	113.881***	106.745***	106.238***	143.336***	133.456***
	(0.135)	(1.079)	(1.105)	(1.115)	(0.220)	(3.179)
Observations	759,951	759,951	759,078	759,078	119,119	118,899
R-squared	0.069	0.074	0.079	0.079	0.084	0.105
Household controls		Yes	Yes	Yes		Yes
Individual controls			Yes	Yes		Yes
Nutritional Intake				Yes		

Standard error clustered at district level in parentheses.
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

6.2 Health impact

Policy Impact on Health

Table 9 and Table 10 report the regression results for equation 9 with hemoglobin and TB, respectively, as health indicators. For ease of presentation, we report only the γ coefficients. The first four columns after the variable columns report regression results with different sets of controls. We can see that the hemoglobin levels were reduced in the targeted group compared to the control group, indicating an increase in oxygen-carrying capacity of hemoglobin. Also, there was a higher decrease in the number of TB cases for the targeted groups compared to the control group. The last two columns of the table report the same regression for male members of the households[9]. We can see that there was no significant change in the hemoglobin levels and TB cases of male household members. This is consistent with the fact

[8] This is in contrast with the finding of Asharaf and Tol (2024), where they find positive impact in rural areas.

[9] Number of observations is smaller for males because fewer households were selected for male hemoglobin records as the survey primarily focuses on women's health.

Table 10: Impact of policy on TB cases

VARIABLES	Female TB cases (1)	(2)	(3)	(4)	Male TB cases (5)	(6)
Post × Tar	-0.00056*	-0.00055*	-0.00059*	-0.00059*	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Post	-0.00028	-0.00031	-0.00015	-0.00016	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Tar	0.00068***	0.00067***	0.00061***	0.00061***	0.001**	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Constant	0.00381***	0.00349***	0.00430***	0.00426***	0.004***	0.008***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.003)
Observations	835,541	835,541	815,898	815,898	142,526	120,300
R-squared	0.002	0.003	0.003	0.003	0.008	0.012
Household controls		Yes	Yes	Yes		Yes
Individual controls			Yes	Yes		Yes
Nutritional Intake				Yes		

Standard error clustered at district level in parentheses.
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

that women spend more time cooking than men in Indian households thus, are more likely to experience the positive impact of LPG adoption. Note that also, in this case, applying the approach by Inverse probability weights, we obtain similar results as illustrated in the appendix table 13.

7 Discussion

The use of LPG as primary fuel was 22% in 2015 for the target group that increase to 44% in 2019. 2.3 percentage point of this 22 percentage point increase (or 10% of the increase) can be attributed to the policy. This increase, although modest, is interesting because it reflects a change in cooking behavior where LPG use not only increased but became the primary mode of cooking in these households. It is also important because, as we know from the literature (Hanna et al. (2016)) how difficult it is to initiate a change in the cooking behavior of households. The presence of intra-household barriers makes the transition even harder. As we show that the impact of the policy was double for households where a woman was head and four times for the households where the wife earned more than the husband. Similarly, the availability of cheaper alternatives makes the transition difficult. For example, we find that the policy did not lead to a switch to LPG as the primary fuel for almost no households that own livestock.

With regard to health, we find a statistically significant positive impact. Tuberculosis (TB) cases among women in the treatment group decreased from 0. 3% to 0. 2% after the intervention, approximately half of this reduction attributed to the policy. This corresponds to roughly 350,000 fewer TB cases among women, and, assuming a 12% TB fatality rate, this translates to an estimated 42,000 lives saved as a direct consequence of the policy. The economic implications are also pronounced. Considering an average cost of TB treatment of 300 euros per case, PMUY has generated estimated healthcare cost savings of approximately 105 million euros, equivalent to 12% of the program's implementation cost from TB prevention cases alone. The impact on TB may seem slightly exaggerated compared to only 2.3 percentage point increase in LPG adoption. However, we again point to the fact that this coefficient only measures the change in LPG as primary fuel. A rudimentary calculation based on hemoglobin and LPG use suggests that LPG use increased by about 9% in the group exposed to treatment.[10]

8 Conclusion

This paper empirically demonstrates that Pradhan Mantri Ujjwala Yojana (PMUY) has had a meaningful, although contextually nuanced, effect on domestic LPG adoption and women's health in India. Using a large nationally representative data set and employing robust econometric techniques, we show that PMUY led to a moderate increase in LPG uptake among poor households - in. The impact of the policy was significantly higher in households where women possessed greater decision-making power. However, the impact of the program was suppressed in rural areas and among households with easy access to traditional fuels such as livestock or agricultural land, validating the importance of the local context and the availability of alternative fuels. This heterogeneity in impact across groups confirms that household bargaining dynamics and external access to alternative fuels are critical within such public policies.

In terms of health outcomes, women in the treatment group saw measurable reductions in tuberculosis incidences and beneficial changes in hemoglobin levels, while male counterparts did not, highlighting both the gendered health burden of dirty fuels and the benefits accrued from clean energy transitions. Importantly, the reduction in tuberculosis prevalence among women not only signifies an immediate health benefit, but also translates into substantial cost savings for both households and the broader healthcare system. These savings compound health gains and provide a compelling case for continued investment in policies such as PMUY.

The findings advocate for the need to pair subsidy-driven clean fuel interventions with parallel efforts to empower women within households and address localized economic and

[10] In appendix we see that LPG use lead to 1.8 grams per liter decrease in hemoglobin, in the difference in difference analysis we find that the targeted group had hemoglobin decrease of 0.2 grams per liters which is about 9% of 1.8.

cultural barriers to fuel transition. In addition, substantial projected healthcare savings and reduced mortality rates associated with improved indoor air conditions lend strong support to the continued expansion and refinement of programs such as PMUY, both within India and in analogous settings worldwide.

References

Afridi, F., Barnwal, P. and Sarkar, S. (2024), Timing the transfer: Liquidity constraints and the transition to clean fuels, Technical report, IZA Discussion Papers.

Afridi, F., Debnath, S. and Somanathan, E. (2021), 'A breath of fresh air: Raising awareness for clean fuel adoption', *Journal of Development Economics* **151**, 102674.

Aitchison, R. and Russell, N. (1988), 'Smoking-a major cause of polycythaemia', *Journal of the Royal Society of Medicine* **81**(2), 89–91.

Alkire, S. and Seth, S. (2013), 'Identifying bpl households: A comparison of methods', *Economic and Political Weekly* pp. 49–57.

Asharaf, N. and Tol, R. S. (2024), 'The impact of pradhan mantri ujjwala yojana on indian households', *arXiv preprint arXiv:2403.17112* .

Bailis, R., Ezzati, M. and Kammen, D. M. (2005), 'Mortality and greenhouse gas impacts of biomass and petroleum energy futures in africa', *Science* **308**(5718), 98–103.

Bourguignon, F., Browning, M. and Chiappori, P.-A. (2009), 'Efficient intra-household allocations and distribution factors: Implications and identification', *The Review of Economic Studies* **76**(2), 503–528.

Browning, M. and Chiappori, P.-A. (1998), 'Efficient intra-household allocations: A general characterization and empirical tests', *Econometrica* pp. 1241–1278.

CAG (2019), Report of the comptroller and auditor general of india on pradhan mantri ujjawal yojana, Technical report, Ministry of Petroleum and Natural Gas, Government of India.

Calzada, J. and Sanz, A. (2018), 'Universal access to clean cookstoves: Evaluation of a public program in peru', *Energy policy* **118**, 559–572.

Chatterjee, M. (1990), *Indian women, health, and productivity*, Vol. 442, World Bank Publications.

Chiappori, P.-A. and Ekeland, I. (2006), 'The micro economics of group behavior: General characterization', *Journal of Economic Theory* **130**(1), 1–26.

Chiappori, P.-A. and Ekeland, I. (2009), 'The microeconomics of efficient group behavior: Identification 1', *Econometrica* **77**(3), 763–799.

Chiappori, P.-A. and Mazzocco, M. (2017), 'Static and intertemporal household decisions', *Journal of Economic Literature* **55**(3), 985–1045.

Dabadge, A. (2018), 'What has the pradhan mantri ujjwala yojana achieved so far?'.

Deng, Y., Steenland, K., Sinharoy, S. S., Peel, J. L., Ye, W., Pillarisetti, A., Eick, S. M., Chang, H. H., Wang, J., Chen, Y. et al. (2024), 'Association of household air pollution exposure and anemia among pregnant women: Analysis of baseline data from'household air pollution intervention network (hapin)'trial', *Environment International* **190**, 108815.

Duflo, E. (2012), 'Women empowerment and economic development', *Journal of Economic literature* **50**(4), 1051–1079.

Gill-Wiehl, A., Brown, T. and Smith, K. (2022), 'The need to prioritize consumption: a difference-in-differences approach to analyze the total effect of india's below-the-poverty-line policies on lpg use', *Energy Policy* **164**, 112915.

Gould, C. F., Bejarano, M. L., De La Cuesta, B., Jack, D. W., Schlesinger, S. B., Valarezo, A. and Burke, M. (2023), 'Climate and health benefits of a transition from gas to electric cooking', *Proceedings of the National Academy of Sciences* **120**(34), e2301061120.

Hanna, R., Duflo, E. and Greenstone, M. (2016), 'Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves', *American Economic Journal: Economic Policy* **8**(1), 80–114.

Honda, T., Pun, V. C., Manjourides, J. and Suh, H. (2017), 'Anemia prevalence and hemoglobin levels are associated with long-term exposure to air pollution in an older population', *Environment international* **101**, 125–132.

IEA (2015), 'World energy outlook', *International Energy Association* **Paris**.

Klasen, S. and Wink, C. (2002), 'A turning point in gender bias in mortality? an update on the number of missing women', *Population and Development Review* **28**(2), 285–312.

Kwag, Y., Ye, S., Oh, J., Lee, D.-W., Yang, W., Kim, Y. and Ha, E. (2021), 'Direct and indirect effects of indoor particulate matter on blood indicators related to anemia', *International journal of environmental research and public health* **18**(24), 12890.

Lindgren, S. A. (2020), 'Clean cooking for all? a critical review of behavior, stakeholder engagement, and adoption for the global diffusion of improved cookstoves', *Energy Research & Social Science* **68**, 101539.

Malenica, M., Prnjavorac, B., Bego, T., Dujic, T., Semiz, S., Skrbo, S., Gusic, A., Hadzic, A. and Causevic, A. (2017), 'Effect of cigarette smoking on haematological parameters in healthy population', *Medical Archives* **71**(2), 132.

MoPN (2023), Features of pmuy, Technical report, Ministry of Petroleum and Natural Gas, Government of India.

Muralidharan, K. and Prakash, N. (2017), 'Cycling to school: Increasing secondary school enrollment for girls in india', *American Economic Journal: Applied Economics* **9**(3), 321–350.

Nandwani, B. and Jain, M. (2024), 'Access to clean cooking fuel and women outcomes', *Indira Gandhi Institute of Development Research, Mumbai, Working Paper* .

Neufeld, L. M., Haas, J. D., Ruel, M. T., Grajeda, R. and Naeher, L. P. (2004), 'Smoky indoor cooking fires are associated with elevated hemoglobin concentration in iron-deficient women', *Revista Panamericana de Salud Pública* **15**(2), 110–118.

NFHS (2017), 'National family health survey (nfhs-4) 2015-16', *International Institute for Population Sciences (IIPS) and ICF* **Mumbai, India**.

NSSO (2013), Key indicators of household consumer expenditure in india, Technical report, National Sample Survey Office, Government of India.

Obore, N., Kawuki, J., Guan, J., Papabathini, S. and Wang, L. (2020), 'Association between indoor air pollution, tobacco smoke and tuberculosis: an updated systematic review and meta-analysis', *Public health* **187**, 24–35.

Odo, D. B., Yang, I. A., Dey, S., Hammer, M. S., van Donkelaar, A., Martin, R. V., Dong, G.-H., Yang, B.-Y., Hystad, P. and Knibbs, L. D. (2023), 'A cross-sectional analysis of ambient fine particulate matter (pm2. 5) exposure and haemoglobin levels in children aged under 5 years living in 36 countries', *Environmental Research* **227**, 115734.

Ranjan, R. and Singh, S. (2020), 'Household cooking fuel patterns in rural india: Pre-and post-pradhan mantri ujjwala yojana', *Indian Journal of Human Development* **14**(3), 518–526.

Sant'Anna, P. H. and Zhao, J. (2020), 'Doubly robust difference-in-differences estimators', *Journal of econometrics* **219**(1), 101–122.

Sen, A. (2017), More than 100 million women are missing, *in* 'Gender and Justice', Routledge, pp. 219–222.

Swain, S. S. and Mishra, P. (2020), 'Determinants of adoption of cleaner cooking energy: Experience of the pradhan mantri ujjwala yojana in rural odisha, india', *Journal of Cleaner Production* **248**, 119223.

Verma, A. P. and Imelda (2023), 'Clean energy access: gender disparity, health and labour supply', *The Economic Journal* **133**(650), 845–871.

WHO (2014), *WHO guidelines for indoor air quality: household fuel combustion*, World Health Organization.

Wooldridge, J. M. (2010), *Econometric analysis of cross section and panel data*, MIT press.

9 Appendix

9.1 LPG use and hemoglobin levels

A few studies that look at the link between hemoglobin levels and indoor air pollution (Deng et al. (2024), Neufeld et al. (2004), Kwag et al. (2021)) are either inconclusive or find a positive correlation between the two. Usually, higher hemoglobin levels reflect better health. However, exposure to toxic gases and pollutants can also lead to higher hemoglobin concentration as the body's compensatory mechanism triggers. The idea is that pollutants like carbon monoxide bind with red blood cells (RBCs), forming carboxyhemoglobin with no oxygen-carrying capacity. The reaction also leads to a reduction in oxygen delivery capacity of (RBCs). To compensate for decreased oxygen-delivering capacity, the body maintains a higher hemoglobin concentration Aitchison and Russell (1988).

Based on the above discussion, τ should be negative in the following model: $hemo_i$ is the hemoglobin level of individual i, LPG_i takes value 1 if the individual is in the household using LPG and 0 otherwise. X_i is the set of control variables and μ_i is random error. However, in this model, LPG and μ could be correlated, leading to inconsistent OLS estimates of τ if the individuals who use LPG are systematically different from those who decide not to use LPG.

$$hemo_i = \delta + \tau LPG_i + \beta X_i + \mu_i \tag{10}$$

Thus, to estimate the above model, we employ a non-linear IV approach, where we use a binary variable BPL_i indicating whether the individual belonged to the BPL household as the instrumental variable for LPG use. The idea is that (after matching) the households in BPL groups are more likely to use LPG as their cooking fuel because of no installation cost. However, in the matched sample, being assigned to a BPL household shall not influence the hemoglobin levels after controlling for nutritional intake (N_i) and individual (I_i) and household (H_i) controls. We use matched cross-section data from the post-policy period for this exercise. To exploit the binary nature of the endogenous variable (LPG) in the model, we use the non-linear IV method described by Wooldridge. The procedure for the method is as follows and adapted from Wooldridge (2010).

1. Estimate the binary response model $Pr(LPG = 1|N, I, H, BPL) = LPG(N, H, I, BPL)$ reported in the lower right corner of table 10
2. IV generation: Obtain the fitted probabilities $L\hat{P}G$.
3. Estimate equation 10 by IV using $L\hat{P}G$ as instrument for LPG. The first stage of this regression is reported in the lower left corner of table 10.

LPG and hemoglobin empirical relation

Table 11 reports the estimation results of model 5 using OLS and an IV approach. The econometric results of all stages of the non-linear IV method are reported in Appendix ...... [11]. Notice that the observations used are about half of those employed in the previous regressions. This is because, for the estimation of model 5, we use only cross-section data from the post-policy period for this analysis. Indeed, the policy variable used as an instrument is only valid for the second analysis period. The results obtained with the non-linear IV estimation are reported in column 2 and indicate that the LPG users have lower hemoglobin than non-LPG users by about 1.5% at a 1% level of significance. We also provide simple OLS estimation. Although OLS also has a negative coefficient, as reported in the appendix, the sign of the controls varies depending on the set of controls used. This highlights the potential endogeneity issue associated with estimating equation 10. This result confirms that there is a negative causal relation between the adoption of LPG and hemoglobin levels. Therefore it supports the use of an approach based on differences in difference method to evaluate the direct impact of policy promoting the use of clean fuel on a health outcome.

Table 11: Non-Linear IV model

VARIABLES	OLS hemo	IV hemo
lpg	-0.136	-1.795***
	(0.098)	(0.416)
Constant	106.241***	106.601***
	(1.454)	(1.463)
Observations	395,566	395,298
R-squared	0.095	0.094
Household controls	Yes	Yes
Individual controls	Yes	Yes
Nutritional Intake	Yes	Yes

Standard error clustered at district level in parentheses.
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

[11] The results of the two regressions used in the non-linear IV confirm the validity of the instruments

First Stage	
VARIABLES	lpg
Predicted LPG	0.969***
	(0.010)
Constant	-0.005
	(0.038)
Observations	395,298
R-squared	0.548
Household controls	Yes
Individual controls	Yes
Nutritional Intake	Yes

Standard errors in parentheses
District Fixed effects included
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Logit regression for IV generator	
VARIABLES	LPG
bpl	0.058***
	(0.011)
Constant	-1.645***
	(0.275)
Observations	401,310
Household controls	Yes
Individual controls	Yes
Nutritional Intake	Yes

Standard errors in parentheses
District Fixed effects included
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

First Stage (Exclusion)	**Second Stage (Inclusion)**	**Third Stage (Scoring with equal weights)**
i. Motorized two/three/four wheeler/ fishing boat	i. Households without shelter	i. Households with only one room, kucha walls and kucha roof
ii. Mechanized three/four wheeler agricultural equipment	ii. Destitute/living on alms	ii. No adult member between the ages of 16 and 59
iii. Kisan credit card with credit limit of Rs. 50,000 and above	iii. Manual scavengers	iii. Female-headed households with no adult male member between 16 and 59
iv. Household with any member as a government employee	iv. Primitive tribal groups	iv. Households with a disabled member and no able-bodied adult member
v. Households with non-agricultural enterprises registered with the government	v. Legally released bonded labourers	v. Scheduled Caste/Scheduled Tribe households
vi. Any member of the family earning more than Rs. 10,000 per month		vi. Households with no literate adult above 25 years
vii. Paying income tax		vii. Landless households deriving a major part of their income from manual casual labour
viii. Paying professional tax		
ix. Three or more rooms with all rooms having pucca walls and roof		
x. Own a refrigerator		
xi. Own a landline phone		
xii. Own 2.5 acres or more of irrigated land with at least one piece of irrigation equipment		
xiii. Five acres or more of irrigated land for two or more crop seasons		
xiv. Owning at least 7.5 acres of land or more with at least one piece of irrigation equipment		

Figure 1: BPL criteria SECC 2011

9.2 Robustness Check

Table 12: Robustness Check: Policy Impact on Adoption of LPG

	Dependent variable: LPG		
VARIABLES	(1)	(2)	(3)
Tar × Post	0.043***	0.032***	0.032***
	(0.005)	(0.004)	(0.004)
Tar	-0.077***	-0.023***	-0.023***
	(0.004)	(0.002)	(0.002)
Post	0.163***	0.167***	0.167***
	(0.004)	(0.004)	(0.004)
Constant	0.278***	0.028	0.031
	(0.004)	(0.019)	(0.019)
Observations	1,117,358	1,117,358	1,117,358
R-squared	0.236	0.532	0.532
Household controls		Yes	Yes
Caste and Religion			Yes

Standard errors clustered at district level in parentheses
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Robustness Check: Impact of policy on hemoglobin levels

	Female Hemoglobin levels			
VARIABLES	(1)	(2)	(3)	(4)
Post× BPL	-0.223	-0.284**	-0.283**	-0.278**
	(0.143)	(0.142)	(0.142)	(0.140)
Post	-1.046***	-1.074***	-1.325***	-1.377***
	(0.209)	(0.210)	(0.209)	(0.206)
BPL	0.046	0.191**	0.224**	0.224**
	(0.096)	(0.095)	(0.095)	(0.095)
Constant	116.104***	113.733***	106.511***	106.069***
	(0.149)	(0.879)	(0.895)	(0.907)
Observations	1,253,436	1,253,436	1,251,945	1,251,945
R-squared	0.069	0.073	0.078	0.078
Household controls		Yes	Yes	Yes
Individual controls			Yes	Yes
Nutritional Intake				Yes

Standard error clustered at district level in parentheses.
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Policy Impact on Adoption of LPG

	Dependent variable: LPG		
VARIABLES	No Housing	No Skill	Some skill
Tar × Post	0.037***	0.026***	0.023***
	(0.004)	(0.005)	(0.003)
Tar	-0.024***	-0.008***	-0.018***
	(0.003)	(0.003)	(0.002)
Post	0.181***	0.157***	0.183***
	(0.005)	(0.006)	(0.004)
Constant	0.110***	0.008	0.107***
	(0.025)	(0.028)	(0.030)
Observations	462,940	195,009	535,848
R-squared	0.525	0.522	0.541
Household controls		Yes	Yes
Caste and Religion			Yes

Standard errors clustered at district level in parentheses
District Fixed effects included
*** p<0.01, ** p<0.05, * p<0.1