

# Adding fuel to fire?

## Social spillovers in the adoption of LPG in India

Suchita Srinivasan\*      Stefano Carattini†

July 1, 2019

### Abstract

In many developing countries, many households, especially in rural areas, are still heavily reliant on solid biomass as a cooking fuel, despite its negative health and environmental implications. Liquefied petroleum gas (LPG) is a clean alternative, but its higher cost implies that its use is often limited to the richer, urban areas of a country. This paper focuses on the Indian context and investigates, over a relatively long time-frame, whether social spillover effects might have played a role in a household's decision to use LPG, and how these effects varied across different sub-populations. Using data from several waves of the National Sample Survey (NSS), the recent ACCESS survey, and the India Human Development Survey (IHDS), this paper provides multiple strands of evidence that, taken jointly, suggest that positive social spillovers are present. Spillovers are also found to be stronger for households that belong to social networks, than for households that do not belong to any network. Our results provide partial evidence on convergence in LPG use rates across subgroups of the Indian population, and have strong implications for policy-makers around the world who could leverage lessons from social learning to encourage consumers to switch to cleaner sources of energy.

**Keywords:** Clean cooking fuels; LPG; Technological adoption; Spillovers; India

**JEL Codes:** D83; Q48; Q53; R12; R23; R29

---

\*We would like to thank Laura Diaz Anadon, Richard Baldwin, Geoffrey Barrows, Francois Cohen, Joëlle Noailly, Timothy Swanson, Sofia Trommlerova and three anonymous reviewers for helpful comments on an earlier version of this paper.

\*Corresponding author. Chair of Energy and Public Economics, Department of Management, Technology and Economics, ETH Zurich and Centre for International Environmental Studies, Graduate Institute of International and Development Studies, Geneva Switzerland. E-mail: suchitas@ethz.ch. Funding from the Swiss Network For International Studies (SNIS) is gratefully acknowledged.

†Department of Economics, Andrew Young School of Policy Studies, Georgia State University, CESifo, and Grantham Research Institute on Climate Change and the Environment and ESRC Centre for Climate Change Economics and Policy, London School of Economics and Political Science. E-mail: scarattini@gsu.edu. Funding from the Swiss National Science Foundation, grant P2SKP1\_165028, and the Centre for Climate Change Economics and Policy, which is funded by the UK Economic and Social Research Council, is gratefully acknowledged.

# 1 Introduction

The use of solid biomass as a cooking fuel is still rampant in the developing world, and is one of the main causes of indoor air pollution, and its deleterious health effects (World Health Organization, 2018).<sup>1</sup> Solid fuel use is still common in many parts of South Asia, Africa, and Latin America. In India, for instance, even as recently as 2015, almost 59% of the population (or about 780 million people) still used biomass for cooking (International Energy Agency, 2017).

The adoption and sustained use of clean cooking fuels, and efficient cookstoves, remains one of the primary means of mitigating the risks of indoor air pollution in countries like India. Clean cooking alternatives, such as liquefied petroleum gas (LPG) have grown in popularity over time, but, for several decades, rather slowly. Only very recently, an acceleration in adoption has been observed, largely due to the Pradhan Mantri Ujjwala Yojana (PMUY) launched in 2016, according to which 80 million free connections of LPG have been distributed to largely poor households in rural areas of India (Jain et al., 2018). The literature abounds on the role of socioeconomic factors in determining which households use clean cooking fuels in developing countries (Lewis and Pattanayak 2012 provide a thorough literature review). Income, education, and urbanization are found to be the most common determinants of the choice to adopt clean cooking fuels, including in India, along with access to cleaner cooking fuels (Reddy 1995, Rao and Reddy 2007, Kumar and Viswanathan 2007, Farsi et al. 2007, Gupta and Kohlin 2006).

However, socioeconomic determinants need not be the only factor influencing households' decisions to adopt clean fuels (or clean technologies more broadly). In this respect, a growing literature has examined the role of social spillovers, or how the decisions of a household's neighbors, social network or friends may influence its own decisions, in the context of energy-related consumption choices. The literature on developed economies

---

<sup>1</sup> The risk of exposure is particularly high among women and children, who mostly stay indoors, and spend considerable amounts of time near open fires.

has looked at the role of spillovers in explaining the adoption of green technologies like solar panels. Bollinger and Gillingham (2012) study the presence of peer effects in the diffusion of solar panels in California, and find that an additional solar panel in a given ZIP code is likely to increase the probability of adoption by households in the same ZIP code by 0.78 %. Graziano and Gillingham (2015) also study the diffusion of photo-voltaic panels in Connecticut, and find a similar pattern. Additional literature confirms the results for solar panels, and provides new evidence for other green technologies, such as hybrid cars (Carattini et al. 2017 provide a review).

A vast empirical literature also exists on the role of spillovers in new technology adoption in developing countries, especially concerning agricultural and health innovations. Broadly, this literature has found that households are more likely to be influenced by other households, if they have strong information linkages with them (Conley and Udry, 2010) and if social interactions offer opportunities to learn how to use a product, rather than just influencing the desire to have it (Oster and Thornton, 2012). Households may also delay their decision to adopt a given technology, if they aim to free ride on others adopting early and generating opportunities for learning (Bandiera and Rasul, 2006). In some cases, households may also delay adoption indefinitely, if other people's adoption create positive externalities from which they could benefit. For instance, Kremer and Miguel (2007) found that positive externalities across households resulted in slow diffusion of de-worming drugs among many households during an experimental study conducted in Kenya.

Our study contributes to the literature on the adoption (and use) of clean fuels in developing countries by investigating the extent to which social spillovers may affect the adoption of LPG in India. This paper is the first, to the best of our knowledge, to provide an empirical estimate for social spillovers in this context, and we do so with a multi-layered empirical design. By explicitly controlling for factors found to be important for LPG adoption in the literature, and incorporating a rich set of socioeconomic and demographic controls, our paper provides multiple strands of evidence on why social spillovers

may act as a possible determinant of a household's decision to adopt LPG. Focus on the use of LPG as a fuel, which is the economic outcome of interest, ensures that we look at the continued use of clean cookstoves that burn LPG. As shown in a recent study by Hanna et al. (2016), mere ownership of improved cookstoves does not necessarily imply that households will reap their health benefits. Regular use (and maintenance) of cookstoves is important to guarantee improved health outcomes (Duflo et al. 2008). This may be particularly relevant, as Mobarak et al. (2012), for instance, find from surveys in Bangladesh that households' willingness-to-pay for improved cookstoves is low, as households tend to underestimate the risk of ill-health from burning solid biomass (cf. also Greenstone and Jack 2015). This may lead to some households only using these stoves if they are provided for free, and thus limiting their regular use.

While sources of fuel such as firewood are available freely, we provide evidence that households can be influenced by other households residing in the same village or urban block to purchase LPG, suggesting that spillovers may be effective not just in inducing investment in technologies (such as cookstoves), but also their regular and continued use, which is the ultimate goal to reduce indoor pollution. Hence, our findings contrast with Beltramo et al. (2015), who implement a randomised control trial in Uganda to study the role of spillovers in clean technologies but do not find any, because households desiring the stoves lack the financial resources to purchase them. By analyzing the context of India over several decades, we can analyze social spillovers in different phases of the market for clean cookstoves and LPG, and at different levels of economic development. Importantly, for both rural and urban areas, our paper covers periods in which the use of LPG was rather non-normative, as well as more recent periods of relatively high adoption. Furthermore, we explore the mechanisms behind the observed social spillovers, by looking at the role of social networks. We show that membership in specific types of social networks may play an important role in determining the average LPG adoption rates in villages and in urban blocks. Specifically, we find that the level of LPG adoption is particularly high among households that belong to certain groups or associations, such as women's groups and

developmental groups, in which flows of information (on LPG and its associated benefits) are likely to be higher. Hence, our results may suggest that social networks play a critical role in dispersing information about LPG use and thus in encouraging households to switch to clean energy sources. This result is consistent with the literature on technology adoption in developing countries, which finds that close social interaction is more likely to lead to stronger flows of information across households (Bandiera and Rasul 2006, Oster and Thornton 2012, Munshi 2004).

To analyze social spillovers in the adoption of LPG over several decades, as well as in the most recent periods, we use three large-scale surveys on household-level consumer expenditure and energy access. Variation in the design of the surveys, including multiple cross sections and panel data, also allow us to employ different empirical strategies to shed light on the mechanisms behind the observed social spillovers. Furthermore, the relatively large sample sizes allow us to compare the adoption of cooking fuels across all areas of the country, and across very heterogeneous sets of households, both in terms of socioeconomic characteristics, and regulatory context. The first dataset is drawn from the National Sample Survey (NSS) on Household Consumer Expenditure, which comprises repeated cross sections from 1983 up till 2011-2012. The source of the second dataset is the very recent Access to Clean Cooking Energy and Electricity Survey of States (ACCESS), which is a cross-sectional dataset of about 8000, relatively homogenous, rural households from 2015. The final dataset is drawn from the India Human Development Survey (IHDS) consumer expenditure survey, a two-year panel (2005-06 and 2011-12) with repeated observations for about 40,000 households.

The LPG market is dynamic in India, and over the last decades has undergone a significant transformation with changes in the economic and social context. Moreover, the policy environment has also changed, with a spate of policies being implemented in more recent years. These changes make the context of India particularly interesting, especially given that most developing countries are still lagging behind in terms of LPG adoption. For instance, more than 90 % of the population in many sub-Saharan African countries

such as Benin, Burkina Faso, Guinea, and Mali still used solid fuel as the primary fuel in 2013 (Van Leeuwen et al., 2017). To fully understand the role of social spillovers in the process of adoption of LPG in India, data collected from a variety of sources, and spanning a long time horizon, are thus necessary. The NSS data, for instance, is collected on a large-scale sample of representative Indian households, and allows us to go back in time to understand the evolution of spillovers. The ACCESS dataset is unique in that it also collects information on supply of LPG, while its relatively homogenous sample allows us to distill the effect of spillovers among rural households that have historically been under-supplied with LPG. Using the IHDS data, we are able to incorporate household fixed effects, which help us to control for time-invariant unobserved heterogeneity. Moreover, the IHDS provides us with information about membership in social networks.

From an empirical perspective, our strategy in this paper is thus to provide multiple pieces of evidence, which, when taken together, provide complementary evidence on the presence of social spillovers in LPG use. We employ an instrumental variable linear probability model (IV-LPM) approach to study the determinants of a household's decision to use LPG as the primary cooking fuel, focusing mainly on the corresponding decision taken by other households in the same village, or urban block, in order to control for known sources of endogeneity. We leverage variation in the underlying data and use multiple instruments to provide further robustness to our empirical strategy, implementing the standard toolbox to test the validity of our instruments.

Our results are relevant to policy-makers operating in similar contexts and aiming at reducing the use of polluting cooking fuels, with subsidies or other measures. Based on our evidence, policy-makers may try to leverage existing social interactions, e.g. by targeting their interventions towards segments of society which are influential, and thus likely to affect the behavior of other households, especially if the extent to which learning occurs depends on the structure of the local social network (cf. Banerjee et al. 2014). In developed countries such as the United States, for instance, 'solar ambassadors' have been introduced in the market for solar panels to try to spur the adoption of non-normative

behaviors (Kraft-Todd et al., 2018). Leveraging social networks can lead to very cost-effective interventions whenever people are influenced, in the adoption of a new technology, by the number of people around them already using it (see Spencer et al. 2019 for a technical assessment).

The structure of the paper is as follows: section 2 provides a background on cooking fuel use in India, section 3 elaborates on the data used as well as the empirical approach, section 4 presents the empirical results and discusses potential policy implications, and section 5 concludes.

## **2 Background on Cooking Fuel Use and Social Structure in India**

The Indian context is particularly interesting to examine the adoption of LPG, especially with a perspective spanning several decades. Several sources of energy are used as cooking fuels in India, and the energy choices have historically varied between rural and urban households. Rural households have tended to have strong preferences for bio-fuels such as firewood, charcoal and agricultural waste, whereas many urban households have switched to electricity, kerosene, and LPG. Fuels derived from solid biomass such as firewood are not only cheaper (sometimes available for free) and more easily accessible, but they are also difficult to wean households off. According to Census data, in 2011 almost 67 % of the overall Indian population relied on solid fuels such as firewood, crop residue, dung cakes and coal as the primary cooking fuel, and the proportion was almost 85 % among rural households. This might have been due to affordability and easy availability, but also due to cooking habits and preferences, which may change rather slowly over time. In this paper, we choose to restrict our attention to the use of LPG as the clean cooking fuel alternative. This is because it is the most widely available clean cooking fuel in India, and the most affordable.

The energy transition has been more sustained in the urban sector than in the rural: in 1987, for instance, consumption of traditional biomass and LPG was not significantly different among rural and urban households, whereas in 2010, 60 % of urban households used LPG, without simultaneously using biomass-based fuels, while only 10 % of rural households did so (Cheng and Urpelainen, 2014). Rural households are often unable to afford the recurrent expenditures needed to acquire the cylinders, and also have difficulties in purchasing cookstoves. LPG users are also required to have a permanent and verifiable residential address, which limits the access of poor or homeless people, or even migrant workers in urban areas (Gupta and Kohlin, 2006), although in recent years the government has relaxed this requirement for smaller LPG cylinders at least in some areas (Press Information Bureau, 2014). LPG is marketed by state-owned petroleum distribution companies, and its price is fixed by the Ministry of Petroleum and Natural Gas. The government has subsidized LPG (and kerosene) since the late 1960s, although in recent times efforts are being made to phase these subsidies out.

Nevertheless, while it seems that LPG is subsidized to meet the requirements of poor households, the benefits of these subsidies have largely accrued to the richer urban households. According to recent estimates, the top 20 % of the population, by consumption expenditure, received 60 % of the total direct subsidy, whereas the bottom 50 % of the population received about 8 % of the subsidy (International Institute for Sustainable Development, 2014). There are also disparities in the distribution of subsidies and LPG connections across Indian states. For instance, five states (Maharashtra, Andhra Pradesh, Tamil Nadu, Uttar Pradesh and Karnataka) account for around 50 % of the total connections of LPG. The same five states, for instance, receive almost 50 % of the subsidies, and, even within these states, the urban areas benefit the most (International Institute for Sustainable Development, 2014).

Acquiring an LPG connection requires a fixed cost to purchase the stove, and install the equipment. Households purchase gas cylinders as and when required. Income and awareness are thus obvious determinants of the choice of a household to consume cleaner fuels



such as LPG. However, the shift to cleaner fuels may not necessarily follow the energy-ladder model, according to which households switch to cleaner cooking fuels in a linear way as the level of income increases. In this respect, we note that fuel-stacking is still commonly observed among many Indian households, where a mixture of modern and traditional fuels are used simultaneously (Cheng and Urpelainen, 2014).

While access to LPG has improved over time, especially in rural areas, data drawn from a very recent study suggests that for a sub-sample of rural households belonging to some of the poorest states of the country, almost 58 % out of 8,566 households stated that LPG availability was a challenge for them (Jain et al., 2018). Data drawn from this study also suggest that almost 84 % of the households covered by the study continued to use firewood, suggesting that transitioning to clean fuels remains challenging in rural areas of India. Moreover, among both urban and rural households, the policy emphasis has shifted to usage; ensuring that households not only acquire connections, but also use them on a regular basis, has become a priority. Accordingly, our study focuses on the role of social spillovers on the decision to use LPG as a primary fuel.

In the last decade, reforms have been undertaken by governments to improve the accessibility of LPG to Indian consumers, both rural and urban, and to try to improve the provision of subsidies. However, either these reforms have often been punctuated with policy reversals, or they have not had a considerable impact on improving the actual disbursement of subsidies. For instance, in September 2012, the central government capped the number of subsidized cylinders that a household can acquire at six per year. In January 2013, however, the limit was increased to nine cylinders per household per annum, which was further increased to 12 by 2014. Successive governments have found it politically infeasible to initiate a phase-out of the subsidies, even though efforts are being made to allocate more resources towards poorer, rural households.

For instance, in March 2015, the central government initiated a policy encouraging rich, urban consumers of LPG to voluntarily renounce their subsidies, which would free up resources for targeting subsidies to poor households. Following this announcement,

almost two million households surrendered their rights to receive subsidies on LPG cylinders. Such measures have had some success in ameliorating the disparities that currently exist in securing access to LPG for all households in India. In the same vein, and as highlighted in the previous section, the Indian government launched the PMUY initiative in the following year, aimed at distributing free connections to rural households living below the poverty line (Jain et al., 2018).

In the Indian context, social networks and groups also play a crucial role in determining people's behavior in many facets of life. While the evidence on the relationship between social networks and the adoption of LPG is scarce, a gap that we contribute to fill with this paper, there is abundant literature on the role of social networks in the adoption of other types of behavior and technologies in India. We shortly describe here relevant literature related to the following social networks, which leads to them being considered in surveys and in this paper: agricultural cooperatives, business associations, development groups/non-governmental organizations (NGOs), credit and saving organizations, women's groups, religious and social groups, and caste associations.

Agricultural cooperatives in India have often been shown to help achieve efficient production scales, and solve market failures arising from indivisibilities and collective action problems. For instance, evidence from Maharashtra suggests that local elites that assembled themselves into cooperatives and bought sugarcane grown by several small-scale producers actually protected the latter against exploitation, while allowing for efficient management of the sugarcane mills (Thorp et al. 2005, Attwood and Baviskar 1988). Business associations are also linked to higher efficiency, for instance by providing enforcement of property rights when formal institutions are lacking, allowing entrepreneurs to engage in profitable activities in under-developed regions (Fisman and Khanna, 2004). Membership in credit and saving organizations, such as microfinance groups, has been shown to contribute to economic development in a wide range of contexts, including India. For instance, Feigenberg et al. (2013) show with a randomized controlled trial realized in West Bengal that (randomly) assigning individuals to "repayment groups" leads

to more interactions and a higher willingness to pool risk, and a much lower likelihood of default.

Women's groups are shown, in a recent program evaluation study in Andhra Pradesh (Deininger and Liu, 2013), to lead to a positive impact on consumption, nutritional intake, as well as asset accumulation. Development groups are shown, among other outcomes, to increase immunization uptake, by spreading information about the benefits of vaccination and by creating social pressure leading more families to have their children fully immunized (e.g. Vikram et al. 2012 in the medical literature). Religious groups are known to affect many aspects of life in India, including the consumption decisions with respect to food (Filippini and Srinivasan, 2018). Caste associations can also be very influential, in several ways. One of them is explored in Barnhardt et al. (2017), who study the effect of a lottery providing improved housing on the quality of life of slum households. While the lottery improved housing conditions, it did not improve other outcomes, and also led to social isolation, because of the caste structure. The authors find that to avoid isolation, most winners ended up renouncing their wins. These are just a few examples on how social networks have played a crucial role in determining several types of economic outcomes in the Indian context, including health outcomes.

### **3 Data and Methodology**

#### **3.1 Data**

In this section, we describe the data and methodology used to test our main hypotheses about the presence of social spillovers in the adoption of LPG by Indian households.

In this paper, we adopt both a historical perspective, examining the adoption of LPG over several decades, and a more contemporaneous perspective, using the most recent available data. We also analyze mechanisms, focusing in particular on social networks. Overall, the objective of the approach adopted in this paper is to provide multiple strands

of evidence on the role of social spillovers in incentivising Indian consumers to adopt LPG.

Specifically, we use three sets of data for the empirical analysis. The first dataset that we employ is that of the National Sample Survey (NSS) of India, which is published by the National Sample Survey Organisation (NSSO), a subdivision of the Ministry of Statistics and Program Implementation of the Indian government (National Sample Survey Office, 2019). The NSSO has been conducting consumer expenditure surveys (CES) on an annual basis (barring some years) since 1983, thereby providing repeated cross sections. Each sample frame is designed to be representative, and comprises households in both the rural and urban areas of the country. The surveys include detailed expenditure data on food items, clothing and footwear, durables, medical and educational expenditure, and other items of daily use such as cooking and lighting fuel.

The NSSO conducts "thick" rounds of the NSS at a frequency of approximately every five years, whereas in the interim, "thin" rounds are conducted, wherein a smaller sample of households is surveyed. The thick rounds that are included in our analysis are the 43<sup>rd</sup>, 55<sup>th</sup>, 61<sup>st</sup> and 66<sup>th</sup> rounds of the surveys (corresponding to the years 1987-88, 1999-00, 2004-05 and 2009-10).

In the empirical estimations, we only use the thick rounds of NSS data to ensure that the sample size is sufficiently large to provide ample geographical heterogeneity in the data, as well as go far back in time to allow for a sufficient historical perspective. The NSS data allow us to attribute to each household the district and the state of residence. In addition, the data provide us with coded information for the urban block or village to which each household belongs. Based on this information, we are able to ascertain which households reside in the same village or urban block, without having to know the exact location of their residence (which is undisclosed due to data privacy concerns).

The second database is the Access to Clean Cooking Energy and Electricity Survey of States (ACCESS), conducted recently by the Council on Energy, Environment and Water in collaboration with Columbia University (Aklin et al., 2016). This is one of the largest surveys on energy access conducted in India, covering more than 8,500 households in six

states (Bihar, Jharkhand, Madhya Pradesh, Odisha, Uttar Pradesh and West Bengal). It is one of the first surveys to provide extensive information on the affordability, availability, duration of use, as well as levels of satisfaction with LPG use in general. Additionally, and in contrast to the NSS, this sample entirely comprises relatively homogenous rural households, thus adding an additional interesting layer to our analysis. This dataset covers 2015, thus providing a very recent perspective on the adoption of LPG among rural households in India.<sup>2</sup>

The third database is from the India Human Development Survey (IHDS), compiled by the University of Maryland and the National Council of Applied Economic Research (Desai and Vanneman 2005, Desai and Vanneman 2011). It is a panel dataset, with two rounds of data available, covering 2005-06 and 2011-12. 83 % of the households sampled in the first round also responded in the second round. The panel nature of the data enables us to track changes in LPG adoption over time. This dataset, composed of about 40,000 households, thus complements the larger cross sections of the NSS, and the recent cross section from ACCESS, while allowing us to control for unobserved heterogeneity.

In the NSS and IHDS data, households are asked detailed questions about their expenditure on items over a "reference period", which is defined by the questionnaire for each item. The reference period often varies across items. For instance, for fuel-related expenditures, in most rounds households are asked for expenditure over the previous 30 days.<sup>3</sup>

The NSS and IHDS contain information on the demographic characteristics of all individuals of a household, including the age, gender, and level of education, whereas the

---

<sup>2</sup> A second round has been conducted in 2018, but the data are not yet ready to be made available to researchers.

<sup>3</sup> The 66<sup>th</sup> round of the NSS comprises two sub-rounds of surveys, which differ in terms of the recall period for some of the items purchased; this was done by the NSS to investigate whether there is a tendency for households to underreport expenditures with a longer recall period. For instance, the first type of data in the 66<sup>th</sup> round uses a recall period of 30 days for food, beverages, and tobacco expenditures, while the second type of data uses a recall period of 7 days for expenditures on the same items. To ensure comparability with the other rounds, we only use observations for which the 30-day window was used.

ACCESS database contains information aggregated to the household level. Information is also provided on land ownership (total land possessed, whether land is rented, irrigated, etc.), the physical characteristics of the house (such as the type of structure, the condition of the house, type of floor, etc.), prices of fuels such as LPG and kerosene, ownership of durables (such as cookstoves) as well as access to electricity.<sup>4,5</sup>

The measure of household-level LPG adoption, and our dependent variable, for the models using NSS data is a binary variable for whether LPG is the primary cooking fuel of a household or not. This information is particularly useful, given that, as mentioned, fuel stacking is commonly observed among households in India, where multiple fuels are used at the same time. For the IHDS and ACCESS datasets, our measure of LPG use is represented by whether the household spent on LPG in the last 30 days, and whether the household uses LPG for cooking purposes, respectively. In the IHDS data, we restrict the sample to the households for whom we have valid information on whether or not they spent on LPG in the last 30 days, and among those that spent on LPG, those that primarily use the fuel for cooking purposes (instead of heating, or other purposes). This eschews the possibility of irregular use of LPG, which is important given our emphasis on regular LPG use, rather than on mere LPG adoption, in line with the findings of Hanna et al. (2016) and the current policy agenda in India.

Table 1 provides information on the mean, maximum, and minimum number of households by village or urban block across our datasets, and the mean LPG adoption rate (at

---

<sup>4</sup> While electricity is not required for using a cookstove with an LPG cylinder, this variable is used as a proxy for economic development, which could in turn enable access to LPG. The urban centers that are chosen are the state capitals, and the tier-I and tier-II cities of the country (where a tier-I city is defined as a city with population above 4 million, while a tier-II city is defined as one with population between 1 and 4 million).

<sup>5</sup> The NSS data do not directly provide a variable for the price paid by consumers to purchase LPG. We derive it by dividing a household's expenditure on LPG by the quantity of LPG purchased by the household. However, this can only be observed for households that actually purchased LPG in the last 30 days, which may be a small fraction of households for several sub-samples. In order to estimate this variable for other households, we follow the procedure outlined by Kumar and Viswanathan (2007), compute the average price in the district, and attribute this as a measure of price for the households for which this information is not available.

Table 1: Sample Size and LPG Adoption Rate by Village/Urban Block

Data source	NSS				IHDS	ACCESS
Round	43	55	61	66 (Type 1)	Overall	Overall
Year	1987-88	1999-00	2004-05	2009-10	2005-06 and 2011-12	2015
Households sampled in village/urban block Mean	9.95	11.96	9.98	7.98	33.88	12
Households sampled in village/urban block: Min.	2	2	3	2	4	11
Households sampled in village/urban block: Max.	10	12	10	8	88	12
LPG adoption rate at village/urban block level: Mean (%)	10.5	25.82	29.79	41.24	65.57	21.6
Observations	104874	103094	97998	67374	43179	8563

Notes: Values reported are calculated only for observations included in the regression sample.

the village/urban block level). We observe that while at the overall level, adoption rates of LPG have increased over time (across the rounds of the NSS as well as the IHDS), they are still relatively low in the rural areas covered by the 2015 ACCESS dataset. While the sample size per village or urban block varies across datasets, it is fixed at around 10 in case of the 43<sup>rd</sup>, 55<sup>th</sup>, and 61<sup>st</sup> rounds of the NSS, 8 in the 66<sup>th</sup>, and at 12 in the case of the ACCESS data. Thus, there is meaningful variation in sample size per village/urban block only in the IHDS data. In our empirical analyses, we leverage the fact of having access to multiple datasets, and use the IHDS data to run sensitivity tests by restricting the sample to only villages and urban blocks having a certain minimum number of households above varying thresholds.

Figure A1 in Appendix A provides a historical perspective on the distribution of households by primary fuel-type (for cooking purposes) in the four thick rounds of the NSS. As is clear from these graphs, firewood was in the 1980s, and still is to a large extent in the most recent round, the primary cooking fuel for a majority of the households. The popularity of LPG has increased over this period, and in 2010 it was the second most popular cooking fuel used by households. Kerosene has also gained in popularity in recent years, primarily in urban areas. Dung cakes have gained popularity in rural areas.

Figures A2 and A3 focus on LPG and plot the evolution of the proportion of households for whom LPG was the main cooking fuel, and show how it has gained popularity, especially in the 2000s. We also observe that during the period covered by the NSS data, the pace of increase in adoption has been much faster in urban areas, thus leading to a

much larger share of LPG users in urban areas than in rural areas.

Figure A4 shows the regions which have contributed most to the increase in adoption of LPG. In 1987, the highest proportion of LPG users were in Delhi and the "union territories" of Goa, Chandigarh, and Daman and Diu, which are all primarily urban areas. Over time, some of the bigger (and richer) states, such as Maharashtra, Tamil Nadu, and Karnataka, experienced an increase in the share of LPG adopters.

Tables 2 and 3 provide summary statistics on some of the demographic characteristics of the households using NSS data, and the ACCESS and IHDS data, respectively. From Table 2, it is clear that the proportion of households using LPG as the primary cooking fuel has been increasing over time (from 10 % in 1987-88 to 39 % in 2009-10). Simultaneously, the proportion of households that have access to firewood has declined over the same period from 73 % to 63 %. Another interesting finding to emerge from this table is the drop in the percentage of households that bought a new cookstove in the last 365 days: it declines from around 4 % in 1987-88 to just about 1 % in 2009-10. This may be attributed to the S-shaped nature of the diffusion process: with time, more households would have already bought cookstoves, thus the proportion of households that buy new cookstoves can be expected to decline. We find from Table 3 that levels of education, household income, and electricity are lower for respondents in the ACCESS dataset. This is primarily because this sample comprises only rural households. Moreover, and consistently with the previous observation, the household size on average is larger, and fewer households have access to electricity.

The statistics presented in Table 3 suggest much higher rates of adoption of LPG compared to those in Table 2: the differences between the NSS data and the IHDS data can be attributed to the difference in the measure of adoption. In Table 3, we are considering whether a household spent on LPG, and not necessarily whether they used it as the primary cooking fuel, which is the reason for the difference in values in these two tables. Households have been found to often spend on LPG, without using it as the primary cooking fuel (Cheng and Urpelainen, 2014).



Table 2: Summary Statistics of NSS Data

Round Year Variable	43 1987-88			55 1999-00			61 2004-05			66 2009-10		
	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs
LPG as primary fuel type (%)	10	29	127,932	23.94	42.67	119,639	28	45	124,637	39	49	100,845
Proportion of rural population (%)	64.65	47.81	118,205	59.40	49.11	119,638	63.62	48.11	124,637	58.62	49.25	100,845
HH size	5.1	2.74	127,932	4.99	2.65	119,638	4.89	2.52	124,637	4.65	2.34	100,845
Monthly per capita expenditure (Rs./month)	230.31	285.53	118,299	757.74	1059.2	119,639	851.51	1160.53	124,637	1527.90	1525.02	100,845
Age of head of households (years)	43.72	13.97	127,932	46.65	11.58	104,540	49.19	10.56	99,156	52.09	9.35	67,882
Whether HH head is female (%)	10	30	127,932	8.8	28.33	119,639	9	29	124,637	8	28	100,845
Whether HH head has at least primary education (%)	41	49	128,026	44.41	49.69	119,639	46	50	124,637	33	47	100,845
Whether HH lives in a district adjoining a district with big urban center (%)	69.78	45.92	118,299	68.42	46.48	119,639	64.63	47.81	124,637	65.42	47.56	100,845
Whether HH has access to electricity (%)	43	49	128,031	64.41	47.88	119,639	73	45	124,637	83	38	100,845
Whether HH purchased a cookstove in the last 365 days (%)	3.69	18.85	118,650	8.47	24.85	119,638	0.29	5.39	124,637	0.6	7.66	100,845
Average price of LPG (Rs./kilogram)	15.62	14.38	113,664	12.77	2.02	117,984	21.12	7.22	124,367	23.82	2.56	100,781
Average price of kerosene (Rs./kilogram)	2.83	0.62	128,039	4.28	20.26	119,626	11.93	68.43	123,638	10.80	1.65	100,079
Whether the HH has access to firewood (%)	73	44	128,039	60.84	48.81	120,127	68	46	124,637	63	48	100,845

Table 3: Summary Statistics of IHDS (2005-06 and 2011-12) and ACCESS (2015) Data

Year Variable	IHDS 2005-06			IHDS 2011-12			ACCESS 2015		
	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs
Whether HH spent on LPG in the last 30 days (%)	59.5	49.1	22703	99.8	3.41	22781	21.61	41.16	8566
Whether HH has access to electricity (%)	95.9	19.8	20717	99.99	0.01	22781	65.86	47.42	8563
Proportion of rural population (%)	60.6	48.9	22703	58	49.4	22781	100	0	0
HH size	5.79	2.95	22703	4.83	2.29	22781	6.74	3.53	8563
Number of years of education for HH head (IHDS)/ Level of education of HH head (ACCESS)	8.72	4.9	22673	9.40	4.97	22772	1.30	1.22	8563
HH income (Rs./year, IHDS)/ HH consumption expenditure (Rs./year, ACCESS)	67506.37	97836.31	22703	155057.3	261935.1	22781	63619	46816	8563

Notes: Levels of education in the ACCESS dataset are defined as 1: no formal schooling, 2: up to 5<sup>th</sup> grade, 3: up to 10<sup>th</sup> grade, 4: 12<sup>th</sup> grade or high school diploma, 5: graduate or above. Income is measured on a monthly basis in the ACCESS dataset, statistics above have been scaled up by a factor of 12 to calculate the annual income.

Table 4: LPG Adoption Rates by Social Network (IHDS)

Type of groups Subsample	LPG Adoption Rates					
	Overall		Rural		Urban	
	Rate	Obs.	Rate	Obs.	Rate	Obs.
Women's groups	0.826	4,172	0.759	2,767	0.957	1,405
Youth and sports associations	0.837	2,160	0.721	1,102	0.957	1,058
Unions and business groups	0.897	2,718	0.808	968	0.947	1,750
Credit and savings associations	0.822	4,426	0.744	2,871	0.964	1,555
Religious and social groups	0.764	6,789	0.656	4,103	0.931	2,686
Caste associations	0.718	4,720	0.608	2,994	0.909	1,726
Developmental groups and NGOs	0.968	2,673	0.948	1,577	0.997	1,096
Agricultural groups and cooperatives	0.638	1,331	0.545	1,018	0.939	313

Notes: Average LPG adoption rates per group are based on both periods in our panel.

Finally, Table 4 provides descriptive statistics, in terms of overall adoption and adoption by rural and urban households, respectively, for each social network considered in our paper. Based on the classifications available in the data, we consider the following social networks: women's groups, youth and sports associations, unions and business groups, credit and savings associations, religious and social groups, caste associations, developmental groups and NGOs, and agricultural cooperatives. These social networks have been found to be important in influencing household consumption expenditure in the sociological literature (Vanneman et al. 2006, Arun et al. 2016), as well as in the economic literature (as highlighted in the previous section). As indicated by Table 4, there is a fair amount of variation in LPG adoption among groups, especially in the rural context.

## 3.2 Methodology

In this paper, our main objective is to understand whether social spillovers have been relevant in facilitating the adoption of LPG among Indian households. To this effect, we adopt a multi-pronged approach by providing evidence from both cross-sectional and panel datasets on the existence of these spillovers. In this section, we shortly distill our main hypotheses and describe the empirical methodology that we adopt to test them.

Our first hypothesis on the role of (positive) social spillovers in the adoption of LPG in India implies that we should observe higher adoption of LPG in areas wherein adoption is already relatively high. In this respect, we refer to the recent and abovementioned empirical literature on the role of spillovers in new technology adoption, especially those related to agricultural and health innovations in developing economies (Conley and Udry 2010, Bandiera and Rasul 2006, Kremer and Miguel 2007, Oster and Thornton 2012) but also to green technologies in developed countries (Bollinger and Gillingham 2012, Graziano and Gillingham 2015, Baranzini et al. 2017).

Our second hypothesis posits that close social interactions (proxied for instance by participation in social networks) are more likely to lead to stronger flows of information across households. As noted in the previous section, in the Indian context, social networks and groups play an important role in the organization of society. Many other developing countries have similar social structures, which have been found to influence the flows of information across households (Bandiera and Rasul 2006, Masera et al. 2000, Oster and Thornton 2012). We expect that certain social networks are likely to have strong spillover effects with respect to LPG adoption, especially those where flows of information related to LPG are more likely to be stronger, such as women's groups, or developmental groups/NGOs, whereas we expect other networks to have weaker spillovers related to LPG use, such as youth or sports groups, or credit and savings organizations. By looking at social networks, we also link to a broader literature examining the role of social capital and membership in social organizations on important outcome variables such as

prosperity and economic development (e.g. Putnam et al. 1994, Knack and Keefer 1997).

We start with our first hypothesis. To determine whether spillovers have been prevalent in LPG adoption, we would need to estimate a model of the following form:

$$A_i = \alpha_0 + \alpha_1 A_{-ij} + \alpha_2 X_i + \mu_i \quad (1)$$

where the dependent variable is denoted by  $A_i$ , a binary variable indicating whether LPG is the primary cooking fuel of household  $i$  (or whether the household spent on LPG in the last 30 days, for the estimations using IHDS data), and the main independent variable is  $A_{-ij}$ , the average LPG adoption rate among all households (other than household  $i$ ) in village/urban block  $j$ .  $X_i$  denotes household-specific controls (such as household size, age, gender and the level of education of the head of the household, whether the household has access to electricity, etc.)  $\mu_i$  denotes the stochastic error term. Standard errors are clustered at the village/urban block level, in order to control for the possibility of errors being correlated across geographical units. Model (1) takes this form for the estimations that use cross-sectional data. In the case of the models estimated using the IHDS data, all variables described in the previous paragraph have a "t" subscript denoting the relevant time period. Moreover, we use household and year fixed effects to control for time-invariant unobserved heterogeneity.

A "naive" regression will not suffice to estimate the effect of social spillovers in the adoption of LPG in India. There are several reasons why there may be potential threats to identification with the estimation of a simple linear probability model due to endogeneity, in particular in relation to the problem of "reflection" or "simultaneity" (Manski 1993, Manski 2000, Moffitt 2001). When studying peers, it may indeed be hard to isolate the effect of agent  $i$  on agent  $j$ , independent of the effect of agent  $j$  on agent  $i$ . Endogeneity may also arise from "homophily" (a form of geographical sorting according to which households with similar preferences cluster together), and common unobservable factors affecting both left and right-hand side variables.

In consistency with the existing literature, we address the issue of endogeneity as follows; first, we take advantage of the very rich data available and make use of a large set of controls in all of our specifications. Second, we systematically combine analyses using both cross-sectional and panel data. In the latter, household-specific fixed effects allow us to capture unobserved heterogeneity across households. Third, and most importantly, we use an instrumental variable (IV) linear probability model (LPM) approach.<sup>6</sup>

We use two different sets of instrumental variables. Recall that three different datasets are used in this study, which implies that our instrumental variable approach needs to be adapted to the available data. Both sets of instrumental variables provide us with strong first stages and pass the standard overidentification tests in all specifications used in this study.

Our first set of instrumental variables (for the endogenous variable capturing average LPG adoption at the village or urban-block level) is the proportion of the sample from each village (or urban block) belonging to the highest monthly per capita expenditure (MPCE) deciles. MPCE is found to be an important determinant of the choice of a household to adopt LPG, thus the average LPG adoption rate in the village or urban block is likely to be highly correlated with the proportion of the population that belongs to the highest MPCE decile. The exogeneity restriction works in a very similar way to the study by Duflo and Saez (2002), which was in turn based on the estimation strategy developed by Case and Katz (1991). Duflo and Saez (2002) study whether there are peer effects among colleagues in the same department of a university in participation in retirement plans, and find that both the choice of employees to enroll in these plans, and the choice of vendor, were influenced by the decisions made by colleagues (belonging to the same department).

To causally assess the existence of peer effects, the authors instrument average par-

---

<sup>6</sup> We opt for IV-LPM over IV-Probit to ensure that our standard errors are correctly measured in all specifications. Given the incidental parameter problem that non-linear models such as the Probit model face when using household-specific fixed effects, our IV-LPM methodology also allows for immediate interpretation and ensures consistency across models. Qualitatively, the results obtained using IV-Probit are the same as those presented in this paper. All additional tables are available by the authors upon request.

ticipation in each peer group by the salary or tenure structure of that group, while controlling for individual wages or tenure. The underlying assumption is that once wage and tenure are controlled for in the second-stage model, the average wage or tenure in the department is unlikely to directly affect individual enrollment decisions. Similarly, the underlying assumption in our context is that once we control, in the second stage, for factors such as income, education, access to electricity and to LPG, the proportion of the population belonging to the highest income deciles is unlikely to directly affect a household's decision to use LPG. The underlying assumption in Duflo and Saez (2002) would not hold if employees were to choose to work for particular departments because of the choice of retirement plans, which is however, highly unlikely. In their study, departments are very similar in many dimensions, especially in terms of salary and years of service of employees, while the participation rates in the retirement plan differ significantly across libraries. In our context, after controlling for factors determining preferences and access to LPG, it is also unlikely that a household's decision to locate (and live) in a village/urban block depends on the proportion of population belonging to the higher income deciles in that village that use LPG.

The second set of instrumental variables that we use are the share of households in the village/urban block in our sample that have highly educated women (at least a university degree), as well as the average size of land holdings at the village/urban block level. The rationale for using the share of households with highly educated women in our second set of instruments is as follows: first, the use of proportions as instruments follows from Duflo and Saez (2002), as with the first set of instruments. Second, we leverage the fact that highly educated women may be more inclined to learn about the benefits of LPG and more sensitive to the health and environmental effects of dirtier fuels. Furthermore, highly educated women have a higher opportunity cost of time, which is also expected to direct them towards investing in (relatively more efficient) LPG cookstoves, and using them on a regular basis, as opposed to using biomass (Muller and Yan 2018, Farsi et al. 2007). Unsurprisingly, in our datasets highly educated women tend to be disproportion-

ately present in urban households. This is not an issue per se, as our analyses are at the village/urban block level and include a large set of control variables controlling for potential confounders.

In a conservative vein, however, we combine this instrument with another one, which may be more relevant for rural areas, namely the average size of possessed land holdings at the village/urban block level, which is measured in acres. Villages with larger amounts of land possessed per household comprise relatively wealthier households, as land is an important asset or source of wealth for agricultural households. In the spirit of our approach relying on income deciles, we expect higher incomes to be associated with a higher likelihood of adopting LPG in the first stage. In 2005, the Indian government initiated the Rajiv Gandhi Grameen Vidyutikaran Yojana (RGGVY) program, which was, essentially, a program that aimed to provide all villages without electricity a connection within five years. The RGGVY program was successful in the context of village electrification: rates of electrification increased from 74 % in 2005 to 91 % in 2011. However, more than two-thirds of the increase in connections to electricity supply went to rich households, and to villages that were richer. Khandker et al. (2014) show that wealthier rural households (and those that owned more land) benefited more from electrification, and households that were electrified in fact spent less time collecting firewood and other forms of biomass, and more time on productive labor and (and in acquiring education, in the case of children). In our model, we control for access to electricity, income, and land ownership in the second-stage. The underlying assumption in our model is that, once we control for the abovementioned factors, larger average land holdings are unlikely to affect a household's behavior directly, and only indirectly through other people's adoption, and thus through social spillovers.<sup>7</sup>

---

<sup>7</sup> Another possible choice of instrument in our panel data setting might have been to instrument adoption in the second wave of 2011 with adoption in the previous period, thus collapsing the data to a cross-sectional dimension. We would like to thank an anonymous reviewer for this suggestion. Given the gap of six years between both rounds, such instrument could not be used in our context because of the weak first stage (Cragg-Donald F-statistics below 2). Such an instrument could, however, be used in future research, using for instance multiple waves of the ACCESS data.

These are likely to be valid sets of instruments only if they do not directly affect households' LPG adoption rates. One way for this concern to be realized is through access to LPG or to dirtier fuels. Households may be able to use LPG, because they live in a neighborhood that is richer, and thus has better supply of LPG. Furthermore, unobserved characteristics linked to the adoption decision could also be correlated with average income at the village/urban block level, which could then have a direct effect on households' LPG adoption rates. For instance, factors like tastes and preferences for cooking methods (and/or food) could be correlated to both. In our empirical approach, we first estimate model (1) using each of the four thick rounds of the NSS data and the proportion of population in the highest MPCE deciles as instruments, while controlling for income in the second-stage estimation, exactly in line with the estimation methodology of Duflo and Saez (2002). Along with many other covariates, we use controls for religion and caste to capture preferences for cooking methods or food. We then repeat this with the ACCESS dataset, using both sets of instruments, and additionally controlling for supply of/access to LPG. In order to alleviate residual concerns with identification, we then exploit the potential of the IHDS panel data to estimate a fixed-effects instrumental variable linear probability model, while using the second set of instruments. The choice of the set of instruments is mainly driven by the variables available in each dataset.

To test our second hypothesis, we use the abovementioned approach with an additional component from the IHDS data, which provides information on the social networks and groups to which households declare to belong. In this way, we are able to offer insights on how stronger social ties are linked to spillover effects in consumption decisions. For these estimations, our endogenous variable is defined as the average LPG adoption among all households that reside in the same village or urban block (except household  $i$ ), and belong to one of the eight different networks that we consider in this paper. In this case, we do not instrument for belonging to one of the networks. While theoretically it is possible that people who use LPG join a given network, we consider it as rather implausible in the social environment that we study. Most of the networks that we consider are not new in



the Indian context, and the IHDS data suggest that membership is rather constant over time.

## 4 Results

### 4.1 Empirical Results

This section presents the results of the empirical estimations. We proceed as follows. First, we use NSS data to test our first hypothesis on social spillovers in the adoption of LPG with multiple waves of cross-sectional data covering multiple decades, which is key for obtaining a historical perspective. Second, we switch to the most recent dataset, obtained from ACCESS, and zoom in on the situation of rural households, testing the very same hypothesis with both sets of instruments. Third, we move to a panel setting, using the IHDS data, and test the robustness of our findings concerning the first hypothesis by incorporating household-specific fixed effects. Fourth, we leverage the full potential of the IHDS data and analyze the role of social networks and social groups in affecting social spillovers in the adoption of LPG, thus formally testing our second hypothesis.

The results from the NSS are presented in Table 5.<sup>8</sup> Our independent variable of interest is the average level of LPG adoption at the urban block or village level. We find evidence to support the existence of social spillovers from the data of the 43<sup>rd</sup>, 55<sup>th</sup> as well as the 61<sup>st</sup> rounds of the NSS (at the 1 % level for the 43<sup>rd</sup> and 55<sup>th</sup>, and 5 % level for the 61<sup>st</sup> round). The coefficients should be interpreted as follows: in the 43<sup>rd</sup> round, a one unit increase in the average village/urban-block LPG adoption rate increases the probability that household  $i$  adopts LPG by about 0.825 units. In other words, if the share of households in a village using LPG increases from 0 to 100 %, the probability of household  $i$  adopting LPG increases by 83 percentage points. The magnitude of the

---

<sup>8</sup> Table 5 displays only the second stage of our IV-LPM approach. Estimates from the first stage are presented in Table B2. A standard comparison between the IV-LPM approach and a naive regression model would rely on Table B1 in the Appendix, showing estimates for the simple LPM.

main coefficients of interest in Table 5 points to decreasing social spillovers over time, which is consistent with a market reaching maturity. The coefficient in the last wave is, actually, very close to zero and statistically non-significant. Note that this coefficient is mainly driven by urban households. In Table 8, which we discuss in the following section, we show that social spillovers in rural areas remain positive even in 2009-2010, which is consistent with what we find later in this section using the ACCESS data. We would obtain the same pattern with the IHDS data. This observation further confirms the intuition that rural areas lagged behind urban areas in the process of reaching market saturation.

We now shortly present an overview of the results with regards to the control variables. We find that households with better access to electricity are more likely to adopt LPG. We also find that households having heads that are older, female, or more educated are more likely to use LPG as the primary cooking fuel. On the other hand, those facing a high price of LPG, or those with access to firewood, are less likely to use LPG. Our results suggest that larger households are more likely to adopt LPG, a common yet not undisputed finding (Lewis and Pattanayak 2012 offer a discussion).

The first-stage results of Table B2 suggest that the larger the population share in the village/urban block belonging to the higher income deciles, the higher is the average rate of LPG adoption. Moreover, the signs of the coefficients of household-level attributes on average LPG adoption are in line with economic intuition, and generally consistent with the coefficients in the second stage. Further, the first-stage estimation yields consistently high Cragg-Donald F-statistics, and satisfies the overidentification tests.

Estimation results from the NSS data provide the first set of evidence on the presence of social spillovers in the historical development of the market for LPG. We now turn to the ACCESS data. Table 6 presents the estimation results using the ACCESS data.<sup>9</sup>

Odd columns (1 and 3) include the results of the estimation of the model using proportion of households belonging to the highest income deciles as instruments, whereas

---

<sup>9</sup> The first-stage results of these estimations are provided in Table B3 in the Appendix.

Table 5: NSS Estimation Results: Second Stage

Round	43	55	61	66
Year	1987-88	1999-00	2004-05	2009-10
Column	(1)	(2)	(3)	(4)
Average LPG use rate (village/urban block)	0.825*** (0.024)	0.315*** (0.025)	0.084** (0.036)	-0.070 (0.047)
Whether bordering an urban center?	-0.027** (0.012)	-0.005* (0.003)	-0.072** (0.035)	0.011 (0.057)
Whether HH had access to electricity?	0.032*** (0.002)	0.055*** (0.003)	0.045*** (0.003)	0.058*** (0.005)
Whether HH lives in a rural area?	0.039*** (0.004)	-0.051*** (0.007)	-0.088*** (0.011)	-0.125*** (0.015)
Whether HH purchased a cookstove in last 30/365 days?	-0.032*** (0.005)	-0.104*** (0.005)	-0.020 (0.022)	0.021 (0.020)
HH size	0.009*** (0.0003)	0.017*** (0.0005)	0.015*** (0.0005)	0.019*** (0.001)
Age of head of HH	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)
Whether head of HH is female?	0.004* (0.003)	0.022*** (0.003)	0.017*** (0.003)	0.021*** (0.004)
Whether head of HH is educated?	0.035*** (0.002)	0.096*** (0.003)	0.090*** (0.003)	0.113*** (0.004)
Price of LPG	(0.00004) (0.0005)	-0.007*** (0.001)	-0.0004** (0.0002)	-0.005*** (0.001)
Price of kerosene	0.005*** (0.001)	0.000005 (0.0001)	-0.00001* (0.00001)	0.002 (0.002)
Whether HH had access to firewood?	-0.115*** (0.004)	-0.233*** (0.005)	-0.431*** (0.007)	-0.507*** (0.010)
Observations	104845	102994	97963	67374
R-squared	0.5183	0.6727	0.7106	0.7666
First-stage Cragg Donald F-Stat	1756.389	1399.271	923.55	309.779
Hansen J-Statistic	4.779	4.435	12.38	3.016
P-value	0.19	0.22	0.01	0.39

Notes: The proportion of population in the same village or urban block in the four highest income deciles are used as instruments. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in Table B2 in the Appendix). All specifications include dummies for MPCEs, and for belonging to districts, religion, and castes. Standard errors are clustered at the village/urban block level (reported in parentheses). \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported. The variable "Whether HH purchased a cookstove in last 30/365 days" includes by design expenditure on repairs in the 55<sup>th</sup> round.

even columns (2 and 4) present the results of the estimation using proportion of households having highly educated women, as well as the average size of the landholdings in the village/urban block, as instruments. In columns 1 and 2, we include as a control a dummy variable for whether the household did not have access to LPG (namely, if they had to travel too far, or if it was unavailable) and use the entire sample for estimation purposes. In columns 3 and 4, we restrict the sample to only those households that had access to LPG and re-estimate the models using both sets of instruments.

We find that the coefficient on the variable measuring village level adoption rate of LPG is positive, and significant at the 1 % level using both sets of instruments. In fact, we find that the coefficients in columns 3 and 4 are larger in magnitude than those in columns 1 and 2 respectively, and that they are persistently positive and significant, highlighting that social spillovers might have been prominent in determining household LPG consumption decisions, even after accounting for its availability among households. Moreover, the coefficients are larger when we use education and land as instruments, than when we use the income deciles. This is likely to be due to how either set of instruments capture variation in the adoption of LPG. As with the previous set of estimations using NSS data, the control variables have the expected signs. The first-stage results in Table B3 are intuitive, and in line with what we observe for the NSS data, and the models exhibit consistently high Cragg-Donald F-statistics and satisfy the overidentification tests. Given the exclusive focus on rural areas of the ACCESS dataset, in even columns most of the variation in the first stage is explained by the land ownership instrument, which, as expected, is associated with higher rates of LPG adoption. The combination of both instruments is, however, crucial in allowing us to conduct overidentification tests.

We now turn to the IHDS data and to our panel specification. The relevant estimations are provided in Table 7.<sup>10</sup> Column 1 presents the results of the baseline model, where we take advantage of the panel structure of the data and incorporate household and

---

<sup>10</sup> The first-stage results of these estimations are provided in Table B4 in the Appendix.

Table 6: ACCESS Estimation Results: Second Stage

Dep.Var.: Whether HH $i$ uses LPG for cooking Column	Income deciles (1)	Education (women) and land (2)	Income deciles (3)	Education (women) and land (4)
Average LPG use rate (village/urban block)	0.285*** (0.078)	0.598*** (0.173)	0.400*** (0.160)	0.670*** (0.274)
Whether HH had access to electricity?	0.077*** (0.008)	0.074*** (0.009)	0.180*** (0.021)	0.174*** (0.022)
HH size	-0.002*** (0.001)	-0.002** (0.001)	-0.005*** (0.002)	-0.005*** (0.002)
Number of years of education of household head	0.042*** (0.003)	0.041*** (0.003)	0.074*** (0.007)	0.074*** (0.007)
Whether HH owns land?	0.035*** (0.009)	0.043*** (0.009)	0.093*** (0.018)	0.100*** (0.019)
Whether HH does not have access to supply of LPG?	-0.454*** (0.013)	-0.449*** (0.013)		
Observations	8563	8563	3603	3603
First-stage Cragg Donald F-Stat	274.76	65.03	96.59	47.47
Hansen J-Statistic	2.76	1.95	2.22	0.49
P-Value	0.43	0.16	0.53	0.48

*Notes:* The proportion of population in the same village or urban block in the four highest income deciles are used as instruments in odd columns. Average land ownership and the proportion of educated women in the same village or urban block are used as instruments in even columns. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in Table B3 in the Appendix). All specifications include dummies for MPCEs, caste, religion and sub-districts. \*,\*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

year fixed effects to control for unobserved heterogeneity. We find evidence consistent with our previous estimations on the existence of social spillovers in LPG use. That is, the results obtained from using panel data reinforce the picture provided by the cross sections using NSS and ACCESS data. The results regarding the control variables and the first stage are again as expected, including consistently high Cragg-Donald F-statistics and estimates satisfying the overidentification tests. Recall that in our IHDS specifications we use average land ownership in a village/urban block and the proportion of educated women as instruments.

After having determined, using three different datasets, that social spillovers tend to drive adoption of LPG, we focus on another feature of the IHDS data, which is that it allows us to look at specific social networks, namely the ones described in the previous section. In column 2 of Table 7, we estimate model (1) for all households who declare that they do not belong to any of the social networks in our list. The endogenous variable, thus, is the average level of LPG adoption among all other households in the same village, or urban block, for this specification. In columns 3-9, we estimate separate models for each of the social networks in our data. As with previous estimations, the variable of interest is the average level of LPG adoption, for a given village, or urban block, but

among the members of specific social networks. According to our hypothesis, the strength of the spillovers may be higher among households belonging to social networks, given that they are likely to experience deeper, and more persistent interactions among themselves. However, some social networks may witness weaker spillovers, either due to especially strong (or traditional) values, or limited information sharing on cooking fuels.

We find evidence to suggest that there may be strong social spillovers among households that belong to social networks. The magnitude of the coefficient for average LPG adoption is 0.973 in column 2, the benchmark, and is significant at the 1% level. In columns 3 to 9, we find most coefficients to be larger than the benchmark. Households that belong to women's associations, unions or business groups, religious and social groups and development groups/NGOs are found to experience particularly strong social spillover effects. It is reasonable to assume that members of these groups are likely to facilitate knowledge sharing on new technologies, so that social networks may contribute towards addressing some of the information asymmetry around the use of LPG. In contrast, and in line with our expectations, we find lower social spillovers, everything else equal, for households who belong to youth or sports groups, where we can expect that information exchange on cooking fuels may be weaker.

The results of columns (1) and (10) suggest that households with access to electricity, and with heads that are more educated, are more likely to adopt LPG, whereas larger households are less likely to do so. The corresponding first-stage results, as with the results using ACCESS data, suggest that the average rate of LPG adoption is higher in villages (or urban blocks) where households own larger plots of land, while controlling for electricity access, ownership of land, and income in the second stage. Moreover, we also find evidence to suggest that the average rate of LPG adoption is strongly correlated with the proportion of households having highly educated (graduate level and above, or those with technical education) women, which is as expected. Also with these specifications, we obtain consistently high Cragg-Donald F-statistics and estimates satisfying the overidentification tests.

Table 7: IHDS Estimation Results: Second Stage

Model Column	Overall (1)	No network (2)	Women's groups (3)	Youth and sports groups (4)	Unions or business groups (5)	Credit and savings groups (6)	Religious or social groups (7)	Caste associations (8)	Developmental groups/NGOs (9)	Agricultural cooperatives (10)
Average LPG use rate (village/urban block)	0.944*** (0.009)	0.973*** (0.015)	1.043*** (0.040)	0.925*** (0.074)	1.138*** (0.114)	0.983*** (0.030)	1.089*** (0.035)	1.024*** (0.039)	1.227*** (0.147)	1.088*** (0.083)
Whether HH lives in a rural area?	0.006 (0.007)	-0.048*** (0.020)	0.112 (0.086)	0.084 (0.022)	0.032 (0.071)	0.0373 (0.041)	0.003 (0.030)	0.026 (0.084)	-0.149*** (0.043)	0.17 (0.140)
Whether HH had access to electricity?	0.114*** (0.014)	0.095*** (0.017)	0.137*** (0.064)	0.160 (0.155)	0.026 (0.121)	0.166*** (0.063)	0.009 (0.045)	0.110* (0.057)	0.304*** (0.138)	0.542 (0.356)
HH size	-0.001 (0.001)	-0.002 (0.002)	-0.003 (0.002)	-0.003 (0.003)	-0.003 (0.003)	0.003 (0.003)	-0.007*** (0.003)	-0.001 (0.003)	-0.001 (0.005)	-0.003 (0.007)
Number of years of education of HH head	0.002*** (0.001)	0.003*** (0.001)	0.002 (0.002)	0.003 (0.003)	0.002 (0.003)	-0.001 (0.002)	0.001 (0.002)	0.003* (0.002)	0.004 (0.004)	0.006 (0.006)
Whether HH owns land?	0.006 (0.006)	0.008 (0.009)	0.027*** (0.014)	0.006 (0.015)	-0.011 (0.020)	-0.004 (0.021)	0.031** (0.013)	-0.003 (0.014)	0.004 (0.031)	0.063 (0.040)
Observations	41090	17576	9582	3964	5622	8550	12166	7868	3272	1476
First-stage Cragg Donald F-Stat	3196.60	1180.173	647.24	186.45	157.93	807.91	710.39	605.90	149.94	131.42
Hansen J-Statistic	0.07	0.96	6.26	1.05	0.06	0.97	0.16	0.85	0.22	0.05
P-value	0.79	0.33	0.01	0.31	0.81	0.33	0.69	0.36	0.64	0.82

Notes: Average land ownership and the proportion of educated women in the same village or urban block are used as instruments. All specifications include household-specific and year fixed effects. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first stage results are provided in Table B4 in the Appendix). All specifications include dummies for income deciles, caste and religion. \*, \*\*, and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

## 4.2 Robustness tests

In this section, we provide the results of some robustness checks. We start with sensitivity tests on the number of households per village/urban block. Our methodology relies on survey data to identify social spillovers in the adoption of LPG in India. To this end, we leverage the fact that household surveys in India tend to be very large. Even so, since our analyses are done at the village/urban block level, for some geographical units, the inference may rely on a relatively limited number of households. Hence, we run our baseline estimation for the IHDS data, with different thresholds of households interviewed per village. We use the following thresholds: 10, 20, and 30 households per village. These thresholds are only applied to the IHDS data, as the sample size per village or urban block is fixed for both the NSS and ACCESS data (at around 10 in case of the 43<sup>rd</sup>, 55<sup>th</sup>, and 61<sup>st</sup> rounds of the NSS, 8 in the 66<sup>th</sup>, and at 12 in the case of the ACCESS data.)

In columns 1-3 of Table B5 in the Appendix, we re-estimate the model of column 1 of Table 7, i.e. the baseline model of the IHDS, using different restrictions for the sample size per village.<sup>11</sup> In column 1, we only include those observations for households residing in village having at least 10 households, in column 2 at least 20 households, and in column 3 at least 30 households. We find that the results of column 1 in Table 7 are broadly corroborated across different sample sizes. We further note that this conclusion would also apply for the estimations using information about social networks.

We then proceed to additional robustness tests, leveraging the panel property of the IHDS data. In Table B6 in the Appendix, columns 1 and 2, we estimate the panel IV-LPM of column 1 of Table 7, but this time use village-specific fixed effects, and village-by-year interactions, respectively. Village-specific, rather than individual-specific, fixed effects capture time-invariant characteristics of the village (or urban block), which may relate (for instance) to the supply of LPG. General trends may also interact with these characteristics, hence the use of the village-by-year dummy variables in column (2). Table

---

<sup>11</sup> The first-stage results of these estimations are provided in Table B8 in the Appendix.



B6 shows that our main findings are robust to these additional specifications.

Another concern with the IHDS results may be that they are being influenced by measurement error in variables that can be expected to not change significantly over time for heads of households, such as income, education, religion or caste. Given that the two periods of the sample are six years apart, it is, however, reasonable to expect income and education to potentially vary between the two periods. India experienced significant growth in the period 2005-2012. For instance, its GDP per capita increased from 31,746 Indian Rupees in 2005 to 78,729 Indian Rupees (at current prices) in 2012 (World Bank Data, 2019). This corresponds to an increase of 148 %. In our dataset, household income increased by about 130 % during this period, on average. Thus, the increase in household income was concomitant with the increase in national income during this period. We also expect education levels for households in our sample to have increased over time, given that the age composition of households is relatively young, which is confirmed by our data. Figure A5 in the Appendix, for instance, shows the age distribution of household children in 2005. In 2005, the enrollment rates in schools and universities over all members of the household was 56 % (including heads of household, their spouses, children in-law, grandchildren, siblings, children as well as other relatives). Hence, it is reasonable to expect that the education variable varies over the 2005-2011 period. However, other variables, such as religion or caste, are less likely (or unlikely) to vary from one period to another. Hence, in Table B7 in the Appendix we present the results of a sanity check to ensure that the results of the IHDS data are not driven by measurement error in religion or caste. Note that these variables vary over time for only 1.4 % of the households in our dataset. Hence, unsurprisingly, our main results, originally presented in column 1 of Table 7, are unaffected when we remove these households from our estimations. Note that in this new estimation we can no longer control for caste or religion, as this variable is perfectly collinear with the household-specific fixed effects.

Table 8: NSS Estimation Results on Rural and Urban Households: Second Stage

Year	1987-88	1999-00	2004-05	2009-10				
Round	43	55	61	66				
Dep.Var:	Whether HH <i>i</i> is LPG (NSS) or whether HH <i>i</i> spent on LPG in the last 30 days (IHDS)							
Column	Rural (1)	Urban (2)	Rural (3)	Urban (4)	Rural (5)	Urban (6)	Rural (7)	Urban (8)
Average LPG use rate (village/urban block)	0.514** (0.215)	0.637*** (0.034)	0.209** (0.104)	0.209*** (0.030)	0.255*** (0.086)	0.080** (0.038)	0.162* (0.095)	0.007 (0.049)
Whether bordering an urban center?	-0.017 (0.035)	0.029 (0.019)	0.002 (0.003)	-0.014* (0.007)	-0.034 (0.035)	-0.134*** (0.031)	0.030 (0.077)	-0.050 (0.064)
Whether HH had access to electricity?	0.017*** (0.002)	0.046*** (0.005)	0.047*** (0.005)	0.099*** (0.007)	0.033*** (0.004)	0.084*** (0.008)	0.032*** (0.006)	0.111*** (0.011)
Whether HH purchased a cookstove in last 30/365 days?	0.005 (0.007)	-0.054*** (0.007)	-0.029*** (0.006)	-0.146*** (0.008)	0.004 (0.027)	-0.046 (0.036)	0.013 (0.022)	0.019 (0.030)
HH size	0.001*** (0.0002)	0.028*** (0.001)	0.006*** (0.0005)	0.041*** (0.001)	0.009*** (0.0005)	0.031*** (0.001)	0.015*** (0.001)	0.028*** (0.001)
Age of head of HH	0.0001*** (0.00004)	0.003*** (0.0002)	0.0003*** (0.0001)	0.002*** (0.0002)	0.0004*** (0.0001)	0.001*** (0.0002)	0.0007*** (0.0002)	0.001*** (0.0002)
Whether head of HH is female?	0.0004 (0.002)	0.012* (0.006)	0.015*** (0.004)	0.045*** (0.006)	0.007* (0.004)	0.032*** (0.006)	0.010* (0.005)	0.033*** (0.006)
Whether head of HH is educated?	0.013*** (0.001)	0.075*** (0.004)	0.063*** (0.003)	0.151*** (0.005)	0.067*** (0.003)	0.130*** (0.006)	0.096*** (0.004)	0.127*** (0.006)
Price of LPG	-0.0002 (0.002)	0.000 (0.0005)	-0.002*** (0.001)	-0.015*** (0.002)	-0.0002** (0.0001)	-0.002** (0.001)	0.000 (0.002)	-0.008*** (0.001)
Price of kerosene	0.002* (0.001)	0.014*** (0.003)	0.00004 (0.0001)	-0.0001 (0.0005)	-0.00001 (0.0001)	-0.00001** (0.00004)	0.001 (0.002)	0.003 (0.004)
Whether HH had access to firewood?	-0.048*** (0.005)	-0.178*** (0.006)	-0.180*** (0.008)	-0.279*** (0.008)	-0.383*** (0.011)	-0.437*** (0.008)	-0.472*** (0.015)	-0.471*** (0.011)
Observations	65287	39558	61097	41897	62937	35026	39915	27459
First-Stage Cragg Donald F-Stat	189.214	555.572	343.533	573.048	316.345	475.349	124.89	166.563
Hansen J-Statistic	1.212	6.809	8.968	2.031	11.827	6.535	4.871	3.044
P-value	0.7501	0.0782	0.0297	0.5659	0.008	0.0883	0.1815	0.3849

Notes: The proportion of population in the same village or urban block in the four highest income deciles are used as instruments. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in Table B9 in the Appendix). All specifications include dummies for MPCs, and for belonging to districts, religion, and castes. \* \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported. The variable "Whether HH purchased a cookstove in last 30/365 days" includes by design expenditure on repairs in the 55<sup>th</sup> round.

### 4.3 Discussion and Policy Implications

In this study, we find that there are positive social spillovers in the decision to use LPG between households residing in the same village or urban block. Our results rely on several specifications and use both cross-sectional and panel data. We control for several factors that have already been proved to be critical determinants of LPG use in the literature, and we provide evidence suggesting that social interaction, in particular through membership in specific social networks, may play an important role in facilitating the adoption of clean cooking fuels.

In this section, we discuss the policy implications of our findings, especially from a distributional perspective. The disparity in adoption between rural and urban areas is of particular policy relevance. Social spillovers are very important in this respect. If it were the case that social spillovers were stronger among urban households rather than rural, and this difference would persist over time, policy-makers will find it increasingly difficult to narrow the gap between rural and urban areas, without implementing specific policies aiming at increasing adoption among rural households (such as leveraging social spillovers there). On the other hand, if rural households experience stronger spillovers (and continue to do so over time), it is not implausible to expect them to catch up to urban households in terms of adoption levels. The question that would then arise is how long catching up may take, i.e. how long will rural households still rely on fuels leading to indoor pollution and other health issues while urban households enjoy better living conditions. While performing this exercise, we also discuss about the observed dynamics of spillovers over time.

To this end, we test for heterogeneity in the magnitude of the spillovers between rural and urban households. To this end, we re-estimate our baseline model using the NSS data for the rural and urban sub-samples: the results are presented in columns (1) to (8) of Table 8.<sup>12</sup> We find that in the 43<sup>rd</sup> round of the dataset, the magnitude of the coefficient

---

<sup>12</sup> The corresponding first-stage results are provided in Table B9.

was higher for urban households, than it was for rural households (as can be discerned from columns 1 and 2). In moving from the results of column 1 towards column 8 of this Table, we find that the magnitudes of spillovers decline for both rural and urban households. However, importantly, it declines faster for the urban sub-sample, with the difference in point estimates being statistically significant at the 5 % level. Note that in the 66<sup>th</sup> round, we find that the spillovers are insignificant for the urban sample (column 8).<sup>13</sup>

Hence, our historical analysis provides evidence suggesting that the strength of social spillovers has decreased over time, but at a slower pace for urban areas than for rural areas. That is, while social spillovers in urban areas were already negligible at the time of the last wave of the NSS data, they were still relatively strong in rural areas. This analysis, combined with our findings from the ACCESS data, which we recall focuses exclusively on rural areas, suggests that social spillovers in rural areas have still been at play over the last decade or so and have contributed to narrowing the gap between rural and urban households in terms of LPG adoption levels, with respect to a counterfactual without social spillovers. However, referring back to Figures A2 and A3, we note that the gap in adoption between rural and urban households has not decreased over time. If anything, it appears to have increased. The more recent ACCESS data confirm that, in 2015, only 22 % of rural households used LPG in the six states covered by its survey. Hence, we conclude that the effect of social spillovers has been dominated by other forces, more conducive to adoption in urban areas. Whether in future years social spillovers may be sufficiently strong to correct the current trend may largely depend on whether their potential is fully exploited by policymakers and practitioners.

While access, affordability and availability remain critical to ensure sustained adoption, our findings suggest that there may be other means to expedite adoption. Addition-

---

<sup>13</sup> These estimates are not directly comparable to those in the baseline estimation results of Table 5, which include a dummy variable for rural households as a control. On dropping this dummy variable from the estimations in Table 5, the results of both tables are directly comparable, and the size of the average effects in Table 5 lie between the estimates for the rural and urban sub-samples obtained from Table 8.

ally, the role of social networks, as documented in this paper, and in particular of women's groups, development groups, and agricultural cooperatives, in facilitating LPG adoption by households, clearly indicates the way for policymakers and practitioners to start leveraging social spillovers. Leveraging social networks has been shown, theoretically, to be a much more cost-effective way to lead to normative behavior when designing interventions compared to policies that neglect the network structure of society (Spencer et al., 2019). Such interventions could be tested with randomized controlled trials, in presence and absence of complementary measures such as subsidies.

## 5 Conclusion

Greater adoption of clean cooking fuels such as LPG by the Indian population is vital for achieving a sustained reduction in indoor air pollution, thus ensuring the consequent improvement of respiratory health. This paper analyses whether there are social spillovers in the adoption of LPG in India, and whether these effects are stronger among households in social networks. We use three sources of data spanning a widely heterogeneous population and several decades, which enable us to provide a broad scope in addressing this research question. We provide multiple pieces of evidence suggesting that social spillovers are present in the adoption of LPG in India, and that social networks play an important role in the dissemination of information across households. We use both cross-sectional and panel data, and control for several household-level characteristics of LPG adoption that have been shown to be important determinants in the literature, while adopting multiple identification strategies to tackle potential endogeneity. Our results point to an important, and most likely largely untapped potential of social networks to spread information about LPG among Indian households. Our historical perspective also informs policymakers and practitioners elsewhere in the developing world, looking to encourage consumers to switch to cleaner sources of energy in contexts similar to urban India a few decades ago, or to rural India in more recent times. Following our results, we consider that in-

terventions leveraging social spillovers should be part of the battery of instruments used by governments with the goal of helping the population switching to cleaner fuels and a cleaner indoor environment.

## References

- Aklin, Michaël, Chao-yo Cheng, Karthik Ganesan, Abhishek Jain, and Johannes Urpelainen (2016) *Access to Clean Cooking Energy and Electricity: Survey of States in India (ACCESS)*: Harvard Dataverse, V1.
- Arun, Shoba, Samuel Ananim, and Thankom Arun (2016) “Do All Networks ‘Work’? The Mediating Role of Social Networks on Consumption Expenditure in India,” *Sociology*, Vol. 50 (3), pp. 522–541.
- Attwood, Donald W. and Baburao S. Baviskar (1988) *Who Shares? Co-operatives and Rural Development*, Delhi: Oxford University Press.
- Bandiera, Oriana and Imran Rasul (2006) “Social Networks and Technology Adoption in Northern Mozambique,” *The Economic Journal*, Vol. 116 (514), pp. 869–902.
- Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson (2014) *Gossip: Identifying Central Individuals in a Social Network*, National Bureau of Economic Research: Working Paper No. 20422.
- Baranzini, Andrea, Stefano Carattini, and Martin Péclat (2017) *What Drives Social Contagion in the Adoption of Solar Photovoltaic Technology?*, Grantham Research Institute on Climate Change and the Environment: Working Paper No. 270.
- Barnhardt, Sharon, Erica Field, and Rohini Pande (2017) “Moving to Opportunity or Isolation? Network Effects of a Randomized Housing Lottery in Urban India,” *American Economic Journal: Applied Economics*, Vol. 9 (1), pp. 1–32.
- Beltramo, Theresa, Garrick Blalock, David I Levine, and Andrew M Simons (2015) “Does Peer Use Influence Adoption of Efficient Cookstoves? Evidence from a Randomized Controlled Trial in Uganda,” *Journal of Health Communication*, Vol. 20 (1), pp. 55–66.
- Bollinger, Bryan and Kenneth Gillingham (2012) “Peer Effects in the Diffusion of Solar Photovoltaic Panels,” *Marketing Science*, Vol. 31 (6), pp. 900–912.
- Carattini, Stefano, Simon Levin, Alessandro Tavoni, and Others (2017) *Cooperation in the Climate Commons*, Grantham Research Institute on Climate Change and the Environment: Working Paper No. 259.
- Case, Anne C and Lawrence F Katz (1991) *The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths*, National Bureau of Economic Research: Working Paper No. 3705.
- Cheng, Chao-yo and Johannes Urpelainen (2014) “Fuel Stacking in India: Changes in the Cooking and Lighting Mix, 1987–2010,” *Energy*, Vol. 76 (1), pp. 306–317.
- Conley, Timothy G and Christopher R Udry (2010) “Learning About a New Technology: Pineapple in Ghana,” *The American Economic Review*, Vol. 100 (1), pp. 35–69.

- Deininger, Klaus and Yanyan Liu (2013) “Evaluating Program Impacts on Mature Self-help Groups in India,” *The World Bank Economic Review*, Vol. 27 (2), pp. 272–296.
- Desai, Sonalde and Reeve Vanneman (2005) *India Human Development Survey-I*, Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research.
- (2011) *India Human Development Survey-II*, Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research.
- Duflo, Esther, Michael Greenstone, and Rema Hanna (2008) “Cooking Stoves, Indoor Air Pollution and Respiratory Health in Rural Orissa,” *Economic and Political Weekly*, Vol. 43 (32), pp. 71–76.
- Duflo, Esther and Emmanuel Saez (2002) “Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues’ Choices,” *Journal of Public Economics*, Vol. 85 (1), pp. 121–148.
- Ezzati, Majid, Bernard M Mbinda, and Daniel M Kammen (2000) “Comparison of Emissions and Residential Exposure from Traditional and Improved Cookstoves in Kenya,” *Environmental Science and Technology*, Vol. 34, pp. 578–583.
- Farsi, Mehdi, Massimo Filippini, and Shonali Pachauri (2007) “Fuel Choices in Urban Indian Households,” *Environment and Development Economics*, Vol. 12 (6), pp. 757–774.
- Feigenberg, Benjamin, Erica Field, and Rohini Pande (2013) “The Economic Returns to Social Interaction: Experimental Evidence from Microfinance,” *The Review of Economic Studies*, Vol. 80 (4), pp. 1459–1483.
- Filippini, Massimo and Suchita Srinivasan (2018) *Impact of Religious Participation, Social Interactions and Globalisation on Meat Consumption: Evidence from India*, CER-ETH Economics: Working Paper Series 18/304.
- Fisman, Raymond and Tarun Khanna (2004) “Facilitating Development: The Role of Business Groups,” *World Development*, Vol. 32 (4), pp. 609–628.
- Graziano, Marcello and Kenneth Gillingham (2015) “Spatial Patterns of Solar Photovoltaic System Adoption: The Influence of Neighbors and the Built Environment,” *Journal of Economic Geography*, Vol. 15 (4), pp. 815–839.
- Greenstone, Michael and B Kelsey Jack (2015) “Envirodevonomics: A Research Agenda for an Emerging Field,” *The Journal of Economic Literature*, Vol. 53 (1), pp. 5–42.
- Gupta, Gautam and Gunnar Kohlin (2006) “Preferences for Domestic Fuel: Analysis with Socio-economic Factors and Rankings in Kolkata, India,” *Ecological Economics*, Vol. 57 (1), pp. 107–121.
- Hanna, Rema, Esther Duflo, and Michael Greenstone (2016) “Up in Smoke: The Influence of Household Behavior on the Long-run Impact of Improved Cooking Stoves,” *American Economic Journal: Economic Policy*, Vol. 8 (1), pp. 80–114.



- International Energy Agency (2017) *Energy Access Outlook 2017*, World Energy Outlook Special Report.
- International Institute for Sustainable Development (2014) *Subsidies to Liquefied Petroleum Gas in India: An Overview of Recent Reforms*, Global Subsidies Initiative.
- Jain, Abhishek, Saurabh Tripathi, Sunil Mani, Sasmita Patnaik, Tauseef Shahidi, and Karthik Ganesan (2018) *Access to Clean Cooking Energy and Electricity Survey of States*, Council on Energy, Environment and Water.
- Khandker, Shahidur R., Hussain A. Samad, Rubaba Ali, and Douglas F. Barnes (2014) “Who Benefits Most from Rural Electrification? Evidence in India,” *The Energy Journal*, Vol. 35 (2), pp. 75–96.
- Knack, Stephen and Philip Keefer (1997) “Does Social Capital Have an Economic Payoff? A Cross-country Investigation,” *The Quarterly Journal of Economics*, Vol. 112 (4), pp. 1251–1288.
- Kraft-Todd, Gordon T., Bryan Bollinger, Kenneth Gillingham, Stefan Lamp, and David G. Rand (2018) “Credibility-enhancing Displays Promote the Provision of Non-normative Public Goods,” *Nature*, Vol. 563, pp. 245–248.
- Kremer, Michael and Edward Miguel (2007) “The Illusion of Sustainability,” *The Quarterly Journal of Economics*, Vol. 122 (3), pp. 1007–1065.
- Kumar, K S Kavi and Brinda Viswanathan (2007) “Changing Structure of Income Indoor Air Pollution Relationship in India,” *Energy Policy*, Vol. 35 (11), pp. 5496–5504.
- Lewis, Jessica J and Subhrendu K Pattanayak (2012) “Who Adopts Improved Fuels and Cookstoves? A Systematic Review,” *Environmental Health Perspectives*, Vol. 120 (5), pp. 637–645.
- Manski, Charles F (1993) “Identification of Endogenous Social Effects: The Reflection Problem,” *The Review of Economic Studies*, Vol. 60 (3), pp. 531–542.
- (2000) “Economic Analysis of Social Interactions,” *The Journal of Economic Perspectives*, Vol. 14 (3), pp. 115–136.
- Masera, Omar R, Barbara D Saatkamp, and Daniel M Kammen (2000) “From Linear Fuel Switching to Multiple Cooking Strategies: A Critique and Alternative to the Energy Ladder Model,” *World Development*, Vol. 28 (12), pp. 2083–2103.
- Mobarak, Ahmed Mushfiq, Puneet Dwivedi, Robert Bailis, Lynn Hildemann, and Grant Miller (2012) “Low Demand for Nontraditional Cookstove Technologies,” *Proceedings of the National Academy of Sciences*, Vol. 109 (27), pp. 10815–10820.
- Moffitt, Robert A (2001) “Policy Interventions, Low-level Equilibria, and Social Interactions,” in Brookings Institution Press ed. *Social Dynamics*: MIT Press.
- Muller, Christophe and Huijie Yan (2018) “Household Fuel Use in Developing Countries: Review of Theory and Evidence,” *Energy Economics*, Vol. 70, pp. 429–439.

- Munshi, Kaivan (2004) “Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution,” *Journal of Development Economics*, Vol. 73 (1), pp. 185–213.
- National Sample Survey Office (2019) *Household Consumer Expenditure Survey*, Ministry of Statistics and Program Implementation: Government of India.
- Oster, Emily and Rebecca Thornton (2012) “Determinants of Technology Adoption: Peer effects in Menstrual Cup Take-up,” *Journal of the European Economic Association*, Vol. 10 (6), pp. 1263–1293.
- Press Information Bureau, Government of India (2014) “Expansion of 5 kg LPG Cylinder Scheme in Delhi,” URL: <http://pib.nic.in/newsite/PrintRelease.aspx?relid=114108>.
- Putnam, Robert D, Robert Leonardi, and Raffaella Y Nanetti (1994) *Making Democracy Work: Civic Traditions in Modern Italy*: Princeton University Press.
- Rao, M Narasimha and B Sudhakara Reddy (2007) “Variations in Energy Use by Indian Households: An Analysis of Micro Level Data,” *Energy*, Vol. 32 (2), pp. 143–153.
- Reddy, B Sudhakara (1995) “A Multilogit Model for Fuel Shifts in the Domestic Sector,” *Energy*, Vol. 20 (9), pp. 929–936.
- Spencer, Gwen, Stefano Carattini, and Richard B. Howarth (2019) “Short-term Interventions for Long-term Change: Spreading Stable Green Norms in Networks,” *Review of Behavioral Economics*, Vol. 6 (1), pp. 53–93.
- Stock, James H. and Motohiro Yogo (2005) “Testing for Weak Instruments in Linear IV Regression,” in Donald W.K. Andrews and James H. Stock eds. *Identification and Inference for Econometric Models*: Cambridge University Press, pp. 80–108.
- Thorp, Rosemary, Frances Stewart, and Amrik Heyer (2005) “When and How Far is Group Formation a Route Out of Chronic Poverty?,” *World Development*, Vol. 33 (6), pp. 907–920.
- Van Leeuwen, Richenda, Alex Evans, and Besnik Hyseni (2017) *Increasing the Use of Liquefied Petroleum Gas in Cooking in Developing Countries*, World Bank, Washington, DC: 2017/74.
- Vanneman, Reeve, James Noon, Mitali Sen, Sonalde Desai, and Abusaleh Shariff (2006) *Social Networks in India: Caste, Tribe, and Religious Variations*, Paper presented at the Annual Meeting of the Population Association of America, Los Angeles, California.
- Vikram, Kriti, Reeve Vanneman, and Sonalde Desai (2012) “Linkages Between Maternal Education and Childhood Immunization in India,” *Social Science & Medicine*, Vol. 75 (2), pp. 331–339.
- World Bank Data (2019) “World Development Indicators,” URL: <https://data.worldbank.org/indicator/NY.GDP.PCAP.CN?locations=IN>.

World Health Organization (2018) “Household Air Pollution and Health,”  
URL: <https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health>.

# Appendix A Figures

Figure A1: Distribution of Households by Primary Cooking Fuel Type (Source: NSS)

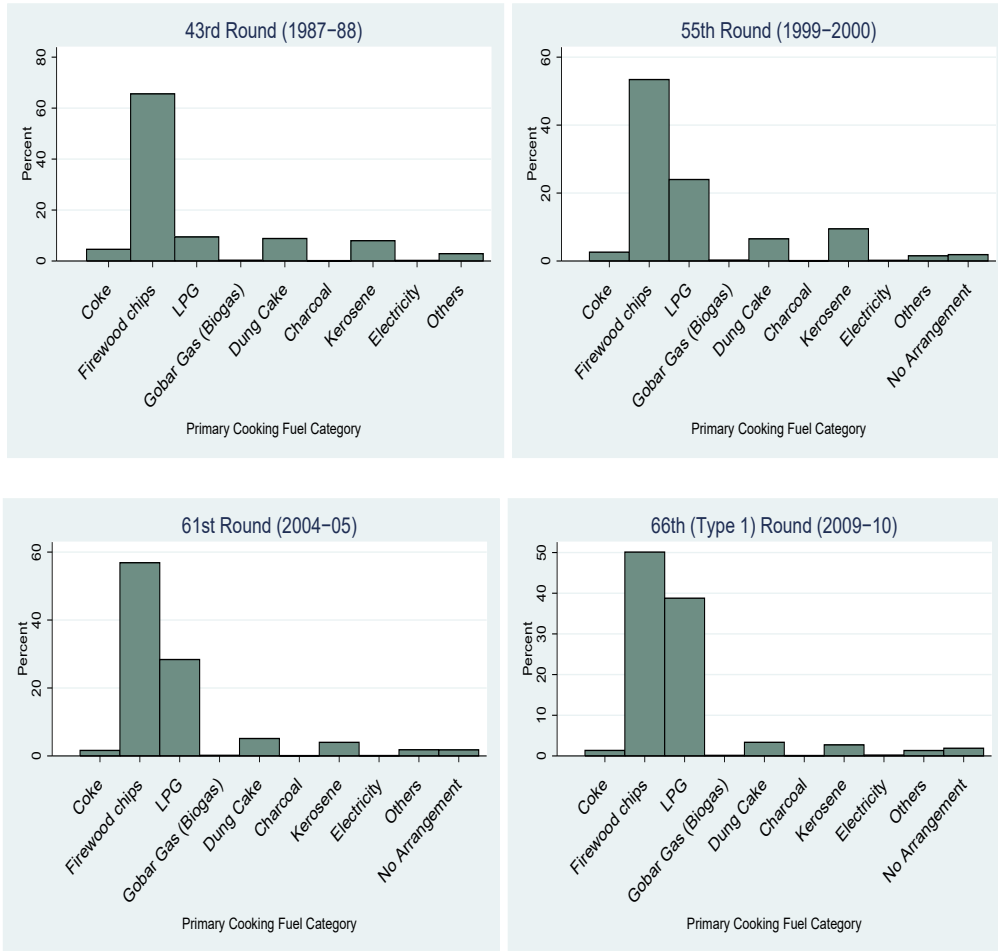


Figure A2: Population Share (%) Using LPG as the Primary Cooking Fuel: 1983 to 2011-12 (Source: NSS)

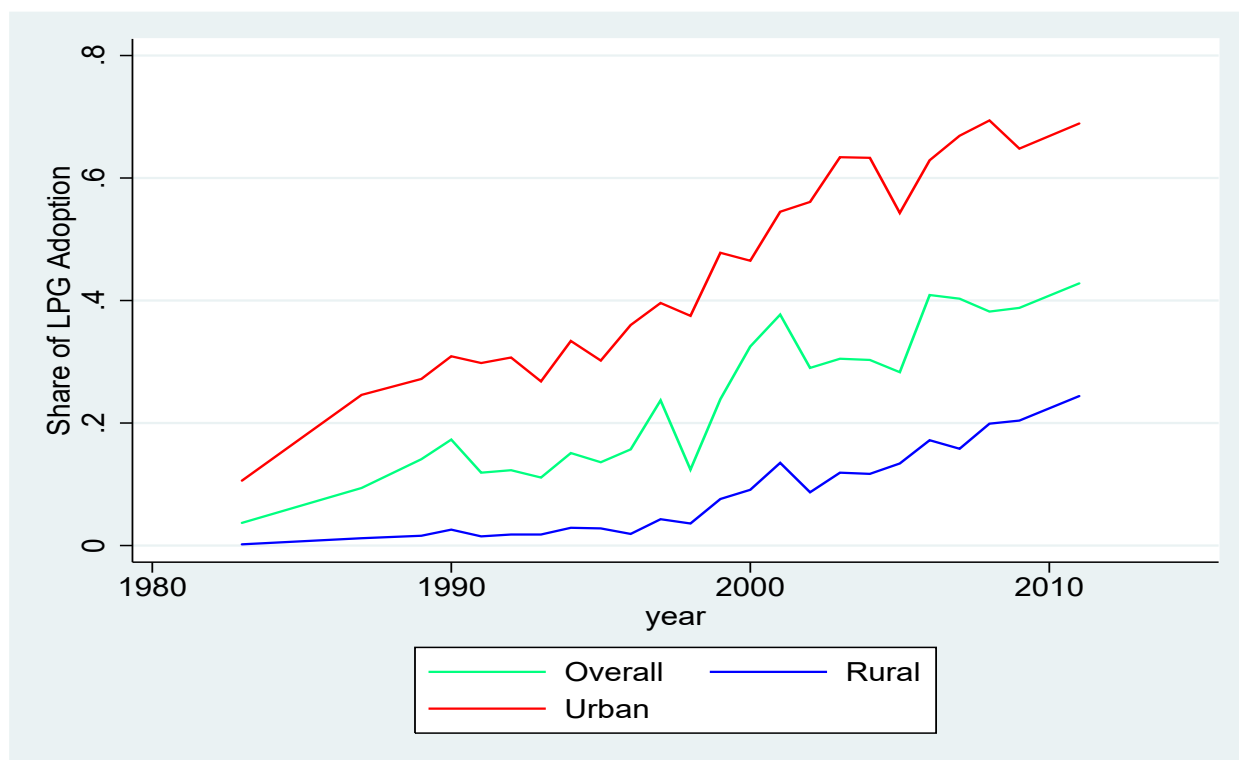


Figure A3: Population Share (%) Using LPG as the Primary Cooking Fuel: Thick Rounds of the NSS (Source: NSS)

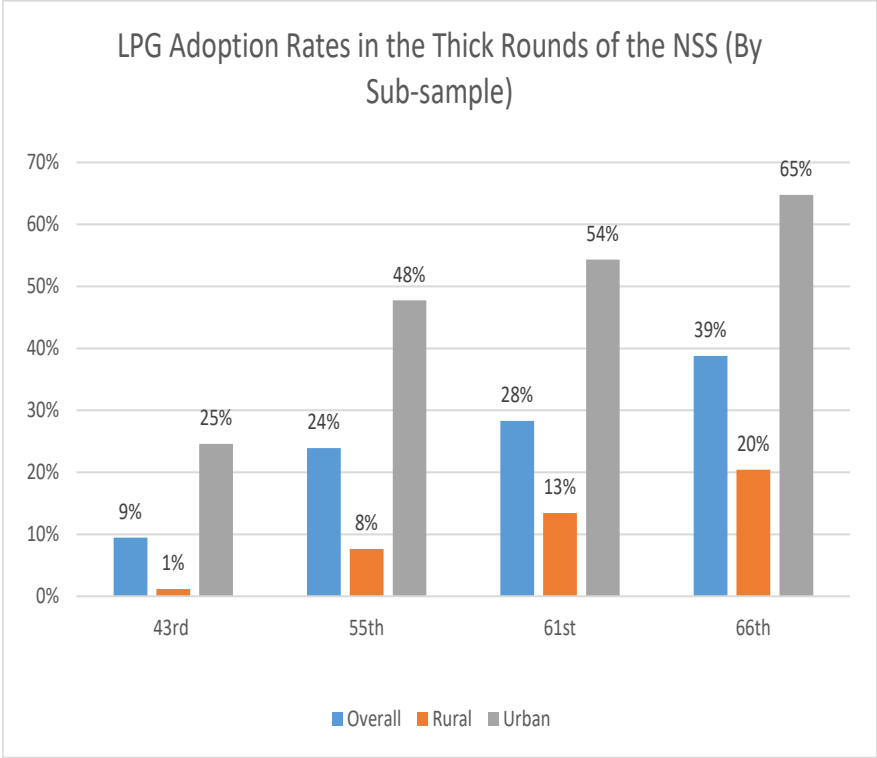
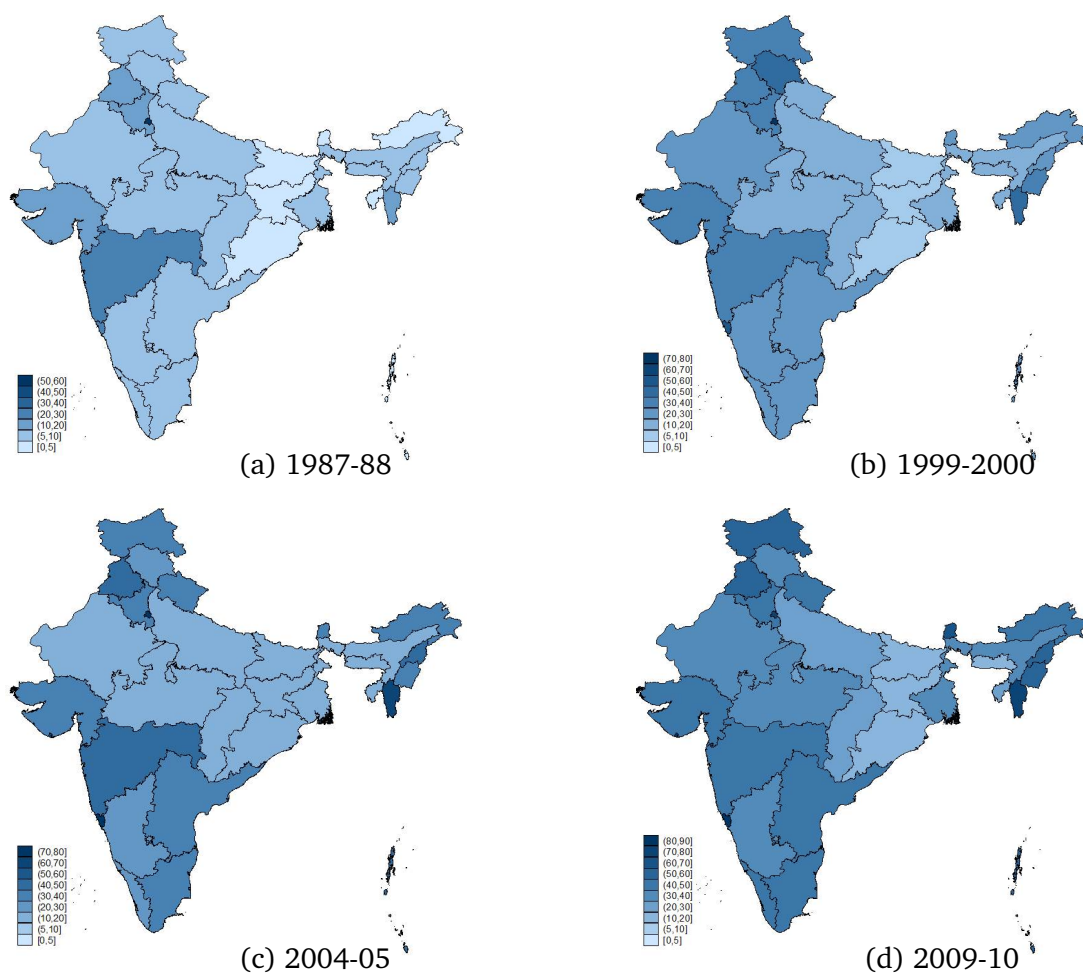
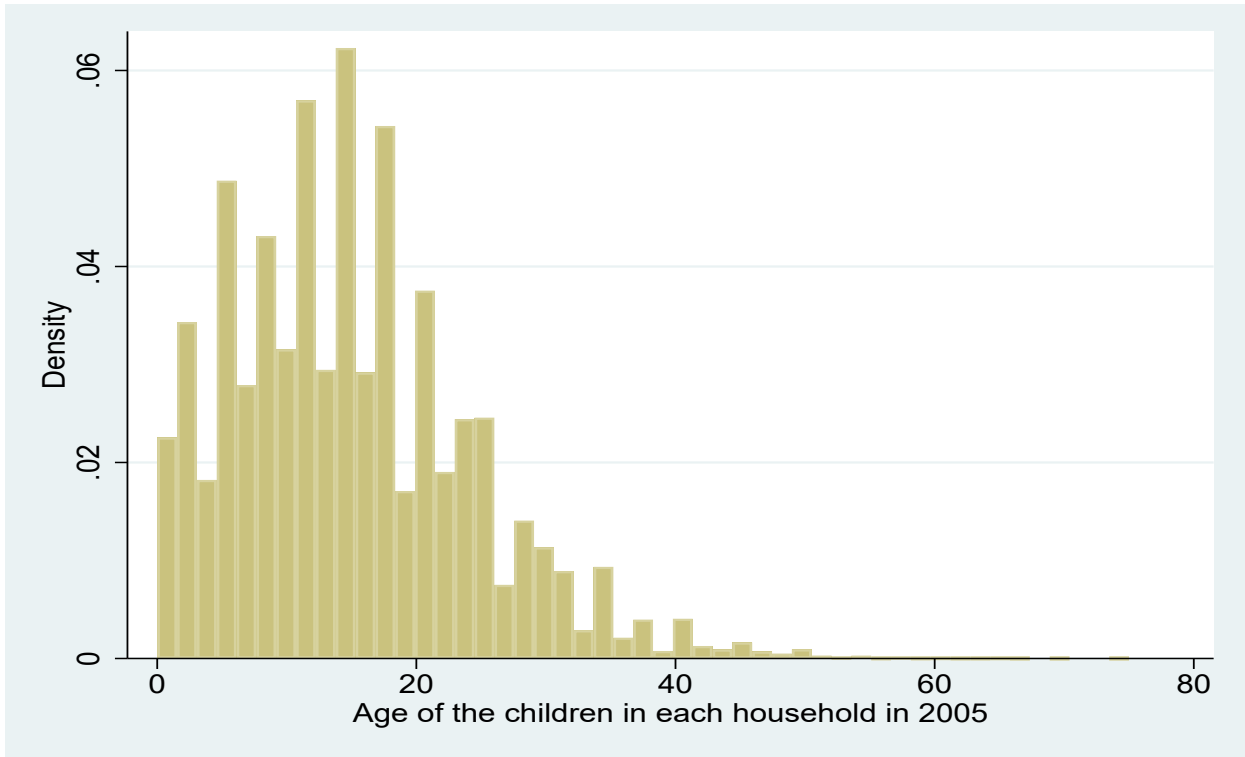


Figure A4: Share of Households Using LPG as the Primary Cooking Fuel (By State)  
 (Source: NSS)



Notes: The maps show the proportion of households (by state) for whom LPG was the primary cooking fuel in the 43<sup>rd</sup>, 55<sup>th</sup>, 61<sup>st</sup>, and the 66<sup>th</sup> rounds of the NSS.

Figure A5: Age distribution of children in households in 2005 in the IHDS data





## Appendix B Tables

Table B1: NSS Estimation Results: Linear Probability Model

Round	43	55	61	66
Year	1987-88	1999-00	2004-05	2009-10
Dep.Var.: Whether prim. cooking fuel of HH <i>i</i> is LPG	(1)	(2)	(3)	(4)
Average LPG use rate (village/urban block)	0.619*** (0.007)	0.468*** (0.007)	0.353*** (0.007)	0.341*** (0.007)
Whether bordering an urban center?	-0.053*** (0.013)	-0.051 (0.065)	-0.131*** (0.060)	-0.025 (0.045)
Whether HH had access to electricity?	0.037*** (0.002)	0.040*** (0.003)	0.035*** (0.003)	0.037*** (0.004)
Whether HH lives in a rural area?	0.011*** (0.002)	-0.007 (0.003)	-0.011*** (0.003)	-0.001 (0.004)
Whether HH purchased a cookstove in last 30/365 days?	-0.033*** (0.006)	-0.107*** (0.005)	-0.021 (0.022)	0.006 (0.017)
HH size	0.009*** (0.0003)	0.017*** (0.0005)	0.015*** (0.0005)	0.019*** (0.0006)
Age of head of HH	0.001*** (0.00007)	0.001*** (0.00009)	0.0008*** (0.0001)	0.0008*** (0.0001)
Whether head of HH is female?	0.006*** (0.002)	0.021*** (0.003)	0.013*** (0.003)	0.011*** (0.004)
Whether head of HH is educated?	0.041*** (0.002)	0.089*** (0.002)	0.083*** (0.002)	0.100*** (0.003)
Price of LPG	-0.00003 (0.0005)	-0.010*** (0.002)	-0.0004** (0.0002)	-0.004*** (0.001)
Price of kerosene	0.004*** (0.001)	0.00004 (0.00009)	-0.00001** (0.000006)	0.003 (0.002)
Whether HH had access to firewood?	-0.134*** (0.003)	-0.229*** (0.004)	-0.396*** (0.005)	-0.435*** (0.006)
Observations	104845	102994	97963	67374
R-squared	0.4729	0.5775	0.6004	0.6381

*Notes:* All specifications include dummy variables for districts, MPCE deciles, religion and cast. Standard errors are clustered at the village/urban block level (reported in parentheses). \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B2: NSS Estimation Results: First Stage

Round	43	55	61	66
Year	1987-88	1999-00	2004-05	2009-10
Dep.Var.: Whether prim. cooking fuel of HH $i$ is LPG	(1)	(2)	(3)	(4)
Proportion of households in 7 <sup>th</sup> income decile	-0.015 (0.034)	0.04 (0.035)	-0.014 (0.028)	0.187*** (0.039)
Proportion of households in 8 <sup>th</sup> income decile	0.056** (0.029)	0.277*** (0.041)	0.264*** (0.035)	0.280*** (0.044)
Proportion of households in 9 <sup>th</sup> income decile	0.186*** (0.032)	0.514*** (0.040)	0.465*** (0.032)	0.388*** (0.031)
Proportion of households in 10 <sup>th</sup> income decile	0.507*** (0.028)	0.549*** (0.019)	0.435*** (0.020)	0.312*** (0.019)
Whether bordering an urban center?	-0.112*** (0.022)	0.021*** (0.004)	-0.019 (0.066)	0.056 (0.058)
Whether HH had access to electricity?	0.030*** (0.002)	0.046*** (0.003)	0.038*** (0.003)	0.052*** (0.004)
Whether HH lives in a rural area?	-0.129*** (0.003)	-0.260*** (0.005)	-0.284*** (0.005)	-0.298*** (0.005)
Whether HH purchased a cookstove in last 30/365 days?	-0.001 (0.008)	-0.026*** (0.006)	0.006 (0.011)	0.039*** (0.017)
HH size	0.001*** (0.0003)	0.002*** (0.0003)	-0.001** (0.0003)	-0.0005 (0.0004)
Age of head of HH	0.0009*** (0.00006)	0.0004*** (0.00007)	0.0003*** (0.00007)	0.0004*** (0.0001)
Whether head of HH is female?	0.009*** (0.002)	0.018*** (0.003)	0.017*** (0.002)	0.025*** (0.003)
Whether head of HH is educated?	0.027*** (0.001)	0.036*** (0.002)	0.023*** (0.002)	0.029*** (0.003)
Price of LPG	-0.0003 (0.0004)	-0.005*** (0.001)	0.00002 (0.00007)	-0.002** (0.0009)
Price of kerosene	-0.006*** (0.001)	0.0001 (0.0001)	0.000005 (0.000003)	-0.002 (0.002)
Whether HH had access to firewood?	-0.083*** (0.003)	-0.107*** (0.004)	-0.131*** (0.004)	-0.173*** (0.005)
Observations	104845	102994	97963	67374

Notes: All specifications include dummy variables for districts, MPCE deciles, religion and cast. Standard errors are clustered at the village/urban block level (reported in parentheses). \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B3: ACCESS Estimation Results: First Stage

Dep.Var.: Whether HH <i>i</i> uses LPG for cooking Column	Income deciles (1)	Education (women) and land (2)	Income decile (3)	Education (women) and land (4)
Proportion of households in 7 <sup>th</sup> income decile	0.085 (0.066)		0.042 (0.091)	
Proportion of households in 8 <sup>th</sup> income decile	-0.049 (0.095)		0.04 (0.108)	
Proportion of households in 9 <sup>th</sup> income decile	0.377*** (0.075)		0.365*** (0.082)	
Proportion of households in 10 <sup>th</sup> income decile	0.563*** (0.093)		0.546*** (0.117)	
Proportion of households with educated women		0.251 (0.268)		-0.127 (0.281)
Average size of land holdings		0.003*** (0.001)		0.005*** (0.002)
Whether HH had access to electricity?	0.010*** (0.004)	0.011*** (0.004)	0.020*** (0.007)	0.023*** (0.007)
HH size	-0.0001 (0.0003)	-0.0003 (0.0003)	0.001** (0.0005)	0.0006 (0.0006)
Number of years of education of HH head	0.002** (0.001)	0.003*** (0.001)	0.0004 (0.002)	-0.0001 (0.002)
Whether HH owns land?	-0.023*** (0.004)	-0.029*** (0.005)	-0.027*** (0.006)	-0.033*** (0.006)
Whether LPG is unavailable?	-0.013*** (0.005)	-0.017*** (0.006)		
Observations	8563	8563	3603	3603

*Notes:* All specifications include income decile, sub-district, religion and caste dummies. Standard errors are clustered at the village/urban block level (reported in parentheses). \*,\*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B4: IHDS Estimation Results: First Stage

Model Column	Overall (1)	No network (2)	Women's groups (3)	Youth and sports groups (4)	Unions or business groups (5)	Credit and savings groups (6)	Religious or social groups (7)	Caste associations (8)	Developmental groups/NGOs (9)	Agricultural cooperatives (10)
Proportion of households with educated women	0.170*** (0.005)	0.161*** (0.007)	0.172*** (0.014)	0.100*** (0.020)	0.087*** (0.013)	0.205*** (0.014)	0.151*** (0.011)	0.179*** (0.015)	0.118*** (0.022)	0.198*** (0.036)
Average size of land holdings	0.0006*** (0.0002)	0.0006** (0.0003)	0.0006*** (0.0002)	0.001*** (0.0004)	0.0002*** (0.00008)	0.001*** (0.0004)	0.0004*** (0.0001)	0.001*** (0.0003)	0.001 (0.0007)	0.0005*** (0.0002)
Whether HH lives in a rural area?	0.077** (0.037)	0.064 (0.056)	-0.011 (0.095)	0.06 (0.071)	0.038 (0.069)	0.076 (0.092)	0.119** (0.057)	0.182*** (0.053)	0.231*** (0.047)	0.012 (0.093)
Whether HH had access to electricity?	0.518*** (0.020)	0.541*** (0.024)	0.380*** (0.076)	0.526*** (0.138)	0.473*** (0.140)	0.335*** (0.066)	0.540*** (0.046)	0.379*** (0.058)	0.249*** (0.108)	0.007 (0.256)
HH size	-0.082*** (0.002)	-0.031*** (0.002)	-0.028*** (0.004)	-0.021*** (0.005)	-0.016*** (0.004)	-0.042*** (0.004)	-0.024*** (0.003)	-0.034*** (0.005)	-0.025*** (0.007)	-0.050*** (0.012)
Number of years of education of HH head	-0.002*** (0.001)	-0.004*** (0.001)	-0.003 (0.002)	0.001 (0.003)	0.002 (0.002)	0.003 (0.002)	-0.002 (0.002)	0.0008 (0.002)	-0.002 (0.004)	-0.003 (0.005)
Whether HH owns land?	-0.019** (0.010)	-0.011 (0.014)	-0.057*** (0.021)	-0.018 (0.028)	-0.020 (0.027)	-0.018 (0.023)	-0.024 (0.018)	-0.056*** (0.024)	-0.002 (0.032)	-0.059 (0.040)
Observations	41090	17576	9582	3964	5622	8550	12166	7868	3272	1476

Notes: All specifications include dummy variables for income deciles, religion and caste, as well as household and year fixed effects. Standard errors are clustered at the village/urban block level (reported in parentheses). \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B5: Robustness Checks: Sample Size Thresholds (Second Stage IHDS)

Column	(1)	(2)	(3)
Average LPG use rate (village/urban block)	0.946*** (0.009)	0.942*** (0.010)	0.943*** (0.011)
Whether HH lives in a rural area?	0.007 (0.007)	0.006 (0.008)	0.001 (0.012)
Whether HH had access to electricity?	0.110*** (0.014)	0.104*** (0.014)	0.088*** (0.015)
HH size	-0.001 (0.001)	-0.002 (0.001)	-0.002 (0.002)
Number of years of education of HH head	0.002*** (0.0006)	0.002*** (0.0007)	0.002* (0.001)
Whether HH owns land?	0.006 (0.006)	0.005 (0.007)	-0.002 (0.009)
Observations	38240	28744	15534
First-stage Cragg Donald F-Stat	3148.208	2806.732	2188.269
Hansen J-Statistic	0.007	0.015	0.409
P-Value	0.9317	0.9022	0.5225

*Notes:* Average land ownership and the proportion of educated women in the same village or urban block are used as instruments. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in columns (1) to (3) of Table B8). All specifications include household-level and year fixed effects. All specifications include dummies for income deciles, and controls for caste and religion. \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B6: Robustness Checks: Different Model Specifications (Second Stage IHDS)

Column	(1)	(2)
Average LPG use rate (village/urban block)	0.895*** (0.007)	0.902*** (0.008)
Whether HH had access to electricity?	0.141*** (0.014)	0.136*** (0.010)
HH size	-0.005*** (0.0007)	-0.004*** (0.0004)
Number of years of education of HH head	0.008*** (0.0004)	0.007*** (0.0003)
Whether HH owns land?	0.022*** (0.004)	0.020*** (0.003)
Observations	43179	43179
First-stage Cragg Donald F-Stat	6535.421	5612.617
Hansen J-Statistic	8.148	0.206
P-Value	0.0043	0.6502

*Notes:* Average land ownership and the proportion of educated women in the same village or urban block are used as instruments. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in columns (4) and (5) of Table B8). Both specifications include household-level and year fixed effects, as well as dummies for income deciles, and controls for caste and religion. \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B7: Robustness Checks: Accounting for Measurement Error (Second Stage IHDS)

Column	(1)
Average LPG use rate (village/urban block)	0.945*** (0.010)
Whether HH lives in a rural area?	0.010 (0.013)
Whether HH had access to electricity?	0.111*** (0.010)
HH size	-0.001 (0.001)
Number of years of education of HH head	0.002*** (0.001)
Whether HH owns land?	0.006 (0.005)
Observations	40008
First-stage Cragg Donald F-Stat	3137.579
Hansen J-Statistic	0
P-Value	0.9825

*Notes:* Average land ownership and the proportion of educated women in the same village or urban block are used as instruments. The Cragg-Donald F-statistics are consistently high, and surpass the rule-of-thumb bound of 10 proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in column (6) of Table B8). The specification includes household-level and year fixed effects, and dummies for income deciles, and controls for caste and religion. \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.

Table B8: Robustness Checks: First Stage

Column	(1)	(2)	(3)	(4)	(5)	(6)
Proportion of households with educated women	0.173*** (0.006)	0.191*** (0.007)	0.242*** (0.010)	0.147*** (0.004)	0.140*** (0.001)	0.141*** (0.001)
Average size of land holdings	0.0006*** (0.0002)	0.0006*** (0.0002)	0.0006*** (0.0001)	0.0007*** (0.0002)	0.0006*** (0.00003)	0.0006*** (0.00003)
Whether HH lives in a rural area?	0.082** (0.038)	0.103** (0.047)	0.186*** (0.060)	0.033 (0.038)	0.013*** (0.004)	0.013*** (0.004)
Whether HH had access to electricity?	0.513*** (0.020)	0.501*** (0.021)	0.451*** (0.024)	0.417*** (0.020)	0.422*** (0.010)	0.423*** (0.010)
HH size	-0.033*** (0.002)	-0.034*** (0.002)	-0.034*** (0.003)	-0.024*** (0.001)	-0.024*** (0.0005)	-0.024*** (0.0005)
Number of years of education of household head	-0.002 (0.0009)	-0.003*** (0.001)	-0.004*** (0.001)	-0.006*** (0.0003)	-0.004*** (0.0003)	-0.004*** (0.0003)
Whether HH owns land?	-0.018* (0.010)	-0.020*** (0.011)	-0.032*** (0.014)	-0.015*** (0.004)	-0.014*** (0.003)	-0.014*** (0.003)
Observations	38240	28744	15534	43179	43179	40008

Notes: Average land ownership and the proportion of educated women in the same village or urban block are used as instruments. All specifications include dummy variables for income deciles, religion and caste as well as household and year fixed effects. Standard errors are clustered at the village/urban block level (reported in parentheses). \*, \*\* and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.



Table B9: NSS Estimation Results on Rural and Urban Households: First Stage

Year Round Dep.Var.: Whether prim. cooking fuel of HH <i>i</i> is LPG (NSS) or whether HH <i>i</i> spent on LPG in the last 30 days (IHDS) Column	1987-88 43		1999-00 55		2004-05 61		2009-10 66	
	Rural (1)	Urban (2)	Rural (3)	Urban (4)	Rural (5)	Urban (6)	Rural (7)	Urban (8)
Proportion of households in 7 <sup>th</sup> income decile	0.01 (0.011)	-0.046 (0.056)	0.069** (0.036)	-0.010 (0.061)	-0.012 (0.031)	-0.022 (0.055)	0.128*** (0.050)	0.220*** (0.051)
Proportion of households in 8 <sup>th</sup> income decile	0.015 (0.012)	0.045 (0.045)	0.121*** (0.046)	0.321*** (0.055)	0.285*** (0.041)	0.198*** (0.055)	0.227*** (0.070)	0.251*** (0.041)
Proportion of households in 9 <sup>th</sup> income decile	0.049*** (0.016)	0.168*** (0.041)	0.456*** (0.068)	0.447*** (0.050)	0.333*** (0.047)	0.472*** (0.039)	0.423*** (0.066)	0.363*** (0.033)
Proportion of households in 10 <sup>th</sup> income decile	0.177*** (0.053)	0.484*** (0.031)	0.500*** (0.093)	0.533*** (0.021)	0.443*** (0.067)	0.457*** (0.021)	0.516*** (0.063)	0.289*** (0.020)
Whether bordering an urban center?	-0.012 (0.008)	-0.139 (0.097)	0.012*** (0.004)	0.037*** (0.010)	0.003 (0.072)	0.105 (0.075)	0.217*** (0.078)	-0.080 (0.078)
Whether HH had access to electricity?	0.007*** (0.001)	0.052*** (0.005)	0.047*** (0.002)	0.051*** (0.007)	0.031*** (0.002)	0.043*** (0.006)	0.042*** (0.004)	0.064*** (0.010)
Whether HH purchased a cookstove in last 30/365 days?	0.010** (0.005)	-0.015 (0.011)	0.015*** (0.005)	-0.042*** (0.009)	0.011 (0.012)	0.004 (0.017)	0.052*** (0.022)	0.017 (0.024)
HH size	-0.0003*** (0.00008)	0.005*** (0.0007)	-0.002*** (0.0003)	0.009*** (0.001)	-0.001 (0.0003)	0.002*** (0.0007)	-0.002*** (0.0005)	0.003*** (0.0008)
Age of head of HH	0.00 (0.00002)	0.002*** (0.0001)	0.00003 (0.00006)	0.0009*** (0.0001)	0.00002 (0.00006)	0.0008*** (0.0001)	-0.00003 (0.0002)	0.0007*** (0.0002)
Whether head of HH is female?	0.003*** (0.0008)	0.014*** (0.005)	0.018*** (0.003)	0.023*** (0.005)	0.011*** (0.002)	0.017*** (0.004)	0.017*** (0.003)	0.033*** (0.004)
Whether head of HH is educated?	0.004*** (0.0001)	0.052*** (0.003)	0.019*** (0.002)	0.057*** (0.005)	0.009*** (0.002)	0.043*** (0.004)	0.013*** (0.003)	0.045*** (0.005)
Price of LPG	-0.0004 (0.0004)	-0.0003 (0.0004)	-0.003*** (0.0007)	-0.009*** (0.002)	0.0005 (0.00007)	-0.00003 (0.00003)	0.00007 (0.0001)	-0.002* (0.001)
Price of kerosene	-0.0006 (0.0006)	-0.002*** (0.003)	0.000006 (0.00006)	0.001*** (0.0006)	-0.00001 (0.00003)	0.00001*** (0.00003)	-0.004** (0.002)	0.001 (0.003)
Whether HH had access to firewood?	-0.013*** (0.002)	-0.097*** (0.005)	-0.055*** (0.004)	-0.137*** (0.007)	-0.096*** (0.005)	-0.126*** (0.006)	-0.136*** (0.006)	-0.155*** (0.007)
Observations	65287	39558	61097	41897	62937	35026	39915	27459

Notes: All specifications include dummies for district, MPCE, religion and caste. \*, \*\*, and \*\*\* respectively denote significance at 10 %, 5 % and 1 % levels. The coefficient of the constant has not been reported.