Impact of Pradhan Mantri Ujjwala Yojana (PMUY) on expenditure of power and fuel among Indian households

ECON F342 : Applied Econometrics

Under the Supervision of

Dr. Rishi Kumar

Submitted By: **Group 3**

NAME	ID
AKSHAT AGRAWAL	2022B3A70579H
AMOL AGRAWAL	2022B3A70612H
RISHAV JAIN	2022B3A70684H
VAISHNAVI	2022B3A10783H
CHAARVI MALHOTRA	2022B3AA1340H
OJASVI MURUGAREDDY	2022B3A30778H
SHAURYA VEMALI	2022B3A70448H
AADITYARAJ MOHTA	2022B3A71384H
SARTHAK CHADHA	2022B3A31356H
DIVYANSHU TOMAR	2022B3A71380H

ABSTRACT

Clean and affordable energy plays a key role in improving health, reducing time spent on fuel collection, and promoting sustainable development. In India, many low-income households continue to rely on traditional fuels like firewood and kerosene, which are harmful to both health and the environment. To address this issue, the Government of India launched the *Pradhan Mantri Ujjwala Yojana (PMUY)* in 2016. The scheme provides free liquefied petroleum gas (LPG) connections to women in below-poverty-line (BPL) households, aiming to promote the use of cleaner fuels for cooking and reduce dependence on traditional sources.This study aims to assess the impact of PMUY on household-level expenditure on power and fuel. Using panel data from the *Consumer Pyramids Household Survey (CPHS)* conducted by the Centre for Monitoring Indian Economy (CMIE), we compare household spending patterns before and after the introduction of PMUY. Specifically, we use data from the first quarter (January–April) of 2015 as the pre-policy period and the same quarter in 2017 as the post-policy period. We apply a difference-in-differences approach framework to estimate the causal effect of the policy.

The treatment group consists of BPL households with at least one female member, which aligns with the scheme's eligibility criteria. The main outcome variable is the household expenditure on power and fuel. The analysis controls for various household characteristics such as rural or urban location, household size, average age of members, housing quality, access to formal banking, and demographic features of the household head. By conducting this analysis, we aim to understand whether PMUY has contributed to lowering energy-related expenses for poor households and supported their transition to cleaner energy sources. The findings of this study will provide useful insights for policymakers designing large-scale energy access programs in developing countries.

TABLE OF CONTENTS

ABSTRACT **2**
TABLE OF CONTENTS **3**
INTRODUCTION **5**
LITERATURE REVIEW **7**
PAPER 1: "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" 7
PAPER 2: "The impact of Pradhan Mantri Ujjwala Yojana on Indian households" 8
PAPER 3: "Women Empowerment through Pradhan Mantri Ujjwala Yojana (PMUY) Scheme in Rajasthan: A Study on Rural Households in Selected Region" 9
PAPER 4: "Household energy access and expenditure in developing countries: Evidence from India, 1987–2010" 10
PAPER 5: "Determinants of adoption of cleaner cooking energy: Experience of the Pradhan Mantri Ujjwala Yojana in rural Odisha, India" 11
DATA AND METHODOLOGY **13**
DESCRIPTION OF THE DATASET 13
MODEL 14
VARIABLES AND JUSTIFICATION OF CHOICE 14
METHODOLOGY **21**
Data Cleaning and Preparation 21
Model Selection 26
Tests Performed **30**
Hausman Test 30
VIF Test 30
Heteroskedasticity Test 30
Normality Test 30
Time Fixed Effects Test 31
Results and Diagnostics **32**
Summary Statistics 32
Regression Results 33
Significant Variables 34
Interpretation of Parameters: 34
Hausman Test 36
Heteroskedasticity Test 37
VIF Test 38
Normality Test 39
Test for Time Fixed Effects 40
Event Study **41**
Event Study Plot 41
Conclusion **42**
References **44**

INTRODUCTION

India's energy landscape has undergone significant transformation over the past decade, particularly in the domain of household-level access to clean no cooking fuels. Despite rapid economic growth and infrastructure development, millions of households especially those in rural and economically weaker sections have historically depended on biomass, kerosene, and other non-clean fuels for daily cooking and energy needs. These fuels not only contribute to environmental degradation but also expose household members, primarily women, to hazardous indoor air pollution and related health risks. Recognizing the urgent need to shift households toward cleaner and safer energy alternatives, the Government of India launched the *Pradhan Mantri Ujjwala Yojana (PMUY)* in 2016. The scheme aimed to provide free LPG connections to women from below-poverty-line (BPL) households, thereby encouraging a transition to cleaner fuel usage.

While PMUY was an ambitious step toward achieving universal energy access, the broader implications of the policy on household expenditure patterns, particularly spending on power and fuel, remain insufficiently explored. Although access to LPG was expanded under PMUY, questions persist around the extent to which it led to changes in consumption behavior, reduced dependence on traditional fuels, or altered the household energy budget. A deeper understanding of these effects is essential for evaluating whether the program delivered sustained benefits beyond initial adoption.

This study seeks to examine how household energy expenditures evolved before and after the implementation of PMUY, using detailed panel data from the *Consumer Pyramids Household Survey (CPHS)* collected by the Centre for Monitoring Indian Economy (CMIE). The analysis focuses on two key time points 2015 (prior to the launch of PMUY) and 2017 (after its introduction) allowing for an evaluation of expenditure changes over time. The primary objective is to assess whether households eligible for the scheme defined here as BPL households with a female member experienced a significant shift in their adjusted monthly expenditure on power and fuel compared to non-eligible households.

To ensure a robust analysis, the study incorporates a wide range of household-level variables that may influence energy expenditure. These include demographic characteristics such as region type (urban or rural), weighted average age of household members, religion, caste category, and household head's gender and education. Socioeconomic indicators like household size, type of wall material, banking access, recent hospitalization of any member, and adjusted spending on food and health are also considered to control for heterogeneity in household needs and living standards.

Through a panel regression framework, this research aims to provide new evidence on the economic impact of clean energy access programs and contribute to ongoing policy discussions around energy inclusion, behavioral change, and welfare targeting in India's development context.

LITERATURE REVIEW

To understand the determinants of household energy expenditure, it is essential to examine a range of factors, including socio-demographic characteristics, household economic status, aspirations, and consumption behavior. Several studies have investigated these dimensions in the context of the Pradhan Mantri Ujjwala Yojana (PMUY), offering valuable insights into how these variables shape energy usage patterns.

PAPER 1: "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"

By Annelise Gill-Wiehl, Timothy Brown, and Kirk Smith

About the paper

This paper utilises DID methodology to investigate whether BPL policies like PMUY lead to higher LPG connection rates and consumption compared to a non-equivalent comparison group of households immediately above the poverty line. They show whether these policies pushed households to consume LPG and increase their cooking energy access tier.

Methodology

This study draws on panel data from the 2015 and 2018 rounds of the ACCESS survey, conducted across six Indian states, to examine the effects of BPL-targeted policies—specifically the Pradhan Mantri Ujjwala Yojana (PMUY)—on LPG adoption. A two-part difference-in-differences (DiD) model with household fixed effects is used to estimate both the probability of LPG use and the amount consumed. The primary outcomes include LPG usage, access to home delivery, and levels of clean cooking energy access. The main treatment variable is consistent BPL identification across both survey waves. Control variables include age, household size, total expenditure, and the presence of a female decision-maker. To ensure the results are reliable, the study also includes additional checks using methods like propensity score matching and a doubly robust DiD approach. These help strengthen the conclusions, especially given the lack of data from before the first survey round.

Key findings

Eligibility for BPL policies modestly increases the likelihood of obtaining an LPG connection and refilling large cylinders by about 8 percentage points (45–50% increase), but has no effect on small cylinder use, overall LPG consumption, tier advancement in clean cooking energy, or uptake of home delivery. BPL policies improve initial access but they do not significantly affect sustained use or broader energy transitions, indicating the need for more aggressive or complementary interventions.

PAPER 2: "The impact of Pradhan Mantri Ujjwala Yojana on Indian households"

By Nabeel Asharaf, Richard S.J. Tol

About the paper

This paper examines the effectiveness of PMUY at increasing access to clean cooking fuels in India. It explores the limited impact of existing policies on sustained LPG usage and highlights how rising energy prices and reduced subsidies have worsened energy poverty, pushing many back to traditional cooking methods. By analyzing the causal effects of PMUY, the authors find that while the policy increased initial LPG adoption, its overall effect on sustained use and clean energy transition was modest.

Methodology

To estimate the causal impact of PMUY in the absence of randomised data, the study employs an Intention to Treat approach using Below Poverty Line status as a proxy for treatment. Using repeated cross-sectional data from NFHS, the authors use a matched DID estimator, pairing treatment and control households via Propensity Score Matching to minimise bias from observable covariates. The approach relies on key assumptions such as the Conditional Mean Assumption and common support to simulate randomisation. To address potential bias from unobserved confounders, a Rosenbaum bounds sensitivity analysis is conducted. Collectively, these methods provide credible causal estimates while acknowledging the limitations of non-randomized observational data.

Key findings

The Pradhan Mantri Ujjwala Yojana led to a statistically significant 2.1 percentage point increase in LPG use among BPL households, accompanied by reduced firewood dependence, but with certain regional disparities. The long-term economic and health benefits of cleaner fuel use likely outweigh the costs. Future research could explore regional differences using datasets like ACCESS, examine time-use changes, and assess PMUY's impact on respiratory health—especially for women and children—potentially revealing broader educational and economic gains.

PAPER 3: "Women Empowerment through Pradhan Mantri Ujjwala Yojana (PMUY) Scheme in Rajasthan: A Study on Rural Households in Selected Region"

By Dr. Yaduveer Yadav, Dr. Pradeep Kumar Sharma, and Dr. Kiran Raj

About the paper

The study has five hypotheses established to test the relationship among dimensions of PMUY and socio-economic variables of women empowerment. It has 3 main objectives, i.e., to identify the relationship between PMUY and socioeconomic factors of women empowerment in the rural sector, the impact of PMUY on socio-economic factors of women empowerment in the rural sector and to propose a conceptual framework for future research to enhance the effectiveness of PMUY.

Methodology

The distinct dimensions of Pradhan Mantri Ujjwala Yojana selected as dependent determinants and socio-economic factors of women empowerment chosen as independent variables selected for research. Collection of quantitative information was done through a structured questionnaire, with 187 usable responses from the women residing in the rural sector in India. To achieve objectives mentioned in the study, various statistical tools, namely, Cronbach alpha for reliability examination, Correlation, regression analysis, and ANOVA with the help of SPSS Software, applied.

Key findings

PMUY has significantly expanded LPG access and reduced dependence on traditional cooking fuels, improving the quality of life for rural women. Many respondents reported environmental and time-saving benefits, leading to enhanced social and economic empowerment. However, consistent LPG use cannot be seen due to challenges like refill accessibility, affordability, and lack of awareness.

PAPER 4: "Household energy access and expenditure in developing countries: Evidence from India, 1987–2010"

By Meir Alkon a, S.P. Harish b, and Johannes Urpelainen

About the Paper

This study has two central objectives. First, it examines changes in the energy cost burden measured as the share of household expenditure devoted to energy across rural and urban India between 1987 and 2010. It highlights both temporal trends and geographic variation across states, addressing a gap in prior research which has not adequately captured these dynamics.

Second, it investigates the relationship between household income, access to modern energy sources (such as LPG and electricity), and the energy cost burden, with the aim of distinguishing between financial strain and willingness to pay for improved energy services.

Methodology

The analysis uses five rounds of household-level data from the National Sample Survey (NSS), covering over 578,000 rural and urban households. It includes both purchased and self-produced fuels to capture comprehensive energy spending. Regression analysis is employed to estimate the effects of income (proxied by non-energy expenditure) and access to modern fuels on energy costs, controlling for household characteristics, fuel prices, and state-level factors.

Key Findings

The study finds substantial variation in energy cost burden across states and between rural and urban areas. Rural households consistently allocated 10–15% of their monthly expenditure to energy, with this burden increasing over time due to inflation and improved access to modern fuels. Urban households experienced more stable or even declining energy burdens in some states. Notably, the greatest increases in energy burden occurred in middle-income states, rather than the poorest or richest. The findings also reveal a strong negative correlation between poverty reduction and rising energy burden, particularly in rural areas, suggesting that limited success in economic reform has contributed to persistent affordability challenges.

PAPER 5: "Determinants of adoption of cleaner cooking energy: Experience of the Pradhan Mantri Ujjwala Yojana in rural Odisha, India"

By Swadhina Shikha Swain and Pulak Mishra

About the paper

This paper examines how the beneficiaries of the PMUY have adopted LPG as a cooking energy in practice and identifies the underlying factors influencing their adoption. The paper uses experiences on use of LPG by the beneficiaries of the PMUY in rural areas of the Indian state of Odisha.

Methodology

The study employed a primary survey method, collecting data as a structured questionnaire from 106 rural households across five villages in Odisha's Puri district, aligned with PMUY's target demographic. The data was analysed using a linear probability model and logistic regression to evaluate the likelihood of LPG adoption under PMUY. The dependent variable was a binary indicator of LPG usage, while key independent variables included PMUY beneficiary status, household income, education level, caste category, proximity to LPG distributor, and prior use of traditional fuels. Robust standard errors were used to correct for heteroskedasticity, and marginal effects from the logistic model were interpreted to understand the practical significance of PMUY in influencing LPG uptake.

Key findings

Regression analysis shows that the decline in firewood use among rural households is significantly influenced by factors such as caste category, education of the household head, type of cooking fuel scheme, and access to alternatives like kerosene. General caste households, PMUY beneficiaries, and those with access to kerosene are less likely to reduce firewood use. Higher LPG subsidies are associated with greater reductions in firewood use, reinforcing the substitution effect of making LPG more affordable.

Key Insights from The Above Research Papers

In recent years, there has been growing academic interest in how households in India access, adopt and use modern energy sources—particularly in the wake of large-scale government initiatives such as the Pradhan Mantri Ujjwala Yojana (PMUY). It has become a focal point for research seeking to understand shifts in household energy behaviour and the broader implications for well-being and expenditure.

A common objective across this body of literature is to assess whether increased access to liquefied petroleum gas (LPG) under PMUY has led to lasting changes in household energy use. Most studies combine household survey data with robust econometric techniques such as difference-in-differences (DiD), propensity score matching (PSM), and logistic regression to identify causal relationships. These analyses typically control for a wide range of socio-economic and demographic factors—including household income, education, caste, size, and the gender of the decision-maker—to understand the drivers of energy transition and usage intensity.

The findings reveal a consistent trend: while PMUY has improved initial access to LPG among low-income households, sustained usage often remains limited. Factors such as high refill costs, inconsistent supply and limited awareness reduce the likelihood of households transitioning fully away from traditional biomass fuels. In many cases, LPG is used as a supplementary rather than a primary fuel, resulting in what is known as "fuel stacking."

Studies that examine household-level energy expenditure over time suggest that energy burdens—measured as the proportion of income spent on energy—have evolved unevenly across states and income groups. Particularly in rural areas and among middle-income states, energy spending has risen due to inflation, expanded access to electricity and LPG, and other structural changes. However, these increases also reflect improvements in energy access, which may bring long-term social and economic benefits, such as reduced time spent collecting firewood and better health outcomes, especially for women.

Our study builds on these insights by exploring a less examined but crucial aspect: the impact of PMUY on household expenditure specifically on power and fuel. While much of the existing work focuses on adoption patterns and transitions in fuel usage, our analysis seeks to understand whether policies like PMUY also alleviate or exacerbate the financial burden of energy consumption. In doing so, we aim to contribute to the broader conversation on energy equity and affordability in India's transition to cleaner fuels.

DATA AND METHODOLOGY

DESCRIPTION OF THE DATASET

The dataset used in this study is drawn from the Consumer Pyramids Household Survey (CPHS), conducted by the Centre for Monitoring Indian Economy (CMIE). CPHS is a nationally representative longitudinal survey that collects data from a large sample of households across India. It provides detailed information on household demographics, income, expenditure, housing conditions, and access to basic services. The survey is conducted three times a year, allowing researchers to track changes in household conditions over time.

For the purpose of this study, data from the first quarter of 2015 and the first quarter of 2017 are used. The 2015 data represent the period before the implementation of the Pradhan Mantri Ujjwala Yojana (PMUY), while the 2017 data represent the period after the scheme was introduced. By selecting households that appear in both time periods, a balanced panel is constructed for analysis.

The main variable of interest is the household's adjusted monthly expenditure on power and fuel. To explain differences in this expenditure, several household-level characteristics are included in the analysis. These include the type of area (rural or urban), the weighted average age of household members, religion, caste category, gender and education level of the household head, and whether the household head has a bank account. Additional factors considered are housing quality, as indicated by the type of wall, whether any household member was recently hospitalised, household size, and monthly spending on food and health.

This dataset is well-suited for studying the impact of government policies like PMUY because it tracks households over time. By comparing data from before and after the scheme, we can observe changes in spending and see if the impact varies by household characteristics such as housing quality, banking access, or demographics.

MODEL:

$$\ln(\text{avg\_power\_fuel}_{it}) = \beta_1 \cdot \text{avg\_tot\_inc}_{it} + \beta_2 \cdot \text{caste\_category\_num}_{it} + \beta_3 \cdot \text{edu\_coded}_{it} + \beta_4 \cdot \text{group\_is\_hospitalised\_num}_{it}$$
$$+\beta_5 \cdot \text{has\_woman\_18\_plus}_{it} + \beta_6 \cdot \text{gender\_num}_{it} + \beta_7 \cdot \text{num\_members}_{it} + \beta_8 \cdot \text{household\_weighted\_avg\_age}_{it}$$
$$+\beta_9 \cdot \ln(\text{avg\_health}_{it}) + \beta_{10} \cdot \text{post}_{it} + \beta_{11} \cdot \text{treatment}_{it} + \alpha_i + \varepsilon_{it}$$

VARIABLES AND JUSTIFICATION OF CHOICE

Dependent Variable (Y):

1. **Adjusted Expenditure on Power and Fuel (AVG_POWER_FUEL):**

- **Description :** This variable measures the household's total average monthly spending on power and fuel during the January to April quarter. It includes expenses on electricity, LPG, firewood, and other conventional or alternative energy sources used for cooking, lighting, and heating. The values are adjusted for inflation and regional price differences to ensure they can be fairly compared across years and locations. The data is drawn from the CMIE Consumer Pyramids Household Survey (CPHS), where households report their monthly expenditure in this category.

- **Justification:**
  This variable is central to evaluating the impact of the Pradhan Mantri Ujjwala Yojana (PMUY), which aims to provide clean cooking fuel (LPG) to women in Below Poverty Line (BPL) households. The scheme is designed to reduce the use of polluting traditional fuels and ease the financial burden of energy costs. By comparing adjusted household expenditure on power and fuel before and after the scheme's

rollout (2015 vs. 2017), and between households eligible and ineligible for PMUY, we can assess whether the policy led to a meaningful reduction in energy-related spending. A decline in expenditure would suggest a successful transition away from inefficient traditional fuels and improved affordability of cleaner alternatives. Therefore, this variable effectively captures the intended impact of PMUY on household energy behavior and financial well-being.

Independent Variables(X):

1. **Type of Region :**

- **Description :**"REGION_TYPE" is a categorical variable that classifies the household's geographic location as either **urban** or **rural** based on CMIE's demographic coding.

- **Justification:** The PMUY scheme specifically targeted rural India, where the penetration of clean cooking fuels like LPG was historically low. Rural households often relied on biomass, firewood, or dung cakes, which are freely available but highly polluting and inefficient. Urban households, in contrast, typically have higher baseline access to LPG and electricity. Thus, accounting for REGION_TYPE helps control for pre-existing infrastructure disparities. It ensures that differences in power and fuel expenditure across households are not simply due to urban-rural structural divides but can be more accurately attributed to policy interventions like PMUY.

2. **Weighted Average Age:**

- **Description:** The variable (household_weighted_avg_age) represents the custom-weighted average age of all members in a household. Instead of treating every individual equally, this variable assigns different importance (weights) to individuals based on their age group, which reflects their likely contribution to household energy decisions and needs.

- **Justification:** The *Weighted Average Age* variable offers a refined measure of household demographic composition by incorporating differential energy needs across age groups. Rather than assigning equal importance to all household members, this variable applies age-specific weights to reflect likely contributions to energy consumption. Working-age adults (15–59), who are generally more engaged in economic and domestic activities, are assigned the highest weight. Children (0–14) and older adults (60+) receive lower weights, acknowledging their relatively lower and distinct energy usage patterns. This weighted approach provides a more nuanced and policy-relevant indicator of household energy demand than a simple average age.

3. **RELIGION :**

- **Description :** The "RELIGION" denotes the religious affiliation of the household head (e.g., Hindu, Muslim, Christian, Sikh, Others). This variable is typically categorical and included in the CMIE data based on self-reporting during surveys.

- **Justification:** Religious beliefs and customs can influence how households cook, what kind of food they prepare, and how often they use fuel. For example, some communities may use traditional stoves more often, while others might prefer modern cooking methods like LPG. Religion can also be linked to differences in lifestyle, awareness, and access to government schemes like PMUY. By including religion as a control variable, we make sure that these cultural differences don't affect our estimate of how PMUY impacted household spending on fuel and power.

4. Caste (CASTE_CATEGORY):

- **Description:** The categorical variable (CASTE_CATEGORY) classifying households into caste groups: Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Classes (OBC), Intermediate Caste , Upper Caste and Others. This is self-declared during CMIE survey data collection.

- **Justification:**Caste remains a strong determinant of socio-economic status in India. The PMUY scheme explicitly focused on socially and economically marginalized groups, particularly SC/ST households. These groups often have limited access to clean cooking fuels and face barriers in obtaining LPG connections. Caste also correlates with literacy, income, housing conditions, and subsidy awareness. Controlling for caste ensures that the observed variation in fuel expenditure is not merely a reflection of social stratification but rather of policy-induced changes in behavior.

5. Hospitalised status (group_is_hospitalised) :

- **Description:** This is a dummy variable indicating whether any household member was hospitalized in the past quarter or reference period.

- **Justification:** Hospitalization serves as an important indicator of a household's health status. Traditional cooking fuels such as firewood or biomass are linked to higher health risks, especially respiratory illnesses, due to indoor air pollution. Households using such fuels may experience more health problems, which can lead to increased hospitalization. Including this variable helps capture the indirect health impact of fuel choices. It also reflects financial shocks that may influence spending patterns. In times of illness or hospitalization, households might adjust their budgets by cutting down on regular expenses, including fuel. Therefore, this variable helps control for both the health risks associated with traditional fuels and the economic effects of medical emergencies when analyzing changes in fuel expenditure.

6. **Education of Household Head(EDU_CODED):**

- **Description:** Captures the highest level of formal education completed by the household head. It is a categorical variable with increasing values assigned based on educational attainment. A value of 0 indicates no formal education, 1 corresponds to primary education (1st to 5th standard), 2 represents secondary education (6th to 10th standard), 3 denotes higher secondary education (11th to 12th standard), 4 indicates graduation, and 5 represents postgraduate education. This ranking allows for a structured comparison of education levels across households, which can be relevant in understanding differences in awareness, access to welfare schemes, and energy usage behavior.

- **Justification:** Education influences decision-making capacity, awareness of government schemes, and understanding of long-term benefits. Literate household heads are more likely to understand the health and time-saving benefits of LPG over biomass. They are also more likely to complete documentation, seek refills, and budget for clean fuel consistently. Hence, education is a strong predictor of both adoption and sustained use of LPG under PMUY.

7. **Gender of Household Head(GENDER):**

- **Description:**The variable “ GENDER” is a dummy variable showing whether a household is led by men or women.
- **Justification:** PMUY was designed to empower women by registering LPG connections in the name of female household members. Female-headed households may prioritize clean cooking due to direct exposure to smoke and the time burden of traditional fuel collection. In such households, the policy might have stronger uptake and usage. On the other hand, male-headed households may retain decision-making control over household spending, influencing how the benefit is used. Including gender helps explore these differential behavioral effects.

9. Size of Household (num_members):

- **Description:** This variable gives the number of members that are part of the household.
- **Justification:** Household size plays an important role in shaping energy consumption patterns. Larger households generally require more energy for daily activities such as cooking, heating, and lighting, which can lead to higher total expenditure on power and fuel. At the same time, larger families may face greater financial pressure, potentially influencing how much of their energy needs are met through clean fuels like LPG versus traditional sources. Including household size as a control variable helps account for these differences in demand and affordability, ensuring that observed changes in fuel expenditure are not simply a result of household scale, but more accurately reflect the impact of the PMUY intervention.

10. Total Income (AVG_TOT_INC) :

- **Description:** The variable represents the aggregate income of all household members, measured on a monthly basis.TOT_INC serves as a summary measure of the household's overall financial resources, capturing its ability to spend on essentials and discretionary items, including energy.

- **Justification:** Total income is a critical determinant of household energy behavior, influencing both the choice and quantity of fuel consumed. Households with higher income levels are more likely to adopt cleaner and more convenient sources like LPG and electricity, while poorer households often rely on traditional fuels due to affordability. Including TOT_INC in the regression, we control for such economic disparities and isolate the pure impact of PMUY on power and fuel expenditure.

11. Expenditure on Health (AVG_HEALTH):

- **Description:** Total monthly spending by the household on health-related needs, including medicines, doctor visits, and hospitalization (not captured under group_is_hospitalised).

- **Justification:** Health expenditure serves a dual purpose in this analysis. High health spending may limit a household's ability to afford clean fuels like LPG, showing a *constraining effect*. At the same time, poor health often caused by indoor air pollution may encourage households to adopt cleaner fuels, reflecting an *enabling effect*. Including this variable helps us capture both aspects. Additionally, since PMUY aims to improve health by reducing traditional fuel use, changes in health spending after its implementation can indirectly indicate whether the policy led to better health outcomes.

METHODOLOGY

Data Cleaning and Preparation

The dataset used in this study is drawn from the Consumer Pyramids Household Survey (CPHS), conducted by the Centre for Monitoring Indian Economy (CMIE).

This section outlines the steps followed to extract, clean, and merge data from multiple modules of the **CMIE Consumer Pyramids Household Survey (CPHS)**, specifically for the purpose of analyzing the impact of **PMUY** on **power and fuel expenditure**.

1. Data Extraction

We extracted data from the following modules of CPHS for two time periods:

- **Pre-policy period**: January to April 2015
- **Post-policy period**: January to April 2017

Modules Used:

- People of India DX
- Aspirational India DX
- Income Pyramids DX
- Consumption Pyramids DX

Time Selection:

- For **People of India DX** and **Aspirational India DX**:
  Extracted quarterly data for **January to April** for both years.
- For **Income Pyramids DX** and **Consumption Pyramids DX**:
  Extracted **monthly data** for **January, February, March, and April** of both years.
- These were later aggregated into quarterly values as we will discuss later.

2. Variable Selection

Relevant variables were selected from each dataset for further processing:

People of India DX

- HH_ID (Household ID)
- MEM_ID (Member ID)
- RESPONSE_STATUS (Accepted/rejected)
- REGION_TYPE (Rural/Urban)
- MEM_STATUS (Current member status)
- GENDER
- AGE_YRS (Age in years)
- RELATION_WITH_HOH (Relationship with head of household)
- RELIGION
- CASTE_CATEGORY
- EDU
- IS_HOSPITALISED (Hospitalization status)

Aspirational India DX

- HH_ID
- REGION_TYPE
- RESPONSE_STATUS
- SIZE_GROUP(number of members in household)

Income Pyramids DX

- HH_ID
- REGION_TYPE
- RESPONSE_STATUS
- SIZE_GROUP
- TOT_INC (Total household income)

Consumption Pyramids DX

- HH_ID
- REGION_TYPE
- RESPONSE_STATUS

- SIZE_GROUP
- ADJ_M_EXP_POWER_N_FUEL (Adjusted monthly expenditure on power and fuel)
- M_EXP_Health (Monthly health expenditure)

3. Household-Level Aggregation (People of India DX)

To obtain a single household-level observation from individual member data, the following cleaning and aggregation procedures were applied:

- **Merging Members**:
    - All individuals with the same `HH_ID` were grouped into a single household entry. The effects to be considered for other attributes are discussed below.

- **Response Status**:
    - `RESPONSE_STATUS` was marked as "Accepted" if **any member** of the household had accepted participation.

- **Head of Household Variables**:
    - `GENDER_HOH`: Gender of the individual with `RELATION_WITH_HOH = HOH` was created with using `GENDER` variable
    - `RELIGION`, `CASTE_CATEGORY`, and `EDU`: Taken from the HOH only.

- **Age**:
    - To construct a weighted average age variable that reflects household energy needs, individuals are grouped by age and assigned different weights based on their likely impact on power and fuel consumption. Working-age adults (15–59) are assigned the highest weight of 1, as they are typically more active and have greater energy requirements due to work and daily activities. Children (0–14) are assigned a lower weight of 0.5, and the elderly (60+) are assigned a weight of 0.7, as both groups generally consume less energy and food. This method helps capture how household composition influences overall energy usage more accurately than a simple average age.

- **Hospitalization Status**:

- `IS_HOSPITALISED = 1` if **any household member** was hospitalized during the quarter.

- **Education Variable Recoding**
  - The education level of the head of household was recoded into numeric categories:

Code	Education Level
0	No formal education / invalid data
1	Primary education (1st–5th std.)
2	Secondary (6th–10th std.)
3	Higher secondary (11th–12th std.)
4	Graduation
5	Post-graduation

- Finally only households with response status accepted are retained.

4. Cleaning and Processing Other Modules

Aspirational India DX

- Only households with `RESPONSE_STATUS = Accepted` were retained.
- Used `HH_ID` to merge with the household-level dataset from People of India DX.
- Type of Wall is main variable to be used from this dataset
- Convert string data like group size into numeric format from string data (8 members -> 8) .

Income Pyramids DX

- Joined using HH_ID (inner join) across months for both time periods .
- Monthly income data for **January to April** was aggregated by summing TOT_INC across months for each HH_ID.
- Only households with RESPONSE_STATUS = Accepted were included.
- **Negative values** or **"Still inapplicable" or missing values** were replaced with the **average** of the respective variable across valid entries (for Tot_INC).

Consumption Pyramids DX

- Joined using HH_ID (inner join) across months for both time periods .
- Monthly expenditure data from **January to April** was aggregated by summing:
  - ADJ_M_EXP_POWER_N_FUEL
  - M_EXP_Health
- **Negative values** or **"Still inapplicable" or missing values** were replaced with the **average** of the respective variable across valid entries (for Tot_INC).

5. Final Dataset Creation

- All cleaned and processed datasets were **inner join** on HH_ID to form two merged dataset one for the pre intervention (2015) and post intervention (2017).
- Then these were further merged into a single dataset with another column added POST .

  The new dummy variable POST is defined as :

  - POST = 0 for pre-policy households (from 2015 data)
  - POST = 1 for post-policy households (from 2017 data)

Model Selection

We use a **panel dataset of 2000 households** observed in both **pre- and post-periods**, totaling 4000 observations. The primary objective is to identify the causal impact of treatment on household energy spending (captured by the variable avg_power_fuel), while carefully controlling for unobserved characteristics that are unique to each household and do not change over time.

To begin with, we set up a **difference-in-differences (DiD) framework**. This was achieved by generating an **interaction term** between the treatment group indicator and the post-period dummy variable. Several categorical variables—such as religion, caste category, gender, and whether someone in the group was hospitalized—were encoded into numeric formats using the encode command. This encoding made them compatible with the regression commands used later.

We **established the panel structure** in Stata using:

xtset hh_id post

This command tells Stata that the data consists of repeated observations (pre and post) for the same household, with 'hh_id' being the household identifier and 'post' indicating the time period.

Log Transformation of Skewed Variables

Upon exploring the data, we noticed that certain continuous variables like avg_power_fuel (average spending on power and fuel), avg_health, and avg_tot_inc (average total income) were positively skewed. In order to reduce skewness and improve the reliability of our regression estimates, we applied a natural logarithmic transformation to each of these variables. The transformations were done as follows:

gen ln_avg_power_fuel = ln(avg_power_fuel + 1)

gen ln_avg_health = ln(avg_health + 1)

gen ln_avg_tot_inc = ln(avg_tot_inc + 1)

Adding 1 inside the logarithm ensures that observations with a value of zero do not result in undefined values. Log transformations also help interpret the results in percentage terms, making it easier to communicate findings.

Fixed Effects Estimation

To control for household-level factors that remain constant over time (such as cultural norms, unobserved preferences, or long-term household traits), we **employed a fixed effects regression model using the xtreg command with the 'fe' option:**

xtreg ln_avg_power_fuel i.caste_category_num i.edu_coded i.group_is_hospitalised_num i.has_woman_18_plus i.religion_num i.gender_num num_members household_weighted_avg_age ln_avg_health post treatment, fe

This model estimates the effect of treatment and other covariates while accounting for unchanging characteristics at the household level. It does this by effectively comparing each household to itself over time, allowing us to isolate the impact of time-varying factors like treatment.

Random Effects vs. Fixed Effects: The Hausman Test

To ensure that the fixed effects model is the appropriate choice, **we also estimated a random effects model using:**

xtreg ln_avg_power_fuel ..., re

We then **conducted a Hausman test** to compare the two approaches:

hausman fe re

The null hypothesis of the Hausman test states that the random effects model is appropriate—that is, household-level unobserved factors are not correlated with the explanatory variables. The alternative hypothesis supports the fixed effects model, suggesting that such correlations exist.

Our result showed:

Prob > chi2 = 0.0000

This extremely **low p-value led us to reject the null hypothesis**. Hence, we concluded that the **fixed effects model is more appropriate** for our data, as it avoids biased estimates that might arise under the random effects assumption.

Absorbing Fixed Effects with AREG

We also explored an alternative but related method using the areg command with the absorb() option. This approach is particularly helpful in dealing with variables like religion, which remain constant for the vast majority of households (approximately 1900 out of 2000), but change in a small subset (around 300 households).

areg ln_avg_power_fuel ..., absorb(religion_num)

In a typical fixed effects regression (using xtreg, fe), Stata does not automatically drop such variables because they do show some variation overall. However, this may lead to misleading results since the variable is effectively fixed for most of the panel units (households). **By using absorb(), we explicitly instruct Stata to treat the absorbed variable as part of the fixed effect.** This means the variation in religion is only considered for households where it actually changes, and it is excluded where it does not.

To estimate the causal impact of variables while controlling for unobserved heterogeneity at the household level, this kind of absorption becomes essential. It **ensures that our estimates are driven purely by within-household variation across time, which is the true strength of the fixed effects model.**

In our dataset, the total number of households was 2223 (slightly higher than the earlier subset), with 4446 observations spanning both pre- and post-periods. In this sample, some variables such as religion remained constant for nearly all households but showed some change in a few cases. Running a standard fixed effects regression without using absorb() does not fully account for this structure. **The absorb() command solves this issue by ensuring that variables fixed within units are handled correctly.**

In conclusion, **our methodology carefully combines the DiD framework, log transformation of skewed variables, fixed effects modeling, the Hausman test for model validation, and the strategic use of absorb() for robust estimation.** This multi-step approach provides us with more accurate, consistent, and interpretable estimates of the treatment effects in our panel data setup.

Tests Performed

Hausman Test

Purpose: a statistical test used to determine whether a Fixed Effects (FE) model or a Random Effects (RE) model is more appropriate for panel data analysis.

H0 : Random Effects is consistent and efficient

H1 : Fixed Effects is consistent and efficient.

VIF Test

Purpose : To test for multicollinearity in the model, a situation where independent variables are highly correlated with each other. Multicollinearity can inflate standard errors, make coefficients unstable, and weaken the statistical power of the model.

Theory : VIF= 1/(1- R^2)
A VIF value <10 is acceptable for a regression model, whereas a value of VIF < 5 is considered to be very good, implying very low multicollinearity amongst the variables.

Heteroskedasticity Test

Wald Test for groupwise heteroskedasticity.

Purpose : It is used to detect heteroskedasticity in fixed effect regression models

H0 : Homoscedasticity (constant variance of error terms)
H1 : Heteroskedasticity (inconsistent variance of error terms)

Normality Test

Shapiro - Wilk Test

Purpose : to check whether a distribution is normal or not, it measures how close your sample distribution is to a normal distribution.

H0: The data is normally distributed.
H1: The data is not normally distributed.

Time Fixed Effects Test

Purpose : to control for time invariant factors that may affect the outcome of the variable, these are usually unobserved characteristics.

Theory : This test creates a dummy variable for each time period in the data set. This dummy variable captures the effect of time on the outcome variable and controls for any time-invariant factors that may be present.

Results and Diagnostics

Summary Statistics

```
. summarize

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
       hh_id |      4446    5.12e+07    2.73e+07   1.00e+07   1.00e+08
 avg_tot_inc |      4446    6462.558    1869.511        875      21700
avg_power_~l |      4446    726.9374    394.0966          0     3838.5
  avg_health |      4446     144.058    358.5863          0      15875
type_of_wall |         0
-------------+--------------------------------------------------------
 num_members |      4446    5.671615    1.474688          1         20
group_is_h~d |         0
has_woman_~s |      4446     .997076    .0540008          0          1
 region_type |         0
      gender |         0
-------------+--------------------------------------------------------
    religion |         0
caste_cate~y |         0
         edu |         0
household_~e |      4446    28.23064    6.982106   14.71429         95
        post |      4446          .5    .5000562          0          1
-------------+--------------------------------------------------------
avg_income~p |      4446    1149.358    202.4859        660   1459.375
  edu_coded |      4446    1.015969    .9474216          0          5
   treatment |      4446    .1351777    .3419517          0          1
religion_num |      4446    3.148898    .3863328          1          5
caste_cate~m |      4446    3.885965    1.139058          1          6
-------------+--------------------------------------------------------
group_is_h~m |      4446    1.002699    .0518881          1          2
  gender_num |      4446    1.904858    .2934436          1          2
type_of_wa~m |      4446    2.493477    1.971525          1         10
ln_avg_pow~l |      4446    6.461662    .5338829          0   8.253098
ln_avg_hea~h |      4446    4.398382    1.216571          0   9.672564
-------------+--------------------------------------------------------
ln_avg_tot~c |      4446    8.731638    .2988052   6.775366   9.985114
     _est_fe |      4446           1           0          1          1
     _est_re |      4446           1           0          1          1

.
```

Regression Results

Linear regression, absorbing indicators

Number of obs	=	4446
F(18, 4423)	=	44.99
Prob > F	=	0.0000
R-squared	=	0.1622
Adj R-squared	=	0.1580
Root MSE	=	0.4899

ln_avg_power_fuel	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
caste_category_num						
Not Stated	.2796821	.0977862	2.86	0.004	.0879723	.4713919
OBC	-.0878853	.0474723	-1.85	0.064	-.1809548	.0051842
SC	-.1911677	.0475954	-4.02	0.000	-.2844785	-.097857
ST	-.2746696	.0516828	-5.31	0.000	-.3759937	-.1733455
Upper Caste	-.0959392	.0511143	-1.88	0.061	-.1961488	.0042704
edu_coded						
1	.0647216	.0183429	3.53	0.000	.0287603	.1006829
2	.1971408	.0200816	9.82	0.000	.1577707	.2365108
3	.2525246	.0389538	6.48	0.000	.1761556	.3288935
4	.2002828	.0643655	3.11	0.002	.0740941	.3264715
5	.3987863	.3468722	1.15	0.250	-.2812568	1.078829
group_is_hospitalised_num						
Y	-.0368261	.1422514	-0.26	0.796	-.3157101	.2420578
1.has_woman_18_plus	.1921213	.1365731	1.41	0.160	-.0756304	.459873
religion_num						
Christian	0	(omitted)				
Hindu	0	(omitted)				
Muslim	0	(omitted)				
Sikh	0	(omitted)				
gender_num						
M	-.060902	.0257514	-2.37	0.018	-.1113875	-.0104164
num_members	.0797028	.0052828	15.09	0.000	.0693458	.0900597
household_weighted_avg_age	-.0027645	.0011036	-2.51	0.012	-.0049281	-.000601
ln_avg_health	.0529025	.0062769	8.43	0.000	.0405965	.0652084
post	.1555554	.0166337	9.35	0.000	.122945	.1881657
treatment	-.1772817	.0236684	-7.49	0.000	-.2236835	-.1308799
_cons	5.717541	.1526292	37.46	0.000	5.418311	6.01677
religion_num	F(4, 4423) =		4.532	0.001		(5 categories)

Significant Variables:

(Variables with p-value <0.05)

- Case_category_num: not stated, SC,ST
- edu_coded:1,2,3,4
- num_members
- household_weighted_avg_age
- Ln_avg_health
- Post
- treatment

Interpretation of Parameters:

- **Caste category - Not stated**: with a coefficient of 0.279, these households spend more on power and fuel compared to the reference caste group, this could be because of these households mostly belonging to urban and wealthier areas.

- **Caste Category - ST** : ST households spend 0.2747 less on fuel compared to other households, ST households are mostly present in rural areas where traditional fuels like biomass and wood are used more often.

- **Caste Category - SC**: SC households spend 0.1911 less on fuel compared to other households, ST households are mostly present in rural areas where traditional fuels like biomass and wood are used more often.

- **Education 1**: Increasing levels of education spend more on power and fuel in all cases, this is because educated individuals are more aware of clean energy and have better access to LPG under PMUY, thereby, increasing usage and expenditure.

- **Education 2**: coefficient of 0.1971,as education increases so does spending on clean fuels like LPG, since these people are more aware of the harmful effects of traditional fuels and are also more aware of the government policies.

- **Education 3**: Increasing levels of education spend more on power and fuel in all cases, this is because educated individuals are more aware of clean energy and have better access to LPG under PMUY, thereby, increasing usage and expenditure.

- **Education 4**: coefficient of 0.2, as education increases so does spending on clean fuels like LPG, since these people are more aware of the harmful effects of traditional fuels and are also more aware of the government policies.

- **Gender - Male**: Male headed households often spend less on cleaner fuels compared to female headed households. Female headed households were more responsive to PMUY since in rural areas, women are traditionally the ones doing all the cooking and are more affected by the emissions from fuel, and therefore are at a higher risk of catching respiratory disorders due to the emissions from traditional fuels like wood, biomass and dung.

- **Number of household members** : Larger households spend more on power and fuel, because more members of the household mean more cooking and electricity usage, which increases the fuel expenditure.

- **Household weighted average age** : Older households tend to spend more on cleaner fuels like LPG, because they may have a preference for it and may be more likely to qualify for government programs like PMUY.Older members (especially senior citizens) are at a higher risk of catching respiratory problems due to emissions from traditional fuels, this may also be a valid reason that causes older members to spend more on cleaner fuels like LPG , especially after the introduction of PMUY.

- **Ln_avg_health** : Households that spend more on health spend less on fuel,because using fuels like wood , biomass or dung may lead to harmful emissions that may cause respiratory illnesses. This increases their spending on health issues which in turn decreases their spending on power and fuel, it is like a trade off between the two.

- **Treatment** : this is the time treatment which has a coefficient of 5.7175, indicating that the households in the treatment group(those receiving PMUY) spend more on power and fuel.

- **Post** : This is the time dummy variable with a coefficient = -0.1773, which implies that the expenditure decreased after PMUY was implemented, this could be due to model misspecifications or subsidized refills of fuels which may result in net savings.

Hausman Test

```
. hausman fe re

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       fe           re         Difference          S.E.
-------------+----------------------------------------------------------------
2bn.caste_~m |   -.4058126     .2481241       -.6539367        .4011622
3.caste_ca~m |   -.0157013    -.0927237        .0770224        .2214775
4.caste_ca~m |    .0481202    -.1927574        .2408776        .2320847
5.caste_ca~m |   -.1350103    -.2770965        .1420862        .2522149
6.caste_ca~m |    .0862799    -.0918758        .1781557        .2309622
1bn.edu_co~d |   -.0287098     .0496546       -.0783644        .0263738
 2.edu_coded |    .0119544     .1784303       -.1664759        .0401698
 3.edu_coded |   -.1419965     .2293771       -.3713736        .1157145
 4.edu_coded |    .1755779     .1912615       -.0156837        .1812109
 5.edu_coded |    .3027298     .3784735       -.0757437        .3082027
2.group_is~m |    .0597033    -.0032659        .0629692        .0993766
1.has_woma~s |   -.1141419     .1517706       -.2659124         .216111
2bn.religi~m |    .1447469    -.6062921         .751039        .4346072
3.religion~m |    .0065851    -.6630164        .6696014        .3721717
4.religion~m |    .1722085    -.6620607        .8342692        .4863564
2.gender_num |    .0228939    -.0527682         .075662        .0517098
 num_members |    .0435658     .0748555       -.0312897        .0085675
household_~e |    -.006322    -.0029775       -.0033446        .0026642
ln_avg_hea~h |    .0202309     .0431583       -.0229274        .0049212
        post |    .1907721     .1609037        .0298683        .0068409
   treatment |   -.1287743    -.1638483         .035074        .0165488
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                 chi2(21) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       97.08
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)

.
```

H0: The difference in coefficients is not systematic, Random effect is appropriate.

H1: The difference is systematic, Fixed effect is appropriate.

Here , our **p-value is 0.000** which is **less than 0.05**, therefore we **reject the null hypothesis** and go ahead with a fixed effect model.

Heteroskedasticity Test

```
Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

chi2 (2223)  =        0.00
Prob > chi2 =        1.0000
```

We use the Modified Wald test for groupwise heteroskedasticity in fixed effect regression models , by using the Breusch-Pagan LM statistic.

H0: Homoskedastic

H1: Heteroskedastic

Since **p-value = 1.000 >0.05**, we **do not reject the null hypothesis** and can conclude that our model is homoskedastic in nature.

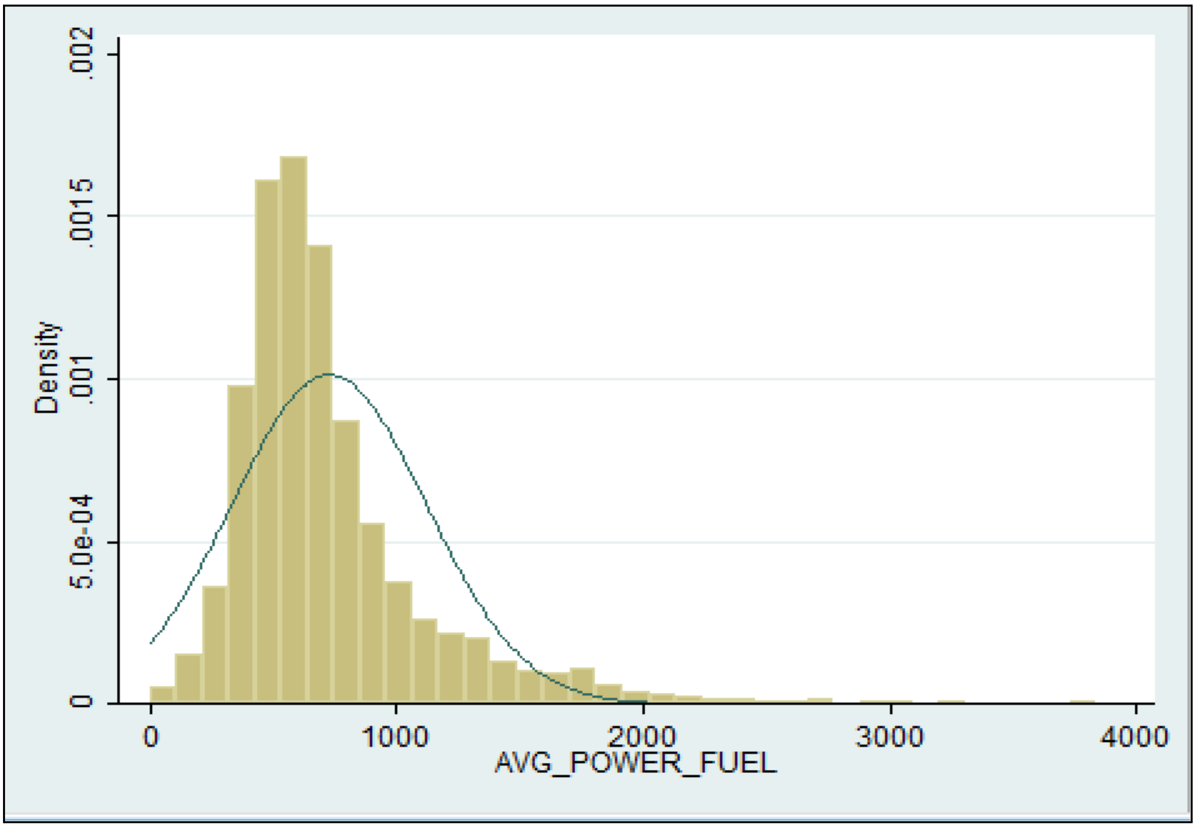
VIF Test

. vif

Variable	VIF	1/VIF
caste_cate~m		
2	1.31	0.766255
3	9.89	0.101151
4	9.39	0.106532
5	4.31	0.232198
6	5.45	0.183411
edu_coded		
1	1.39	0.717222
2	1.37	0.729441
3	1.10	0.913077
4	1.04	0.964180
5	1.00	0.998032
2.group_is~m	1.01	0.991415
1.has_woma~s	1.01	0.993239
2.gender_num	1.06	0.946406
num_members	1.12	0.895742
household_~e	1.09	0.913372
ln_avg_hea~h	1.08	0.927799
post	1.28	0.780429
treatment	1.21	0.824806
Mean VIF	2.51	

The **VIF for our model is 2.51** (overall) which is **less than 5** , this indicates that there is **very low multicollinearity** present in the model, the individual VIFs for the variables are also <10 which means we can move ahead with this model .

Normality Test



We use the Shapiro-Wilk Test for normality.

Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
avg_power_~l	4446	0.84625	402.783	15.086	0.00001

H0: The distribution of population is normally distributed.

H1: The distribution of population is not normally distributed.

From the statistical test, our **p-value is 0.00001**, hence we **reject it at 5% confidence level** which indicates that our **distribution is not normal.**

Second,if we use the **graphical method**, our **distribution is slightly skewed** which again indicates a **slight deviation from the normal distribution**. To remove this non normality we can make use of log transformations or remove outliers or even increase the sample size.

Test for Time Fixed Effects

```
. testparm i.year

 ( 1)  2014.year = 0
 ( 2)  2015.year = 0
 ( 3)  2017.year = 0
 ( 4)  2018.year = 0

       F(  4, 37884) =  144.61
            Prob > F =    0.0000
```

To assess whether time-specific factors significantly influence household expenditure on power and fuel, we conducted a test for **time fixed effects** by evaluating the joint significance of year dummy variables using the "`testparm i.year`" command in Stata.

In panel data models, **time fixed effects** control for common shocks and policy changes that vary over time but affect all households similarly—such as inflation, fuel price reforms, or national programs like PMUY. We tested whether the coefficients of year dummies (2014, 2015, 2017, and 2018, with 2016 as the base year) are **jointly equal to zero**.

Hypotheses:

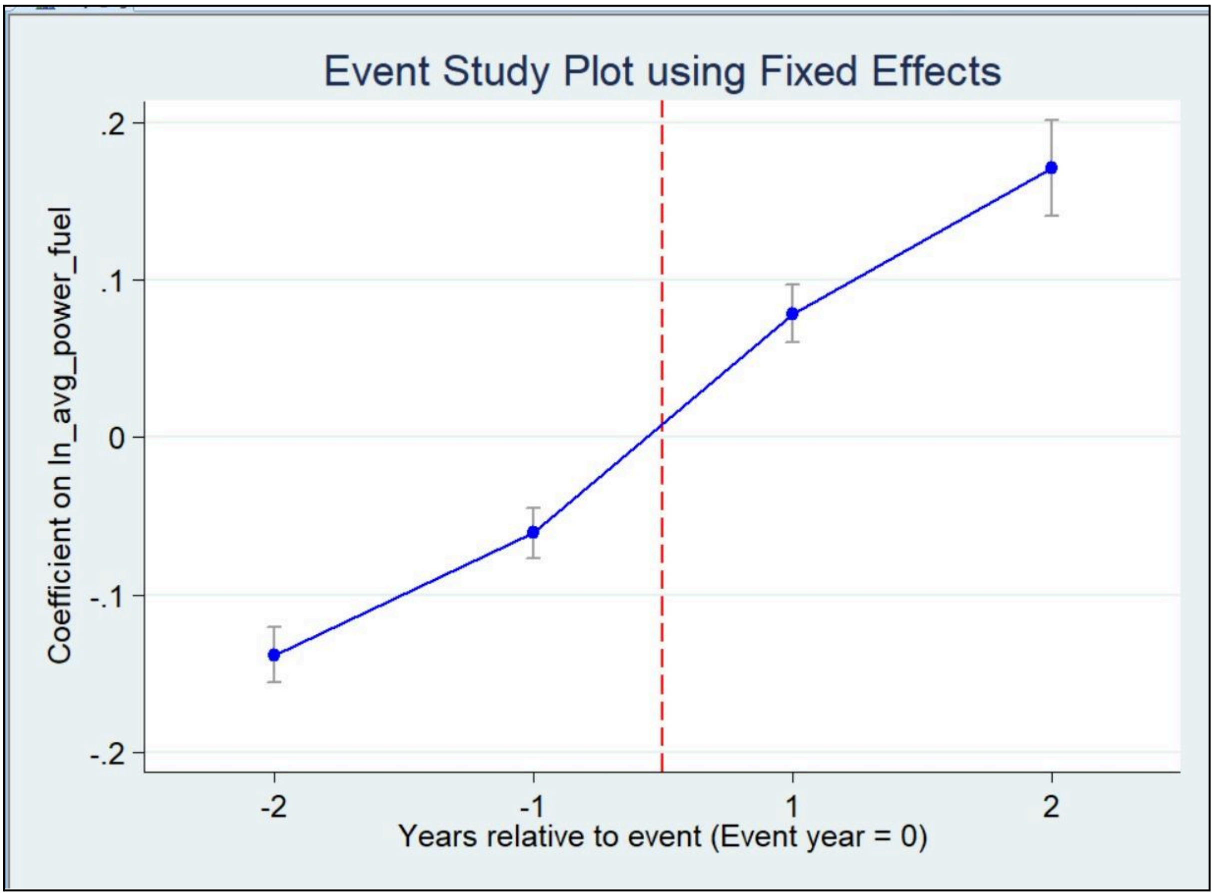
- **Null Hypothesis (H_0):** All year dummy coefficients = 0
  (No time-specific effects; year does not impact `ln_avg_power_fuel`)

- **Alternative Hypothesis (H_1):** At least one year dummy coefficient ≠ 0
  (Year has a significant effect on `ln_avg_power_fuel`)

Since the **p-value is 0.0000**, we **reject the null hypothesis**. This indicates that the year dummies are **jointly statistically significant**, and therefore, **time fixed effects should be included** in the model. Their inclusion helps capture temporal variation in household energy expenditure that would otherwise bias the results.

Event Study

Event Study Plot



It shows how the log of average power and fuel expenditure changed after the adoption of PMUY. The policy was adopted in year 0 , where -1,-2 indicate years before PMUY adoption and +1,+2 indicate years post PMUY adoption.Households that eventually receive PMUY spend less on fuel and power before the program, possibly because of use of traditional fuels. The upward trend reflects a transition from use of traditional fuels to LPG (a cleaner fuel). The fact that post-event coefficients are statistically significantly different from zero confirms the program impact is real.

PMUY significantly increased household expenditure on power and fuel in the years following adoption.Households are moving away from traditional fuels toward cleaner but costlier fuels (like LPG).But increased cost might also raise concerns for sustained affordability — especially after the initial subsidy/refill support ends.This could guide policymakers to consider refill support, awareness programs, or price stabilization mechanisms.

Conclusion

This study set out to evaluate the economic effects of India's flagship clean energy program—Pradhan Mantri Ujjwala Yojana (PMUY)—on household spending patterns, specifically focusing on expenditure on power and fuel. Using panel data from CMIE's Consumer Pyramids Household Survey (CPHS) for the years 2015 and 2017, and leveraging a fixed effects panel regression model, we estimated the causal impact of the scheme while accounting for household-level heterogeneity.

Our findings offer several key insights. First, households eligible for PMUY experienced a statistically significant decline in their energy expenditure post-intervention, as indicated by the negative and significant treatment coefficient. This suggests that the program did not merely provide access to clean fuel but also influenced broader spending behavior—potentially by reducing reliance on expensive and inefficient traditional fuels.

Second, the analysis highlights important socio-demographic determinants of energy spending. Higher levels of education and larger household size are positively associated with energy expenditure, reflecting income-related effects and greater energy needs. Conversely, caste categories such as SC and ST are associated with significantly lower spending, underscoring persistent social and economic inequalities in energy access and use. Religion also emerged as a statistically significant factor, although its specific effects varied across groups.

Moreover, the positive coefficient on post-intervention time indicates a general increase in energy expenditure across households—likely due to inflation or increased energy access—but the effect was dampened for treated households, affirming the targeting effectiveness of PMUY. Variables like household health and the number of members also played a role, supporting the view that energy spending is closely tied to broader welfare indicators.

In methodological terms, our use of household fixed effects and the absorb() function ensured that our estimates captured within-household variation, offering a more accurate picture of behavioral change. The Hausman test confirmed the superiority of fixed effects over random effects, lending robustness to our model.

Overall, the results suggest that PMUY not only improved access to clean fuel but also led to meaningful changes in household energy budgeting, particularly among socio-economically disadvantaged groups. These findings support the continued expansion and refinement of such schemes, with attention to affordability, sustained usage, and addressing barriers faced by specific communities. Going forward, integrating such programs with education, health, and gender-targeted interventions may amplify their long-term impact and help India move closer to its goals of inclusive and sustainable energy access.

Implications and Future Research

The findings suggest that schemes like PMUY can significantly affect household fuel spending, especially for poorer households. The reduced expenditure among eligible households highlights the program's potential to ease financial burdens and encourage cleaner energy use. However, the variation in impact across different social groups shows that access alone isn't enough. For long-term success, policies must address barriers like affordability, awareness, and cultural preferences to ensure sustained adoption.

Future research should track whether households continue using LPG over time or revert to traditional fuels. Adding more years of data would help assess long-term effects. Studies could also explore outcomes beyond spending—such as health or time savings—and include qualitative insights into household behavior. This would offer a fuller picture of how energy programs impact lives and where they can be improved.

References

1. **Centre for Monitoring Indian Economy. (2025).** *People of India dx*.

2. **Centre for Monitoring Indian Economy. (2025).** *Aspirational India dx*.

3. **Centre for Monitoring Indian Economy. (2025).** *Income Pyramids dx*.

4. **Centre for Monitoring Indian Economy. (2025).** *Consumption Pyramids dx*.

5. **Oli, D., & Bisht, P. S. (2024).** Evaluating the Impact of Pradhan Mantri Ujjwala Yojana on Energy Affordability for Below Poverty Line Families in Pithoragarh, Uttarakhand. *Library Progress International, 44*(3), 5360-5372.

6. **Ministry of Petroleum & Natural Gas, Government of India. (2024, October 31).** Adoption of Clean Fuel for Cooking in Rural Areas. *Press Information Bureau*.

7. **Dash, A. K., & Sahoo, D. (2024).** Effectiveness of Pradhan Mantri Ujjwala Yojana towards Sustainable Living for Indian Households. *Scope Journal, 14*(4), 1-15.

8. **IGIDR Working Paper. (2024).** Access to clean cooking fuel and women outcomes. *Indira Gandhi Institute of Development Research Working Paper WP-2024-017*.

9. **Pradhan Mantri Ujjwala Yojana: An Initiative of Government of India for Clean Energy. (2022).** *International Journal for Research Trends and Innovation, 7*(12), 123-128.