Course Title: Empirical Methods for Public Policy

Course Code: PP455E

Essay title:

From Access to Impact: Evaluating the Causal Effect of Pradhan Mantri Ujjwala Yojana on LPG Adoption, Sustained Use, Indoor Air Pollution, and Public Health in India

March 2, 2025

Author:

Arun Arya, EMPA Candidate
London School of Economics

From Access to Impact: Evaluating the Causal Effect of Pradhan Mantri Ujjwala Yojana on LPG Adoption, Sustained Use, Indoor Air Pollution, and Public Health in India

1. Introduction

Access to clean cooking fuels remains a major challenge in India, where about 660 million people still rely on traditional biomass fuels (IEA, 2021). The widespread use of firewood, coal, and dung cakes contributes to severe indoor air pollution, which accounted for 1.67 million deaths (17.8% of total fatalities) in 2019 (Pandey et al., 2021). Despite efforts to promote modern cooking technologies, high costs, poor stove quality, and behavioral inertia continue to limit adoption (Mobarak et al., 2012).

In response to this crisis, the *Pradhan Mantri Ujjwala Yojana (PMUY)* was launched in 2016 to provide subsidized Liquefied Petroleum Gas (LPG) connections to below-poverty-line (BPL) households (Government of India, 2016). However, access to LPG did not guarantee its sustained use. Prior evaluations found that PMUY increased LPG adoption but did not ensure sustained use or fully displaced firewood (CAG, 2019; Asharaf & Tol, 2024). Many beneficiaries reverted to traditional biomass fuels due to cost constraints and behavioral factors (Gaikwad et al., 2024; Selvam et al., 2021). These evaluations, however, lacked insights into long-term environmental and health impacts.

This study proposes to conduct a causal evaluation of PMUY's impact on LPG adoption, sustained LPG use, indoor air pollution levels, and respiratory health outcomes, and contributes to the causal evaluation literature on large-scale social programmes by combining nationally representative surveys with administrative LPG data in a quasi-experimental framework. The following sections present policy details, evaluate alternative methodologies, outline the empirical approach, and discuss potential evaluation challenges and mitigation strategies.

2. Policy Background & Past Evaluations

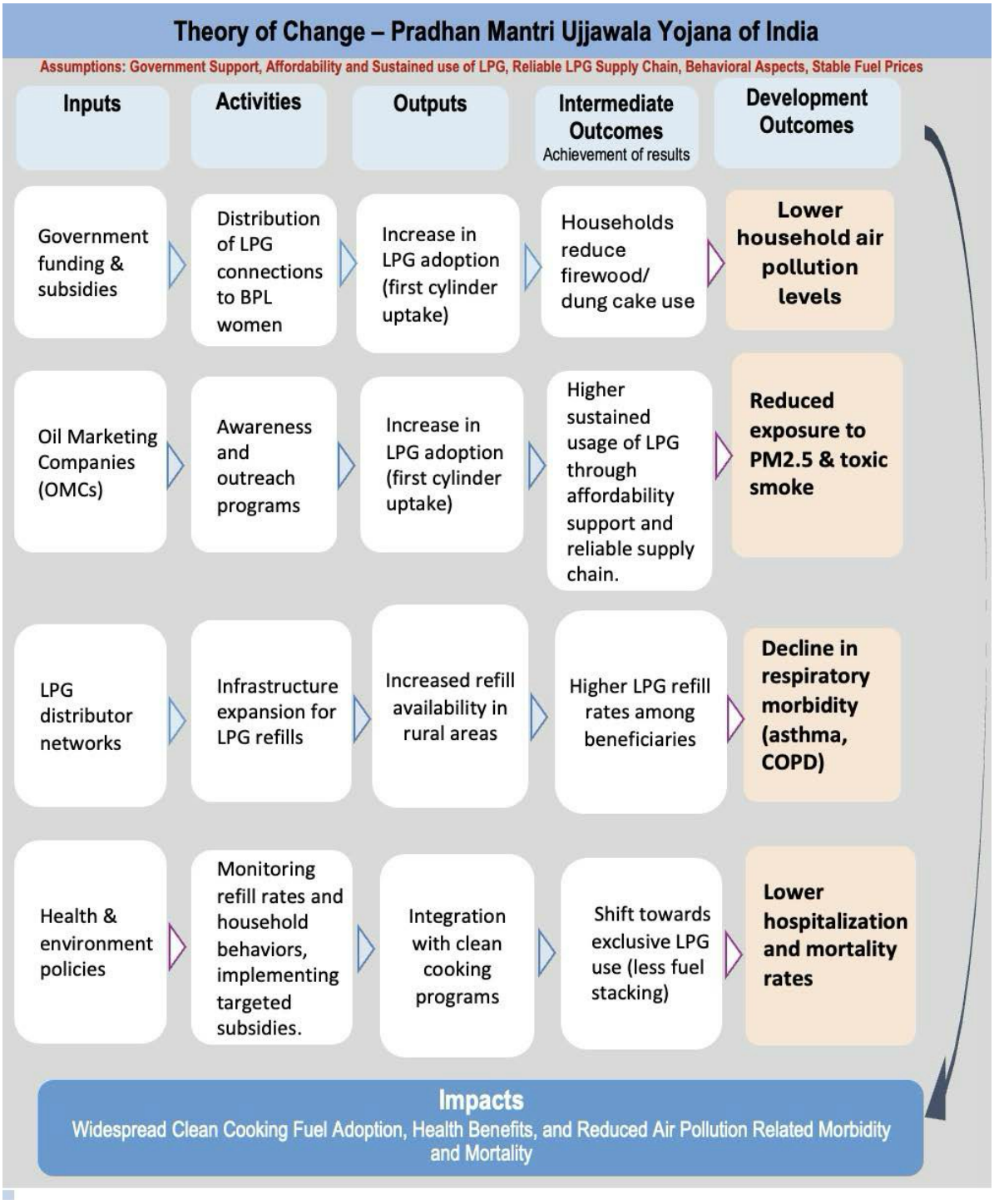
Launched in 2016, PMUY aimed to provide five crore deposit-free LPG connections to adult women of BPL households across India over the next 3 years. The target of the

scheme was increased to 8 crores by 2019-20. While the scheme offered an initial connection subsidy, it did not include guaranteed refill subsidies (Government of India, 2016). By providing LPG connections, PMUY aimed to encourage a shift toward clean cooking energy for reducing indoor air pollution and improving health outcomes.

2.1 Theory of Change

The Theory of Change for PMUY illustrates the program's intended pathway from LPG adoption to long-term health and environmental outcomes. It outlines the sequential linkages between policy implementation, behavioral shifts in fuel usage, improvements in indoor air quality, and reductions in respiratory diseases and mortality (Government of India, 2016; Government of India, 2025).

Figure 1: Theory of Change for PMUY



Source: Developed by the author based on an analysis of PMUY policy documents, past evaluations, and related literature

2.2 Past Evaluations of PMUY

Findings from previous studies provide mixed evidence of PMUY's effectiveness. They

confirm that PMUY increased LPG adoption and improved air quality, but challenges remain in sustained usage, economic benefits, and targeting effectiveness.

Table 1: Past Evaluations - Findings and Limitations

Study	Findings	Limitations
Ashraf & Tol (2024)	Difference-in Differences (DiD) & Propensity Score Matching (PSM) analysis found that the Average Treatment Effect on the intendedly Treated is 2.1 percentage point increase in LPG consumption due to PMUY, with a parallel decrease in firewood consumption. There were differential impacts by regions and social groups. Sustaining usage was challenging.	The study kept all BPL households in the intended treatment group and non-BPL households as control group. It calculated the ATE on the 'intendedly' treated not the ones 'actually' treated. It was focused only on LPG adoption not covering sustainable use, pollution and health outcomes. Used NFHS self-reported data, which may suffer from recall bias.
IIT Kanpur (2023)	PMUY successfully expanded LPG access; however, a significant proportion of beneficiaries continued fuel stacking. In terms of exposure of PM2.5 particles to the primary cooking person, the high PMUY connection villages indoor environment have 10 to 20 percent less average exposure than the low connection village.	Lack of Causal Attribution. Limited Geographic Coverage of only 6 states. Small sample size of only 2,366 households. Did not account for external factors like seasonal air pollution variations. Self-Reported Health Data.
Gaikwad et al. (2024)	PMUY contributed to improvements in mental health, economic empowerment, and time savings for women. LPG use under PMUY led to improved respiratory health, reducing exposure to indoor air pollution. Economic benefits remained unclear.	Limited Geographic Scope focusing on only five districts in Western Maharashtra. Survey-Based Approach that relied on self-reported data, which is prone to recall bias and social desirability bias. Lack of Causal Inference.
CAG Audit (2019)	PMUY contributed to a significant increase in LPG coverage, from 61.9% in May 2016 to 94.3% by March 2019. Found implementation issues, such as low refill rates (3.66/year) and connections issued to ineligible men.	No causal analysis, only descriptive audit. The study analysed a sample of 18,558 beneficiaries and 1,662 field surveys, which may not be fully representative of all PMUY users. Absence of Third-Party Validation. Limited Assessment of Health & Environmental Benefits.
Selvam, et al (2021)	PMUY led to behavioural changes among rural women in three districts in Tamil Nadu. The transition to LPG has resulted in perceived health improvements. Women saved time previously spent collecting firewood. Many beneficiaries do not regularly refill LPG cylinders due to financial constraints.	The study is confined to three districts in Tamil Nadu. The study primarily uses a descriptive research methodology rather than a rigorous econometric approach. Small sample size of only 112 respondents. The study relies on self-reported data from beneficiaries, which can introduce recall bias or social desirability bias.

Source: Compiled by the author based on Asharaf & Tol, 2024; Gaikwad et al., 2024; IIT Kanpur, 2023; Selvam et al., 2021; CAG, 2019

3. Proposed Evaluation Methodology

3.1 Study Approach

The Pradhan Mantri Ujjwala Yojana (PMUY) targeted households identified as socioeconomically disadvantaged using the Socio-Economic Caste Census (SECC 2011) — including deprivation indicators such as caste, income, and asset ownership. However, as enrolment was voluntary**,** self-selection bias may arise if households choosing to enrol differ systematically from those who do not. To mitigate this, the study will combine Difference-in-Differences (DiD) estimation with Propensity Score Matching (PSM) to improve comparability between treatment and control groups. While prior studies relied on cross-sectional or descriptive approaches, this study exploits **pre/post structure + household fixed effects.**

The DiD approach is most appropriate for evaluating PMUY's causal effects, as it estimates policy impact by comparing pre- and post-intervention outcomes across treated and matched untreated units, controlling for time-invariant unobservables. Other empirical strategies were considered. While Randomized Controlled Trials (RCTs) are the gold standard for causal inference, implementing one retrospectively for a national rollout would be infeasible due to ethical, logistical, and cost constraints**.** Regression Discontinuity (RD) designs require a sharp eligibility threshold, which PMUY lacks due to multiple overlapping selection criteria. Accordingly, a DiD framework supplemented with matching is both methodologically robust and practically feasible for this evaluation.

The analysis will use secondary data drawn from robust national and administrative datasets. LPG adoption and sustainability will be assessed using Oil Marketing Companies' (OMC) administrative data (MoPNG, 2023). Survey-based methods will measure PMUY's effects on sustained LPG usage, indoor air pollution, and respiratory health outcomes using data from NFHS and NSSO **(**NFHS, 2021; NSSO, 2012**)**. Indoor air quality changes will be tracked via PM2.5 concentration levels from CPCB data **(**CPCB, 2020**)**, while health effects such as asthma and COPD incidence will be analysed using morbidity data from hospital surveys and epidemiological studies (Gaikwad et al., 2024; Pandey et al., 2021).

Primary data collection is not proposed, though limited qualitative interviews may be used

to complement quantitative findings, where appropriate, to explore behavioural factors or beneficiary perceptions that are not fully captured in administrative or survey datasets.

3.2 Sampling Strategy

This study leverages existing nationally representative datasets to evaluate PMUY's impact using a household-level Difference-in-Differences (DiD) design. Beneficiaries (treatment group) will be compared with eligible non-beneficiaries (control group), based on eligibility criteria from the Socio-Economic Caste Census (SECC) and LPG connection records.

The primary datasets include:

- **National Family Health Survey (NFHS)** waves from 2015–16 and 2019–21 for household energy use, demographics, and health indicators.
- **NSSO Consumption and Social Consumption Surveys** for fuel usage patterns and socio-economic profiling.
- **OMC administrative data** for LPG distribution and refill rates at district and household levels.

To ensure robust representativeness, the analysis will cover a stratified sample of districts available within these datasets, reflecting variation across:

- High vs. low LPG penetration states.
- Rural vs. urban areas.
- Socio-economic categories (caste, income, education).

Propensity Score Matching (PSM) will be used to construct comparable treatment and control groups, based on observable characteristics from NFHS and NSSO. This helps mitigate selection bias in uptake.

3.3 Pre- and Post-PMUY Periods

To estimate both short- and long-term programme effects, the analysis will compare outcomes across two time periods:

- **Pre-PMUY Period (2010–2015):** Establishing baseline trends in LPG adoption, indoor air pollution, and health conditions.

- **Post-PMUY Period (2016–2021)**: Capturing changes in outcomes following the programme's rollout, including indicators of sustained use.

The parallel trends assumption will be tested using pre-intervention NFHS and NSSO data. If valid, the DiD framework will allow causal interpretation of observed changes

3.4 Econometric Model

The standard DiD estimator offers a simple and intuitive approach for identifying policy impacts. However, when applied to household-level panel data, it may be biased if unobserved household characteristics or nationwide time shocks affect the outcome.

To improve robustness, the study adopts a Difference-in-Differences model with Two-Way Fixed Effects (TWFE), controlling for both household-specific unobservables and time trends (Angrist & Pischke, 2009). While recent literature has noted limitations of TWFE in settings with staggered adoption or treatment heterogeneity (Goodman-Bacon, 2021), PMUY's nationwide rollout in 2016 allows the TWFE framework to remain valid and informative.

Figure 2: Econometric Model

Econometric Model

$$Y_{it} = \beta_0 + \beta_1 Post_t + \beta_2 Treatment_i + \beta_3 (Post_t \times Treatment_i) + X_{it}\gamma + \lambda_t + \mu_i + \epsilon_{it}$$

where:

- Y_{it} = Outcome variable for household ii at time tt (**LPG adoption, sustained LPG use, PM2.5 levels, or respiratory morbidity**).
- $Post_t$ = 1 if the time period is after PMUY introduction (2016 or later), 0 otherwise.
- $Treatment_i$ = 1 if the household received PMUY LPG connection, 0 otherwise (eligible but non-beneficiary households).
- ($Post_t \times Treatment_i$) = **DiD estimator** capturing the **causal effect of PMUY**.
- X_{it} = Control variables (household size, income, education, rural/urban, caste, previous LPG access).
- λ_t = **Time fixed effects** (absorbs nationwide trends, e.g., inflation, fuel price fluctuations).
- μ_i = **Household fixed effects** (controls for household-specific unobserved factors).
- ϵ_{it} = Error term.

Source: *Author's formulation based on standard DiD methodology, adapted from Angrist & Pischke, 2009*

This enhanced DiD model provides more credible causal estimates by controlling for both household fixed effects (capturing unobserved time-invariant characteristics) and year fixed effects (absorbing national-level shocks such as fuel prices or inflation), strengthening internal validity.

4. Key Outcome Measures

The evaluation will focus on four key outcomes to comprehensively assess PMUY's effectiveness:

1) **LPG Adoption**: Measured by whether eligible households received an LPG connection, this serves as an indicator of program's reach and expanded access to clean cooking fuel.

2) **Sustained LPG Use**: This outcome assesses whether households purchased a third LPG refill independently, capturing long-term behavioral shifts and potential affordability constraints that may hinder sustained use.

3) **Reduction in Indoor Air Pollution**: This will be measured using PM2.5 concentration levels, evaluating whether LPG adoption led to a measurable reduction in indoor pollution.

4) **Health Improvements**: Given the strong link between household air pollution and respiratory diseases, this outcome tracks changes in respiratory morbidity (e.g., asthma, COPD, pneumonia) among beneficiaries.

Together, these outcomes allow for a comprehensive evaluation of PMUY—ranging from access and sustained behavioral change to environmental and health impacts—thereby capturing the full causal chain of program benefits.

5. Data and Sources

To implement the proposed evaluation, the study will integrate multiple data sources from government records, surveys, and environmental monitoring datasets.

Table 2: Data Sources

Data Source	Type	Usage in Evaluation
Oil Marketing Companies (OMC) Records at MoPNG	Administrative	LPG connection and refill patterns (DiD estimation).
NFHS & NSS Household Surveys	Survey	Household fuel usage, health outcomes (self-reported).
Hospital & Health Records	Medical	Respiratory morbidity trends (COPD, asthma, pneumonia).
Central Pollution Control Board (CPCB) Data	Environmental	PM2.5 concentration trends in sample districts.
Satellite-based Air Quality Models	Remote Sensing	Supplementary data for indoor air pollution exposure estimates.

Source: Compiled by the author based on Government of India, 2016, 2025; NSSO, 2020; NFHS, 2021; OMC MoPNG Reports, 2022

6. Evaluation Threats and Mitigation Strategies

Conducting a DiD evaluation of PMUY requires addressing threats related to causal attribution (internal validity) and the generalizability of results (external validity) that could bias the findings.

6.1 Internal Validity Threats (Causal Attribution Issues)

1. **Self-Selection Bias in Programme Uptake**
   PMUY's eligibility is determined by socio-economic criteria under the Socio-Economic and Caste Census (SECC), but enrolment is voluntary — meaning households self-select into the programme. This introduces a risk of selection bias, as more motivated, health-conscious, or better-informed households may be more likely to apply and sustain LPG usage (Asharaf & Tol, 2024).

   Mitigation Strategy:

   - Propensity Score Matching (PSM) will be employed to balance observable household characteristics between treated and untreated eligible groups.
   - Household fixed effects in the DiD model will control for unobservable time-invariant factors.

- Robustness checks will explore how results vary across groups with differing propensities to enroll, and heterogeneity analysis will test for selection on unobservables.

2. **Programme Targeting Bias**

   Even within eligible households, administrative discretion or information asymmetries may have led to the early inclusion of relatively wealthier or better-connected BPL households.

   Mitigation Strategy:

   - Baseline characteristics of early adopters vs. late adopters will be compared.
   - Matching and stratification will reduce bias from this targeting pattern.

3. **Parallel Trends Assumption Violation**

   A key DiD assumption is that absent PMUY, treatment and control groups would have followed similar LPG adoption and health trends. If pre-existing differences exist, estimates may be biased.

   Validation Strategy:

   - Pre-treatment trends in LPG usage and respiratory illness (2010–2015) will be examined.
   - **Placebo tests** using false treatment years (e.g., 2014) will be conducted.
   - An **event-study design** will be used to trace dynamic treatment effects over time.

4. **Spillover Effects & Contamination**

   Some non-beneficiary households may acquire LPG independently (MoPNG, 2023), diluting the observed treatment effects of PMUY.

   Mitigation Strategy:

   - The control group will exclude any households that acquired LPG during the study window.
   - Sensitivity tests will be run with varying control group definitions.

5. **Attrition Bias**

   If households drop out due to migration or non-response, the Longitudinal household

datasets may become non-random over time (NFHS, 2021).

Mitigation Strategy:

- Attrition will be assessed for randomness.
- Multiple imputation techniques will correct for missing data if necessary.

6.2 External Validity Threats (Generalizability Issues)

Even with randomization, several factors could limit the applicability of findings beyond the sample.

1. **Geographic and Socio-Economic Representation**: The study's findings may not be generalized if the sampled districts and households are not representative of India's

   diverse socio-economic and geographic conditions.

   Mitigation Strategy:

   - Stratified sampling across high vs. low LPG penetration states, rural vs. urban areas, and different income groups.
   - Heterogeneity analysis will assess variation in programme impact across subgroups (e.g., SC/ST, gender, geography).

2. **Policy Implementation Differences**: PMUY's effectiveness may be influenced by LPG distributor density, administrative efficiency, and awareness campaigns, which vary across states.

   Mitigation Strategy:

   - State fixed effects will control for time-invariant regional implementation differences.
   - Interaction terms will test whether stronger administrative contexts improved outcomes.

3. **Dependence on Survey-Based Data for Pollution & Health Outcomes:** Unlike LPG adoption (which can be tracked through administrative records), pollution and health impacts rely on self-reported or survey-based measures, which may suffer

from measurement errors and recall bias.

Mitigation Strategy:

- Administrative and hospital-based health data (Gaikwad et al., 2024; MoHFW, 2022) will supplement survey data.
- PM2.5 concentration estimates from satellite and remote sensing data will strengthen pollution metrics (CPCB, 2020).

4. **Temporal Boundaries of the Study:** This study evaluates PMUY's impact based on historical data (2016-2021). Future changes in LPG pricing, subsidy structures, or alternative clean cooking technologies may influence household behavior in ways not captured in this analysis. Findings will, therefore, be interpreted within the context of PMUY's implementation period rather than assumed to apply indefinitely.

   Mitigation Strategy:

   - The findings will be contextualized within this implementation period and not assumed to extrapolate to future policies.

7. Limitations of this Methodology

While the proposed methodology (DiD + TWFE using administrative and survey data) strengthens causal inference, several limitations remain:

1. **Reliance on Secondary Survey Data**
   Health and pollution outcomes depend on survey-based or administrative sources (NFHS, NSSO, hospital records), which may suffer from recall bias, measurement error, or limited temporal coverage (NFHS, 2021; NSSO, 2012). Despite triangulation with hospital and satellite data, these issues may attenuate effect sizes.
2. **Unobserved Time-Varying Confounders**
   Although household fixed effects address time-invariant unobserved heterogeneity, the DiD approach may still be vulnerable to **time-varying shocks** (e.g., concurrent health campaigns, LPG price changes) that differentially affect treatment and control groups.

3. **Policy Mechanisms Not Directly Tested**
   While the study estimates PMUY's overall impact, it does not directly test the **effectiveness of policy levers** like targeted subsidies, phased refills, or direct benefit transfers (DBT), which could influence long-term LPG use (Pandey et al., 2021). Experimental designs (e.g., RCTs) would be needed to evaluate such mechanisms.
4. **Limited Generalizability Beyond Study Period**
   The analysis covers the 2016–2021 implementation window of PMUY. Changes in LPG pricing, subsidy structures, or alternative fuel access may affect the relevance of findings for future policy decisions.

8. Conclusion

The Pradhan Mantri Ujjwala Yojana (PMUY) represents a landmark initiative to expand access to clean cooking energy among BPL households in India. While it achieved remarkable scale in LPG adoption, questions remain about sustained use, refill affordability, and continued reliance on biomass fuels. This study employs a Difference-in-Differences (DiD) design with Two-Way Fixed Effects (TWFE) to estimate PMUY's long-term impact on four critical outcomes: LPG adoption, sustained usage, indoor air pollution, and respiratory health.

Relying on secondary survey and administrative data, the methodology enables robust causal inference while avoiding the cost and bias associated with new data collection. Still, challenges remain — particularly in measuring pollution and health impacts, which depend on survey and hospital data that may introduce recall or reporting errors. While the study identifies program-level effects, it does not experimentally test specific policy mechanisms that may enhance long-term usage.

International experiences such as Indonesia's Kerosene-to-LPG transition suggest that targeted refill subsidies, behavioural nudges, or integration with health programs may further strengthen clean fuel adoption. While outside this study's empirical scope, such interventions merit future evaluation through experimental designs. Future research — especially through randomized controlled trials (RCTs) — can help test the effectiveness of such complementary approaches and guide the design of next-generation clean cooking policies.

By shifting attention from access to **sustainability and health impact**, this study contributes to evidence-based policymaking in India's clean energy transition. It lays the groundwork for understanding how flagship schemes like PMUY can be strengthened to deliver deeper social and environmental returns — not just in India, but across other low- and middle-income countries pursuing similar goals.

The analytical approach adopted here reflects the type of policy-driven, methodologically rigorous research I intend to deepen and extend during doctoral training

References

Angrist, J. & Pischke, J., 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Asharaf, N. & Tol, R.S.J., 2024. The impact of Pradhan Mantri Ujjwala Yojana on Indian households. *Economic Analysis and Policy, 84*, pp.878-897. Available at: https://doi.org/10.1016/j.eap.2024.09.026.

Central Pollution Control Board (CPCB), 2020. *National Ambient Air Quality Status and Trends.* Available at: https://cpcb.nic.in/openpdffile.php?id=UHVibGljYXRpb25GaWxlLzk5Ml8xNjUwNjA4OTgyX05hcW1zMTkucGRm.

Comptroller and Auditor General of India (CAG), 2019. *Performance Audit of Pradhan Mantri Ujjwala Yojana (PMUY).* New Delhi: CAG of India. Available at: https://cag.gov.in/en/audit-report/details/55961.

Gaikwad, H., Pandey, S. & Patil, K., 2024. Empowering rural Indian women by clean fuel initiatives: an impact assessment of Pradhan Mantri Ujjwala Yojana. *Cogent Social Sciences.* Available at: https://doi.org/10.1080/23311886.2024.2428823.

Government of India, 2016. *Pradhan Mantri Ujjwala Yojana: Scheme Guidelines.* Ministry of Petroleum and Natural Gas. Available at: https://mopng.gov.in/files/marketing/lpg/ujscheme.pdf [Accessed 8 Feb. 2025].

Government of India, 2025. *Pradhan Mantri Ujjwala Yojana (PMUY) – Official Website.*

Available at: https://pmuy.gov.in/ [Accessed 5 Feb. 2025].

Indian Institute of Technology (IIT) Kanpur, 2023. *Evaluation of Pradhan Mantri Ujjwala Yojana (PMUY) in Six States of India.* Available at: https://nhsrcindia.org/sites/default/files/IRHSS%20%E2%80%93%20PMUY%20Evaluation%202023.pdf.

International Energy Agency (IEA), 2021. *India Energy Outlook 2021.* Paris: IEA. Available at: https://iea.blob.core.windows.net/assets/1dc6d91c-e23f-4e02-b1fb-51fdd6283b22/India_Energy_Outlook_2021.pdf.

Ministry of Petroleum and Natural Gas (MoPNG), 2023. *Annual Report 2022-23.* Government of India. Available at: https://mopng.gov.in/files/TableManagements/IPNG-Annual-Report-2022-23-web.pdf.

Mobarak, A.M., Dwivedi, P., Bailis, R., Hildemann, L. & Miller, G., 2012. Low Demand for Nontraditional Cookstove Technologies. *Proceedings of the National Academy of Sciences, 109(27),* pp.10815–10820. Available at: https://www.pnas.org/doi/pdf/10.1073/pnas.1115571109.

National Family Health Survey (NFHS), 2021. *National Family Health Survey (NFHS-5), 2019-21 Compendium of Fact Sheets.* New Delhi: Ministry of Health and Family Welfare. Available at: https://mohfw.gov.in/sites/default/files/NFHS-5_Phase-II_0.pdf.

National Sample Survey Office (NSSO), 2012. *Household Consumption of Various Goods and Services in India.* New Delhi: Ministry of Statistics and Programme Implementation. Available at: https://mospi.gov.in/sites/default/files/publication_reports/nss_report_541.pdf.

Pandey, A. et al., 2021. Health and Economic Impact of Air Pollution in the States of India: The Global Burden of Disease Study 2019. *The Lancet Planetary Health, 5(1),* pp.e25–e38. Available at: https://www.thelancet.com/journals/lanplh/article/PIIS2542-5196(20)30298-9/fulltext?ref=insights.onegiantleap.com.

Selvam, V., Ashok, D., Rajamanoharan, I.D., Rajalakshmi, V. & Vidhya, K., 2021. Impact of Ujjwala Yojana Scheme and Its Effect on Behavioural Changes Among Rural Women. *International Journal of Asian Business and Information Management, 13(1).* Available at: https://www.igi-global.com/article/impact-of-ujjwala-yojana-scheme-and-its-effect-on-behavioural-changes-among-rural-women/315752.

World Bank, 2013. *Indonesia Toward Universal Access to Clean Cooking.* World Bank. Available at: https://documents1.worldbank.org/curated/ar/105441468044144806/pdf/792790ESW0P1290ox0377371B00PUBLIC00.pdf.