Research Paper 48 **| 2016**

ADDING FUEL TO FIRE? SOCIAL SPILLOVERS AND SPATIAL DISPARITIES IN THE ADOPTION OF LPG IN INDIA

Suchita SRINIVASAN, Stefano CARATTINI

Adding fuel to fire? Social spillovers and spatial disparities in the adoption of LPG in India

Suchita Srinivasan* Stefano Carattini†

November 24, 2016

Abstract

The Indian population is still heavily reliant on solid biomass as a cooking fuel, especially in the rural areas, despite its negative health 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 the country. This paper investigates whether social spillovers might play a role in a household's decision to use LPG, how these effects vary across different subpopulations, and whether they exacerbate or ameliorate existing spatial disparities in LPG use. Using data from the National Sample Survey (NSS) and the India Human Development Survey (IHDS), this paper provides multiples strands of evidence, which when analysed in conjunction, suggest the presence of positive social spillovers. Hence, households are more likely to adopt LPG if other households residing in the same village or urban block do so. We find divergence in the strength of this effect between rural and urban households, with more persistent spillovers amongst rural households. We also find that spillovers are stronger in states that have previously had low rates of adoption, supporting the idea of an S-shaped pattern of technological adoption. Spillovers are also found to be stronger for households that are members of associations or social networks such as caste associations. Our results provide evidence on convergence in LPG use rates across subgroups of the Indian population, and have strong implications for policy-makers who could leverage lessons from social learning to encourage consumers to switch to cleaner sources of energy in developing countries.

Keywords: Clean cooking fuels; LPG; Technological adoption; Social learning; India

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

---

*This version: November 2016

*Corresponding author. Department of International Economics, Graduate Institute of International and Development Studies, Geneva Switzerland and Centre for International Environmental Studies, Geneva, Switzerland. E-mail:suchita.srinivasan@graduateinstitute.ch. Funding from the Swiss Network For International Studies (SNIS) is gratefully acknowledged.

†School of Forestry & Environmental Studies, Yale University 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: stefano.carattini@yale.edu. Funding from the Swiss National Science Foundation, grant P2SKP1_165028, 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 (WHO, 2016). Indoor air pollution still remains one of the leading causes of death in low and middle-income countries. Smoke generated from burning wood contains harmful pollutants such as carbon monoxide and particulate matter. Almost three billion people in the world still cook and heat their homes using open fires and stoves that burn biomass such as firewood, animal dung, crop waste and coal, and almost 4.3 million people die prematurely each year due to illnesses that are directly related to the inefficient use of solid fuels (WHO, 2016).[1] IHME (2013) estimates that 2.9 million deaths were caused by ambient air pollution in 2013 due to PM 2.5 (particulate matter with diameter less than or equal to 2.5 micrometers). Causes of death range from pneumonia, stroke, heart disease, chronic obstructive pulmonary disease, to lung cancer. Approximately 50% of premature deaths due to pneumonia among children under five are caused by soot that is inhaled due to indoor air pollution (WHO, 2016).[2]

In addition to health implications, the use of solid biomass also has implications for the global environment. Bond et al. (2007) estimate that cooking with traditional biomass accounts for almost 18% of greenhouse gas emissions. In addition, their use also degrades local forests and ecological systems. For instance, burning of firewood to produce charcoal has been found to expedite the degradation of land, including arable land(OECD/IEA, 2006). These environmental concerns are particularly pressing in light of India's commitments in the recently concluded Paris Agreement, where it pledged to create a carbon sink of 2.5 to 3 billion tonnes of CO_2 equivalent by increasing forest and tree cover, and to reduce its energy emissions intensity by 30-35% by 2030 compared to 2005 levels (UNFCCC, 2015).

The adoption and sustained use of clean cooking fuels, and efficient cook-stoves, 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 rather slowly. From the beginning of its entry in the Indian market, LPG was better supplied

[1] Solid fuel use is still common in South Asia, Africa and Latin America: in India, for instance, according to WHO estimates, almost 81% of the rural population still use solid fuels for cooking, and 26% of the urban population still relied on their use in 2013 (WHO, 2016).

[2] The risk of exposure is particularly high amongst women and children, who mostly stay indoors, and spend considerable amounts of time near open fires.

to urban areas, where its use was thus more prevalent, while rural adoption rates lagged behind. The Indian government, in order to incentivise consumers to switch to LPG, subsidised the cooking fuel considerably, but these subsidies have been mainly targeted to urban areas, which further contributed to widen the gap between rural and urban areas in terms of adoption rates.[3]

The literature abounds on the role of socioeconomic factors in determining which households use clean cooking fuels in developing countries (see Lewis and Pattanayak 2012 for a literature review). However, one aspect of this decision that has not been extensively studied is whether households are influenced by households in making this decision, i.e. whether social interactions between households could explain the wide disparities in adoption of clean cooking fuels that is observed in countries like India. While it is clear that income and access to LPG have played a significant role in revealing the pattern of LPG adoption in India, in this paper, we examine whether social learning and social spillovers could also affect a household's choice of cooking fuel, and thus contribute to the wide variations we observe in its adoption.

In particular, the objective of this paper is to investigate whether social spillovers exist in the use of LPG in India, and if they do, how they vary across rural and urban areas, states and within pre-existing social networks. By explicitly controlling for factors found to be important in the literature, and incorporating a rich set of socio-economic 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. We provide complementary evidence on the presence of positive social spillovers, i.e. a household is more likely to adopt LPG if other households residing in the same village or urban block do so. In addition, we find that these effects are more persistent amongst rural households than urban households, and that they are weaker in states which have high prior rates of LPG adoption. Lastly, we find that social spillovers may be stronger for households that belong to certain groups or associations such as caste associations, suggesting that social networks might play an important role in encouraging households to switch to clean energy sources.

This paper uses two sets of survey data on household-level consumer expenditure, which are nationally representative and large-scale. The biggest benefit of using this data is that its large scale

[3] Despite significant reductions in subsidies over the years, as of 2015-16, there is still an LPG subsidy of Rs. 150.82 (approximately USD 2) per litre per cylinder (PPAC, 2015).

allows us to compare the adoption of cooking fuels across all areas of the country, and across very heterogeneous sets of populations, both in terms of socio-economic characteristics, and in terms of governance and policy implementation. The first dataset that we use is the National Sample Survey (NSS) Household Consumer Expenditure Survey, which comprises repeated cross-sections. The second is the India Human Development Survey (IHDS) Consumer Expenditure dataset, which is a two-year panel (2005-06 and 2011-12).

Our strategy in this paper is to provide multiple pieces of evidence, which when taken together, provide complementary evidence on the presence of social spillovers in LPG use. Using the NSS data, we first estimate a linear probability model 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. Furthemore, we employ an instrumental variable two-stage least squares (IV-2SLS) approach to control for potential endogeneity in this estimation. With the IHDS data, we are able to use fixed effects in carrying out these estimations which helps us to control for time-invariant unobserved heterogeneity. Our results are robust across specifications, and hence support the existence of social spillovers in the use of LPG in the Indian context.

This paper contributes to a nascent 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 credible and transparent multi-layered empirical design. Focus on the use of LPG, and thus the continued use of clean cook-stoves, is crucial in this case. As shown by Hanna et al. (2016), ownership of cook-stoves does not necessarily imply that the cookstoves are used over a long period of time. Use of clean cook-stoves require indeed special maintenance, which may require skills and knowledge that can be obtained through social learning. While sources of fuel such as firewood are available freely, and at significantly lower costs, we provide evidence that households can be influenced by other households residing in the same village or urban block to purchase and use regularly clean energy alternatives.

These findings suggest the possibility that social interactions may contribute to explaining the spatial disparities that are observed in the adoption of LPG in India. Our results are relevant to all 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 behaviour 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).

The structure of the paper is as follows: section 2 provides a background on cooking fuel use in India and a review of the literature, section 3 provides a theoretical framework, section 4 elaborates on the data used and the empirical methodology, section 5 presents the empirical results and discuss potential policy implications, and section 6 concludes.

2 Background and Literature Review

2.1 Background on Cooking Fuel Use in India

Several sources of energy are used as cooking fuels in India, and the energy choice typically varies between rural and urban households. Rural households have strong preferences for biofuels 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 data provided in the 2011 Census, almost 67% of the overall Indian population still relies on solid fuels such as firewood, crop residue, dung cakes and coal as the primary cooking fuel, and the proportion is almost 85% among rural households. This may have to do with affordability and easy availability, but also with cooking habits and preferences, which have not changed over time.

Ample scientific evidence suggests that burning traditional biomass as a cooking fuel in homes leads to indoor air pollution, and that fuels like LPG are much cleaner in terms of their environmental impact compared to sources such as firewood (Boy et al. 2000, WHO 2016, Singh and Gundimeda 2014). In this paper, we choose to restrict our attention to use of LPG as the clean cooking fuel alternative. This is because it is the most widely available clean cooking fuel, and the most affordable among clean options. LPG is currently being used by most urban households, and increasingly by many rural ones.

Acquiring an LPG connection requires an initial expenditure to purchase the stove, and install the equipment. Households are required to 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 commonly observed amongst many Indian households, where a mixture of modern and traditional fuels are used simultaneously Cheng and Urpelainen (2014).

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 amongst rural and urban households, whereas in 2010, 60% of urban households used LPG, without stacking biomass fuels, while only 10% of rural households did so Cheng and Urpelainen (2014). Rural households are often not able to afford these recurrent expenditures needed to acquire the cylinders, and also have difficulties in purchasing the cooking stoves. LPG users are also required to have a permanent and verifiable residential address, which limits the access of poor or homeless people, even in urban areas (Gupta and Kohlin, 2006). 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 subsidised LPG (and kerosene) for many years, although in recent times efforts are being made to phase these subsidies out.

It seems however clear that even though LPG is subsidised to meet the requirements of poor households, the benefits of these subsidies have largely accrued to the richer urban households. According to some 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. 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 (IISD 2014).

Recent 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 considerable impact in improving the actual disbursement of subsidies. For instance, in September 2012, the central government capped the number of subsidised cylinders that a household can acquire at six per year. On January 2013, however, the limit was increased to nine cylinders per household per annum, which was further increased to 12 by 2014. 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 this paper we analyse whether the presence of social spillovers may justify additional initiatives targeting subsidies to specific sub-populations, and so increasing their effectiveness.

2.2 Literature Review

There is a growing literature looking at the adoption of clean cooking fuels and improved cookstoves (ICS) in developing countries, including in the Indian context. A significant strand of this literature has focused on air pollution borne out of the continuous use of solid biomass for cooking, and the associated negative health implications (Ezzati et al. (2000), Boy et al. (2000), Zhang and Smith (2007), Romieu et al. (2009)). A key finding that has emerged is that insufficient use of ICS, and their improper maintenance, is prevalent in developing countries, which limits their health benefits. 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.

However, providing the stoves for free may also not be sufficient. In a recent paper, Hanna et al. (2016) use experimental data for India and find that distributing clean cookstoves to poor, rural households leads to lower pollution and improved health outcomes, but only in the short run. This result was borne by the fact that households in the sample were not maintaining the cookstoves, and used them irregularly. Regular maintenance of cookstoves is crucial to guarantee improved health outcomes (Duflo et al. 2008).

The second significant strand of the literature on cooking fuels and ICS has focused on the role of socio-economic determinants of the adoption of clean cooking fuels by households. Lewis and Pattanayak (2012) provide a comprehensive summary of several studies which have looked at the determinants of cooking fuel choice in low and middle-income countries. Income, education, and urbanisation 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).

Another aspect of the transition to clean cooking fuels which has been studied in the Indian context is the phenomenon of fuel stacking: Cheng and Urpelainen (2014) find that from 1987 to 2010, many Indian rural households began using LPG, but continued to use firewood as well. One of the most important reasons for this behaviour is the need to diversity and rely on multiple sources of cooking, hedging against variations in the price of the fuel, and uncertainty in its supply.

Socio-economic determinants need not be the only factor influencing household adoption decisions of clean fuels, or clean technologies. In this respect, a growing literature has examined the role of social spillovers, or how the decisions of a household's neighbours, social network or friends may influence its own decisions, in the context of energy-related consumption choices. The literature on developed economies has looked at the role of spillovers, or "peer-effects", in explaining the adoption of green technologies like solar panels, for instance. 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%. The authors use the lag between the time of adoption and delivery of the panel for identifying the magnitude of this effect, and find that it increases over time. Graziano and Gillingham (2015) also study the diffusion of PV panels in Connecticut, and find a similar pattern. They also use a rich set of controls related to the built environment, and socio-demographic factors. Additional literature confirms the results for solar panels, and provides new evidence for other green technologies, such as hybrid cars (see Carattini 2015 for a review). There is also a branch of the literature on social capital and economic determinants which looks at whether membership in organisations could play a role in increasing prosperity (cf. e.g. Putnam et al. 1994; Knack and Keefer 1997).

The literature on social spillovers in the adoption of new technologies in developing countries has mainly focused on agricultural issues. For instance, Munshi (2004) finds evidence of social learning amongst farmers in the adoption of wheat and high-yield varieties of rice and wheat in India. Bandiera and Rasul (2006) study the adoption decisions of farmers in Mozambique, and find that farmers are more likely to adopt a new crop if a few farmers in their network adopt, but it may no longer be in their interest to adopt it if too many farmers in their network do so. The only paper, to our knowledge, dealing with social spillovers in the adoption of green technologies in developing countries is Beltramo et al. (2015). The authors study the adoption of efficient cookstoves using data from a randomised control trial in Uganda, and do not find evidence of peer effects or social interactions.

In this paper, we contribute to this recent literature, by focusing on the adoption of LPG in India. We use pan-Indian data to study whether social spillovers in LPG use exist for households that reside in the same geographical area, i.e. the same village or urban block. We then investigate whether these effects vary for rural and urban households, for households residing in states with high prior levels of LPG adoption and states where the use of LPG is beginning to diffuse, and for households belonging to social networks, such as credit and savings organisations, religious and social groups, and caste associations. We use multiple datasets to provide evidence that spillovers may be an important, and thus far overlooked, factor in explaining LPG adoption by Indian households. We use a rich set of controls in the empirical estimations, and estimate different models to support our hypotheses.

To introduce the mechanisms that may lead social spillovers to affect the household-level cooking fuel adoption, we outline a theoretical framework in the next section.

3 Theoretical Framework

In this section, we develop a theoretical framework of social spillovers in the use of LPG. This model builds on flows of information across households about LPG, and how this may influence a particular household's decision to adopt it. In characterising these flows, we do not restrict them to be one-time transfers of knowledge, i.e. households can continuously learn from each other about the existence or availability of LPG, but also about its use, maintenance and benefits. In this respect, our framework is consistent with Hanna et al. (2016), who suggest that it is not sufficient to inform households about clean cook-stoves as a cleaner option, but it also necessary to ensure that they are regularly used, and maintained.

3.1 Notation and Model Setup

The theoretical framework developed in this section is an extension of the social learning model of innovation diffusion developed in Young (2009). Following this model, the various sources of heterogeneity across agents (households, in this case) are reduced to a single threshold, which summarises the likelihood of the household adopting LPG, given the information that has been generated by other LPG users in the sample. The example which this model is built on is that of standard normal-normal belief updating (De Groot (1970)).

To retain the notation of Young (2009), let X be a random variable which denotes the payoff gain from using the new technology (or cooking fuel, in this case) compared to the incumbent one, distributed with mean μ and variance σ^2 (independent and identically distributed across households, and time periods). Let c_i denote the household-specific cost of adoption, such that if μ is known, and μ is greater than c_i, then household *i* adopts the technology, otherwise it does not. If μ is unknown, each household formulates beliefs about it. Let μ_{i0} denote household *i*'s initial beliefs about μ, and let τ_i denote its "rigidity" of beliefs, such that low values of τ_i indicate that household *i* is very amenable to adapting its beliefs.

The essence of the model is that it allows the adoption of LPG to depend upon how much "social interaction" household *i* has with other households (denoted as β_i >0). In Young's model, this parameter denotes the extent to which household *i* "gets around", and it is assumed to be time-invariant. As in Young (2009), let N_{it} denote a Poisson random variable which is *i*'s total number of observations of household *i* up to period *t*. In divergence to Young's model, we express the cumulative information generated up to time *t* as the sum of two terms

$$E[N_{it}] = \beta_i({Y_t}^{\mathrm{A}} + {Y_t}^{\mathrm{N}}) \tag{1}$$

where Y_t^{A} denotes the number of adopters up to time *t*, and Y_t^{N} denotes the number of non-adopters up to time *t*. In the context of this paper, these can represent the number of adopters and non-adopters in a certain geographic entity such as state, district or village/urban block, or even in a particular social network. N_{it} thus denotes the number of households that household *i* has observed till period *t*. We assume the regularity condition that $\tau_i \geq E[N_{it}]$, i.e. that the households initial beliefs are sufficiently rigid.

Retaining the Bayesian updating model of Young (2009), the expression for the posterior estimate of the payoff gain μ_{it} can be written as

$$\mu_{it} = \frac{n_{it}\bar{x_{it}} + \tau_i \mu_{i0}}{n_{it} + \tau_i} \tag{2}$$

where n_{it} denotes a particular realisation of N_{it}, and $\bar{x_{it}}$ denotes the sample payoff gain from using the new technology among n_{it} observations (it follows the normal distribution, with mean μ and variance $\frac{\sigma^2}{n_{it}}$).

We will now use this framework to derive a set of hypotheses, which we will test in the empirical section of the paper.

3.2 Derivation of Expression for Expected Benefit of Adoption

Our objective in developing this model is to focus on these flows of information across households, and to pinpoint how spillovers may vary across different categories of households. We denote these spillovers as "social spillovers" and remain agnostic about their exact nature. Since the decision to adopt LPG is driven in this model by the expected benefits from its use, we derive the following expression:

$$B_{it} = \mu_{it} - c_i \tag{3}$$

Thus, household *i* will adopt the new technology if B_{it} is greater than or equal to zero, i.e. $\mu_{it} \geq c_i$. Substituting the expression for μ_{it} derived in (3.1) above, B_{it} can be rewritten as

$$B_{it} = \frac{n_{it}\bar{x_{it}} + \tau_i \mu_{i0}}{n_{it} + \tau_i} - c_i \tag{4}$$

Given that $\bar{x_{it}}$ is normally distributed with mean μ and variance $\frac{\sigma^2}{n_{it}}$, this is equal to

$$B_{it} = \frac{n_{it}(\frac{\sigma(z_{it})}{\sqrt{n_{it}}}) + \tau_i \mu_{i0}}{n_{it} + \tau_i} - c_i \tag{5}$$

Substituting the expression for N_{it} derived above, we get the final expression for B_{it} as

$$B_{it} = \frac{\sigma(z_{it})\sqrt{\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t)} + (\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t))\mu + \tau_i \mu_{i0}}{\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t) + \tau_i} - c_i \tag{6}$$

We now assume that the household only takes into account the expected benefit of using the new technology, i.e. the expression becomes

$$E[B_{it}] = \frac{\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t)\mu + \tau_i \mu_{i0}}{\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t) + \tau_i} - c_i \tag{7}$$

3.3 Hypotheses

We assume that the parameter β_i is higher for urban households than for rural households, i.e. $\beta_i^U \geq \beta_j^R$, for any households *i* and *j*. This is based on the observation that, given the higher population density in urban areas, urban households tend to have more opportunities to interact with other households, and may be more easily exposed to new technologies and habits. Define the share of adoption up to time *t*, λ_t, as

$$\lambda_t = \frac{{Y^{\mathrm{A}}}_t}{{Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t} \tag{8}$$

Hypothesis 1: A household is more likely to adopt LPG, if more households residing in the same geographical area, or belonging to its social network, adopt it, i.e. the social spillovers are positive.

Proof: Given expression (7) above, we can derive the following expression for $\frac{\partial E[B_{it}]}{\partial {Y^{\mathrm{A}}}_t}$:

$$\frac{\partial E[B_{it}]}{\partial {Y^{\mathrm{A}}}_t} = \frac{\tau_i \beta_i(\mu - \mu_{i0})}{(\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t) + \tau_i)^2} \tag{9}$$

It is straightforward to show that $\frac{\partial E[B_{it}]}{\partial {Y^{\mathrm{A}}}_t} \geq 0$, if we assume the regularity condition that $\mu \geq \mu_{i0}$, i.e. that the actual population payoff gain is higher than household *i*'s initial beliefs. If this condition holds, then household *i*'s expected benefit from adopting LPG increases as the number of adopters increase, and thus the social spillovers are positive in nature.

Corollary 1: Social spillovers are stronger amongst households interacting in social networks. Proof: This follows from the proof of Hypothesis 2, i.e. the social spillovers are stronger for households that have higher β_i. This would imply that households participating in some kind of network (where they have the possibility of closely interacting with more households) will experience stronger spillovers.

Hypothesis 2: Spillovers will be weaker in areas that already have high rates of adoption of LPG, i.e. the spillover effect will be stronger amongst households residing in areas where diffusion of LPG is at its preliminary stages.

Proof: Substituting for ${Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t$ in terms of λ_t and ${Y^{\mathrm{A}}}_t$ from expression (8) into (7), it can be shown that $\frac{\partial E[B_{it}]}{\partial {Y^{\mathrm{A}}}_t \partial \lambda_t} \leq 0$, i.e. the magnitude of the spillover decreases as the rate of adoption at period t (λ_t) increases (this assumes that the regularity condition on τ_i holds).

This implies that spillovers may be stronger in states which start out having higher rates of adoption, and is consistent with the S-shaped curve commonly used to explain the diffusion of new technologies (Bass, 1969).

Hypothesis 3: The spillovers decreases in strength as the number of adopters increase, with the decline being greater for urban households.

Proof: The expression for $\frac{\partial E[B_{it}]}{\partial {Y^{\mathrm{A}}}_t}$ derived in Hypothesis 1 above can be used to evaluate the second-order derivative

$$\frac{\partial^2 E[B_{it}]}{\partial {{Y^{\mathrm{A}}}_t}^2} = \frac{-2\tau_i \beta_i^2 (\mu - \mu_{i0})}{(\beta_i({Y^{\mathrm{A}}}_t + {Y^{\mathrm{N}}}_t) + \tau_i)^4} \tag{10}$$

From expression (10), it is clear that $\frac{\partial^2 E[B_{it}]}{\partial {{Y^{\mathrm{A}}}_t}^2} \leq 0$, i.e. while the spillovers themselves are still greater than or equal to zero in magnitude, they can be expected to weaken as more households become users of LPG. This is intuitive, and closely related to the S-shaped diffusion curve for new technologies (Bass, 1969).

Given expression (10) above, we can show that $\frac{\partial \frac{\partial^2 E[B_{it}]}{\partial {{Y^{\mathrm{A}}}_t}^2}}{\partial \beta_i} \leq 0$. Using our assumption at the beginning of the section ($\beta_i^U \geq \beta_j^R$), the implication is that the weakening of spillovers is stronger in urban areas with a higher population density.

4 Methodology

4.1 Data

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, and thus to test the hypotheses that we derived in the previous section. We use two sets of data for the empirical analysis. The first set of data employed is from 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 Programme Implementation of the Indian government (National Sample Survey Office and Programme Implementation, 2199). 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 43rd, 55th, 61st and 66th 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 (over 100000 households) to provide ample geographical heterogeneity in the data, in order to examine spatial disparities. The NSS data allows us to attribute to each household the district and the state of residence. In addition, the NSS data provides us with coded information for the urban block or village to which each household belongs. From this, 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 we use in this paper 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 (2009), Desai and Vanneman (2015)). It is a panel dataset, with two rounds of data available (2005-06 and 2011-12). 83% of the households sampled in the first round also respond in the second round. The panel nature of the data enables us to track changes in LPG adoption over time. This panel dataset, composed of about 40000 households, thus complements the large cross sections of the NSS.

In both datasets, 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, most of the rounds ask the households for expenditure over the previous 30 days.[4]

Our empirical approach uses data from the NSS on expenditures of all types of fuels purchased by the households, along with the respective quantities and values of the purchase), and information on which is the primary fuel used by the household, both for cooking and lighting purposes. This information is particularly useful, given that fuel stacking is commonly observed amongst households in India, where multiple fuels are used at the same time. In the IHDS data, we restrict the sample to the households for whom we have valid information on whether they spent on the LPG fuel in the last 30 days, and those that primarily use the fuel for cooking purposes.

The measure of LPG adoption (our dependent variable) in the NSS data is a binary variable for whether LPG is the primary cooking fuel of a household or not. The IHDS data do not have the exact same variable. Our measure of adoption for this data is represented by whether the household spent on LPG in the last 30 days. Both variables eschew the possibility of irregular use of LPG, which may be a potential problem with using initial LPG adoption as a measure.

[4] The 66$^{\text{th}}$ 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^{th} 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 expenditure on the same items. To ensure comparability with the other rounds, we only use observations for which the 30-day window was used.

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

Statistic	NSS				IHDS
Round	43	55	61	66 (Type 1)	Overall Panel
Year	1987-88	1999-00	2004-05	2009-10	2005-06 and 2011-12
Households Sampled in Village/urban block : Mean	9.95	11.96	9.98	7.98	33.88
Households Sampled in Village/urban block : Min.	2	2	3	2	4
Households Sampled in Village/urban block : Max.	10	12	10	8	88
LPG Adoption Rate at Village/Urban Block level: Mean (%)	10.5	25.82	29.79	41.24	65.57
Observations	104874	103094	97998	67374	18179

Notes: Values reported are calculated only for observations included in the regression sample. Villages and urban blocks in the IHDS data comprise 150-200 households.

Table 1 provides information on the mean, maximum and minimum number of households by village or urban block in both datasets, and the mean LPG use rate (at the village/urban block level). Figure A1 in Appendix A shows the distribution of households by primary fuel-type (for cooking purposes) in the four thick rounds of the NSS. As it is clear from these graphs, firewood was, and still is, the primary cooking fuel for a majority of the households. The popularity of LPG has increased over this period, and as of 2010, it is 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.

Figure A2 focuses on LPG only and plots the evolution of the proportion of households for whom LPG is the main cooking fuel, and shows how it has gained popularity, especially in recent periods. We also observe that 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 A3 shows the regions which have contributed most to the increase in popularity 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 states, such as Maharashtra, Tamil Nadu and Karnataka, experienced an increase in the share of LPG adopters.

Households are also asked about expenditure on durable goods (such as cookstoves) in the last 365 days. Both datasets compile information on the demographic characteristics of all the members of the households surveyed, including the age, gender, marital status, industry of occupation, and level of education. Information is also provided on land ownership (total land possessed, whether land is rented, irrigated, etc.), and the physical characteristics of the house (such as the type of structure, the condition of the house, type of floor, etc.).

Table 2 and Table 3 provide summary statistics on some of the demographic characteristics of the households using NSS and IHDS data, respectively.

4.2 Empirical Approach

In order to investigate the presence of spillovers in LPG use, we adopt a multi-pronged approach by providing evidence from both cross-sectional and panel datasets on the hypotheses developed in the theoretical framework. In this section, we describe the econometric models that we estimate using both cross-sectional and panel data. Our first model uses cross-sectional data for testing the presence of social spillovers across households located in the same geographical area, i.e. the same village or urban block.

Using data from the four thick rounds of the NSS, we first estimate for each round a linear probability model (LPM) of the form[5]:

$$A_i = \alpha_0 + \alpha_1 A_{-ij} + \alpha_2 X_i + \mu_i \tag{11}$$

for each round, where the dependent variable is denoted by A_i, a binary variable indicating whether LPG is the primary cooking fuel of household *i*, and the main independent variable is A_{-ij} , the average LPG adoption rate amongst all households (other than household *i*) in village/ urban block *j*. X_i denote household-specific controls, such as household size, age, gender and level of education of the head of the household, whether the household has access to electricity, firewood, monthly per capita expenditure (MPCE) dummies, and prices of LPG and kerosene. It also includes

[5] For robustness, we also estimate a non-linear logit model (see Table B1 in the Appendix). Coefficients for the main variables of interest remain unchanged with respect to those obtained from the LPM. This is also valid for the Probit model (all additional results are available by the authors upon request.

Table 2: Summary Statistics of NSS Data

Round		43			55			61			66	
Year		1987-88			1999-00			2004-05			2009-10	
Variable	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
Household 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 household head is female (%)	10	30	127,932	8.8	28.33	119,639	9	29	124,637	8	28	100,845
Whether household 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 household lives in a district adjoining a district with big urban centre (%)	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 the household has access to electricity (%)	43	49	128,031	64.41	47.88	119,639	73	45	124,637	83	38	100,845
Whether the household purchased a cook-stove 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 household 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 Data (2005-06 and 2011-12)

Year	2005-06			2011-12		
Variable	Mean	Std. Dev.	Obs	Mean	Std. Dev.	Obs
Whether household spent on LPG in the last 30 days (%)	59.5	49.1	22703	99.8	3.41	22781
Whether household has access to electricity (%)	95.9	19.8	20717	99.99	0.01	22781
Proportion of rural population (%)	60.6	48.9	22703	58	49.4	22781
Size of household	5.79	2.95	22703	4.83	2.29	22781
Number of years of education for household head	8.72	4.9	22673	9.40	4.97	22772
Income (Rs./year)	67506.37	97836.31	22703	155057.3	261935.1	22781
Whether household uses a non-biomass cook-stove (%)	33.6	47.2	22703	32.1	46.7	22781
Number of hours of cook-stove use (/Day)	3.25	1.68	22675	2.92	1.30	22552
Time taken to collect fuel (mins/one-way trip)	29.77	33.47	13825	55.37	46.10	6230
Whether kitchen has a vent (%)	0.68	0.47	20490	0.67	0.47	22781

a control for whether the household resides in a district which is adjoining a large urban centre.[6,7] μ_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.

In this framework, potential threats to identification may be due to endogeneity, in particular in relation to the problem of "reflection" or "simultaneity" (Manski 1993, Manski 2000, Moffitt et al. 2001): a household's choice to use LPG as the primary cooking fuel, may in turn, influence the other households' choices. 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*. In addition, in spite of the large set of controls used in our specifications, common unobservable factors may, in principle, still affect the observed decision of households to adopt LPG. Lastly, there may be endogenous group formation, even though our detailed data should address most of the concerns on self-selection of peers.

To address potential threats to identification, we apply an instrumental variable approach following Duflo and Saez (2002), which is an adaptation of an earlier application of instruments used by Case and Katz (1991). Duflo and Saez (2002) study whether there are peer effects among

[6] 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 enable access to LPG. The urban centres 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 > 4 million, while a tier-II city is defined as one with population between 1 and 4 million).

[7] The NSS data does 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 subsamples. 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 that did not actually purchase LPG in a given year.

colleagues in the same department of a university in participation in retirement plans, and find that the choice of employees to enrol in these plans, and the choice of vendor, were influenced by the decisions made by colleagues. To causally assess the existence of peer effects, they instrument average participation in each peer group by the salary or tenure structure of that group. We follow their methodology, and use the proportion of population of each village (or urban block) belonging to the highest monthly per capita expenditure (MPCE) decile as an instrument for average LPG adoption at the village/urban-block level, which is the endogenous variable in the second-stage.[8] The model that is estimated is thus the same as above, but A_{-ij} is treated as endogenous. We apply this approach to each of the four thick rounds of the NSS data.

To alleviate residual concerns with identification, we exploit the potential of the panel IHDS data to estimate a fixed effects linear probability model. The model that we estimate is:

$$A_{it} = \alpha_0 + \alpha_1\delta_i + \alpha_2 A_{-ijt} + \alpha_3 X_{it} + \alpha_4\tau_t + \mu_{it} \tag{12}$$

where A_{it} is a binary variable indicating whether household *i* spent on LPG in the last 30 days prior to the date of the survey (as of time period *t*). δ_i is the household-specific fixed effect which captures time-invariant unobservable characteristics of every household.[9] The independent variable of interest is A_{-ijt} , the average LPG adoption rate amongst all households (other than household *i*) in village/ urban block *j* in time period *t*. X_{it} now include potentially time-varying household level characteristics, such as the size of the household, level of education of the household head, income dummies, and some cooking fuel and cook-stove related controls such as whether the household uses a non-biomass cook-stove, the amount of time the household spends in collecting firewood, how many hours a day that the cook-stove is used, and whether the household has a kitchen with a vent. τ_t denotes a household-specific time trend, while μ_{it} denotes the error term.

For the reasons mentioned before, we also estimate an IV-2SLS model following Duflo and Saez (2002). In choosing the instruments, we adopt the same methodology as with the NSS estimations,

[8] 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.

[9] In order to be able to partially capture the effects of time-varying unobservables, we also estimate models using village-by-year time trends. These results are included in the appendix

i.e. we use the proportion of population in the same village or urban block belonging to the highest income deciles.

In order to test Corollary 1, we also estimate model (12) above only for households that declare to participate to certain networks or social groups within the same village or urban block. These include credit and savings associations, caste-based groups and religion-based communities. Given a non-negligible loss in observations associated with testing for these effects, we perform only standard OLS estimations.

Hypothesis 2 of the theoretical framework suggests that spillovers might be weaker in states which have a historical advantage in terms of LPG adoption, because these states are further along the S-shaped diffusion curve. We test this hypothesis by using the 61st round of the NSS data (from 2004-05) because it represents the last round of the NSS just before the time period of the IHDS sample begins, which allows us to get a sense of the pre-sample trends. Following the distribution as observed in the descriptive statistics, we create four dummies for the following adoption rates: below 20%, between 20-30%, between 30-40% and more than 40%. We then interact these dummies with the observed level of adoption in the village or urban block, to analyse how the spillovers change across the different states. We estimate a linear probability model with fixed effects to study test Hypothesis 2.

In order to test Hypotheses 3, we estimate the models (4.2) and (12) for rural and urban sub-populations. In the NSS estimations, this involves using data from the four thick rounds.

5 Results

5.1 Empirical Results

This section presents the results of the empirical estimations. Table 4 below presents the results of the models estimated using NSS cross-sectional data from the four thick rounds of the survey. This includes the estimations of the linear probability models (in the odd-numbered columns) and the instrumental variable probit models (in the even-numbered columns).[10]

Our variable of interest is the level of LPG use at the urban block or village level. In all rounds

[10] The results of the logit model are included in Table B1, and the first-stage results are included in Table B2, both in the appendix.

and specifications, we find that the coefficient on this variable is positive, and often significant at the 1% level. Hence, this provides evidence for Hypothesis 1, which posits that social spillovers across households in LPG use are positive. The coefficients should be interpreted as follows: in the 43rd round, a one unit increase in the average village/ urban-block LPG adoption rate increases the probability that household i adopts LPG by about 0.62 units in the LPM, and by about 1.75 units according to the IVM. The magnitude of the coefficient is higher in the IVM than in the LPM, which suggests the linear estimates may be underestimating the strength of this effect. The results of Table 4 also suggest that the spillovers are larger in magnitude in the older round, and decrease in the most recent ones.

We now present an overview of the results with regards to the control variables. All coefficients have the expected signs. We find that households in more developed areas, as proxied by access to electricity, are more likely to adopt LPG, as our initial hypothesis suggested. Proximity to a big urban centre is insignificant in most specifications. Households having heads that are older, female, or more educated are also more likely to use LPG as the primary cooking fuel. On the other hand, houses facing a high price of LPG, or those with access to firewood are less likely to use LPG. our results also suggest that larger households are more likely to adopt LPG, a common yet not undisputed finding (cf. Lewis and Pattanayak 2012 for a discussion).In this model, we also control for income using dummies for monthly per capita expenditure deciles, which are significant in every round. This supports previous findings of the literature, suggesting that income is an important determinant of the decision of a household to switch to cleaner cooking fuels.

Table 5 below presents the estimation results using the IHDS panel data. The estimates of columns (1) and (2) in Table 5 indicate that the variable for average LPG adoption at the village/urban-block level has a positive coefficient, confirming the results of Table 4, and providing further evidence in support of Hypothesis 1. All coefficients are significant at the 1% level. The coefficients are comparable across models; the results suggest that as with the NSS data, the magnitude of the coefficient is larger in the IV-2SLS estimation than in the LPM. The magnitude of the coefficient is comparable to those obtained using NSS data, namely a one unit increase in the adoption rate of LPG at the village level leads to about a one unit increase in the probability of household i spending on LPG in the LPM, and to about a 1.16 unit increase in the probability of household i spending on LPG in the IVM.

Table 4: NSS Data Linear Probability Model (LPM) and Instrumental Variable Probit Model (IVM) Results

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	LPM	IVM	LPM	IVM	LPM	IVM	LPM	IVM
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average LPG use rate (Village/ Urban Block)	0.619***	1.748***	0.468***	1.080***	0.353***	0.801***	0.341***	-0.472
	(0.007)	(0.140)	(0.007)	(0.158)	(0.007)	(0.270)	(0.007)	(0.435)
Whether bordering an urban centre?	-0.053***	0.085	-0.051	-0.038**	-0.131***	-0.259	-0.025	-0.043
	(0.013)	(0.159)	(0.065)	(0.021)	(0.060)	(0.194)	(0.045)	(0.278)
Whether HH has access to electricity?	0.037***	0.635***	0.040***	0.703***	0.035***	0.657***	0.037***	0.636***
	(0.002)	(0.029)	(0.003)	(0.024)	(0.003)	(0.028)	(0.004)	(0.036)
Whether HH lives in a rural area?	0.011***	-0.483***	-0.007	-0.249***	-0.011***	-0.323***	-0.001	-0.647***
	(0.002)	(0.034)	(0.003)	(0.045)	(0.003)	(0.075)	(0.004)	(0.111)
Whether HH purchased a cookstove in last 30/365 days?	-0.033***	-0.169***	-0.107***	-0.415***	-0.021	-0.066	0.006	0.088
	(0.006)	(0.035)	(0.005)	(0.026)	(0.022)	(0.116)	(0.017)	(0.097)
Household size	0.009***	0.128***	0.017***	0.119***	0.015***	0.102***	0.019***	0.107***
	(0.0003)	(0.004)	(0.0005)	(0.003)	(0.0005)	(0.003)	(0.0006)	(0.006)
Age of head of household	0.001***	0.018***	0.001***	0.009***	0.0008***	0.006***	0.0008***	0.005***
	(0.00007)	(0.001)	(0.00009)	(0.0006)	(0.0001)	(0.001)	(0.0001)	(0.0008)
Whether head of HH is female	0.006***	0.157***	0.021***	0.187***	0.013***	0.142***	0.011***	0.126***
	(0.002)	(0.034)	(0.003)	(0.023)	(0.003)	(0.021)	(0.004)	(0.023)
Whether head of HH is educated	0.041***	0.732***	0.089***	0.638***	0.083***	0.623***	0.100***	0.564***
	(0.002)	(0.027)	(0.002)	(0.017)	(0.002)	(0.017)	(0.003)	(0.021)
Price of LPG	-0.00003	0.0004	-0.010***	-0.042***	-0.0004**	-0.007***	-0.004***	-0.019***
	(0.0005)	(0.002)	(0.002)	(0.006)	(0.0002)	(0.002)	(0.001)	(0.005)
Price of Kerosene	0.004***	-0.009	0.00004	-0.0007	-0.00001**	-5.26 0.003	-0.005	
	(0.001)	(0.018)	(0.00009)	(0.002)	(0.000006)	(4.17)	(0.002)	(0.010)
Whether HH has access to firewood	-0.134***	-1.147***	-0.229***	-1.084***	-0.396***	-1.600***	-0.435***	-1.874***
	(0.003)	(0.029)	(0.004)	(0.022)	(0.005)	(0.030)	(0.006)	(0.025)
Observations	104845	104148	102994	102994	97963	97933	67374	67372
R^2	0.4729		0.5775		0.6004		0.6381	
Wald test of endogeneity (Chi^2)		0.21		31.64		10.47		21.28
P-value		0.6438		0		0.0012		0

Notes: The proportion of population in the same village or urban block in the highest income decile is used as an instrument in even columns. For the IVM results, 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 social groups. The dummy for monthly per capita expenditure of the 10th decile is the variable of reference. 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 for 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 55th round. The IVM in the 61st round does not include controls for religion, as they would prevent the convergence of the model.

The results on the controls corroborate those obtained using the NSS data. Larger households, and households that have higher MPCE are more likely to spend on LPG, as are households with more educated heads. In addition, households that use a non-biomass cookstove, or have a vent in the kitchen, are also more likely to spend on LPG.

In column (3) of Table 5 we include an interaction term between indicator variables for the levels of LPG adoption in 2004-05, and our main independent variable. The results from the LPM estimation suggest that the spillover effects are significant at the 1% level, and of the highest magnitude, for those states which start with LPG adoption rates which are in the "middle" of the distribution, namely between 20% and 30%. Next are those states which have the lowest LPG adoption rates prior to the data sample period, namely those with adoption rates of less than 20%. The interaction term is insignificant in the states that are observed already with relatively high rates of adoption, namely those where the adoption rate was between 30-40%, or greater than 40%. This pattern is thus compatible with the S-shaped diffusion curve found to be relevant for many technologies. At low levels of adoption, diffusion is relatively slow, but as adoption increases, forces of contagion kick in, and at moderate levels of adoption, diffusion becomes relatively fast. Once the level of adoption increases further, the market tends to be saturated, and diffusion rates flatten out.

Table B3 provides the results of a set of robustness estimations using village fixed effects, and village-by-year time trends. As mentioned in the previous subsection, we do not have information on the access of households to LPG, or on the availability of retailers by village/urban block. Household level fixed effects are effective in accounting for the time-invariant unobservables at a household level, however supply of LPG may be better controlled for by using village/urban block-level fixed effects, as done in column (1). Village-by-year time trends are also used, in column (2), to further capture the effect of time-varying unobservables at the village/urban-block level. The IV-2SLS methodology is used for both models, and the first-stage results are provided in Table B5. The variable for the average adoption of LPG in the village/urban block is significant at the 1% level, and has a positive coefficient, as before. The magnitude of the coefficient is also larger than it was before, suggesting that if anything, the results of Table 5 provide a lower-bound estimate of the true effect.

We then analyse whether households active in certain social networks are more likely to adopt

Table 5: IHDS Data Baseline Results

Dependent Variable: Whether HH i spent on LPG in the last 30 days Column	LPM (1)	IV-2SLS (Overall) 2nd-Stage (2)	LPM (Using 1999-00 NSS LPG Adoption Rates) (3)
Average LPG use rate (Village/ Urban Block)	1.009***	1.157***	
	(0.006)	(0.032)	
Avg. LPG Use Rate * Indicator for 1999-2000 LPG adoption rate less than 20%			0.95***
			(0.013)
Avg. LPG Use Rate * Indicator for 1999-2000 LPG adoption rate between 20-30%			1.00***
			(0.012)
Avg. LPG Use Rate * Indicator for 1999-2000 LPG adoption rate between 30-40%			0.984
			(0.017)
Avg. LPG Use Rate * Indicator for 1999-2000 LPG adoption greater than 40%			-0.720
			(1.247)
Whether HH lives in a rural area?	-0.089	-0.109**	-0.050
	(0.062)	(0.060)	(0.061)
Whether HH has access to electricity?	0.032***	-0.028**	0.048***
	(0.008)	(0.017)	(0.015)
Household size	-0.002	0.007***	-0.005**
	(0.002)	(0.003)	(0.003)
Number of years of education of household head	0.004***	-0.00007	0.003***
	(0.001)	(0.001)	(0.001)
Whether household has non-biomass cookstove?	0.074***	0.016	0.075***
	(0.011)	(0.016)	(0.012)
Hours of cookstove use (/ Day)	-0.011***	-0.007***	-0.011***
	(0.003)	(0.003)	(0.003)
Time spent in collecting firewood (/ Trip)	0.00004	-0.0002***	0.00001
	(0.00007)	(0.0001)	(0.00009)
Whether household has vent in kitchen?	0.0134**	0.018***	0.013*
	(0.007)	(0.008)	(0.008)
Observations	18590	9350	17072
Hansen J-statistic		7.93	
P-value		0.1601	

Notes: All specifications include household-level fixed effects and time trends. Results in columns (1) and (2) include dummies for whether the household belongs to the 7th, 8th, 9th or 10th income deciles, whereas the result in column (3) includes all income decile dummies. Controls are included for social group and religion in all specifications. For the results in column (2), the first-stage F statistics to check for weak instruments are higher than the Stock and Yogo (2005) threshold values, and the value of 10 suggested by Staiger and Stock (1997) (cf. Table B5 in the appendix).*,** and *** respectively denote significance at 10%, 5% and 1% levels. The coefficients of the constant are not reported.

LPG. Not all respondents provide information on their belonging to social networks. Due to decline in the number of observations, we do not estimate an IVM in this case. Table 6 includes the results of the LPM estimation. According to Corollary 1 derived in the theoretical framework, the strength of the spillovers may be higher amongst households belonging to social networks, given that they are likely to enjoy more social interactions. Based on the classifications in the data, we consider the following social networks: women's groups, self-help groups, credits and savings organisations, religious and social organisations, and caste associations.

To test Corollary 1, we compare the coefficients provided in Table 6 for the relevant variable, average LPG use rate at the urban block or village level, with the equivalent coefficients in column (1) of Table 5. In the models presented in Table 6, the coefficients of this variable are all positive and significant at the 1% level. However, only in one case the coefficient of Table 6 is statistically higher than the respective coefficient in Table 5. This is the case of caste associations, in column (5). This result confirms the importance of castes as a determinant of social capital in Indian society, and the level of trust that may exist among members belonging to caste associations (Bouma et al., 2008).

Finally, we test for heterogeneity in spillovers between rural and urban households, thus testing Hypothesis 3 of the theoretical model. The results presented in columns (1) to (8) of Table 7 present the results of the NSS estimations, whereas columns (9) and (10) present the results derived from the IHDS data. The theoretical model hypothesised that while rural and urban households are both expected to experience positive social spillovers, these can be expected to decline as the number of adopters increases, and the decline is expected to be steeper amongst urban households, where population density is higher and thus the social learning process may proceed with greater speed.

Both sets of models (NSS and IHDS) are estimated using OLS. The results from columns (1), (3), (5) and (7) of the table suggest that the spillovers are positive for rural households in all rounds of the NSS, but there is no monotonic trend in their magnitudes across rounds. On the contrary, the results from columns (2), (4), (6) and (8) indicate that for urban households, these spillovers are weakening over time. All coefficients are significantly different from each other at the 5% level.

The IHDS results also broadly support these findings. The results in column (9) show that amongst rural households, the spillovers are positive and significant at the 1% level in 2005-06, but they weaken by 2011-12, as indicated by the negative coefficient on the interaction term. For urban households, the results of column (10) suggest that the positive (and significant) spillover effect in 2005-06 is no longer significant by 2011-12. That is, social spillovers decrease in both urban and rural contexts, but to the point of becoming statistically non-significant in the case of urban spillovers. Urban households with better access to LPG experience positive social spillovers, but these deteriorate relatively quickly as the number of adopters reaches a saturation level. Rural households, on the other hand, experience more persistent spillovers, given that they started using LPG relatively later. These spillovers can be expected to first strengthen, as adoption rates increase, and then slowly decline as rural LPG adoption rates also approach saturation point. This result provides additional evidence in support of a convergence in the rates of LPG adoption across regions in India.

Table 6: IHDS Data Social Network (LPM) Results

Dependent Variable: Whether HH *i* spent on LPG in last 30 days Column	Women's Group (1)	Self-Help Group (2)	Credit/Savings Organisation (3)	Religious/Social Organisation (4)	Caste Association (5)
Average LPG use rate (Village/ Urban Block)	1.004***	1.005***	0.867***	0.934***	1.066**
	(0.047)	(0.030)	(0.053)	(0.059)	(0.059)
Whether HH has access to electricity?	-0.196	-0.029	0.183**	0.194**	0.176
	(0.146)	(0.038)	(0.092)	(0.109)	(0.171)
Household size	0.0004	-0.003	0.001	-0.013**	-0.0006
	(0.011)	(0.007)	(0.010)	(0.007)	(0.014)
Number of years of education of household head	0.0004	-0.007	0.01	-0.003	0.006
	(0.007)	(0.004)	(0.008)	(0.007)	(0.006)
Whether household has non-biomass cookstove?	0.100*	0.083	0.154***	0.112*	0.077
	(0.060)	(0.053)	(0.053)	(0.063)	(0.080)
Hours of cookstove use (/ Day)	-0.016*	-0.002	-0.035***	-0.004	-0.011
	(0.010)	(0.010)	(0.010)	(0.011)	(0.013)
Time spent in collecting firewood (/ Trip)	-0.00002	-0.00007	0.0006	-0.0002	-0.00002
	(0.0004)	(0.0004)	(0.0006)	(0.0003)	(0.0006)
Whether household has vent in kitchen?	0.090*	0.052	0.034	-0.066	-0.107*
	(0.048)	(0.034)	(0.037)	(0.044)	(0.060)
Observations	1576	2278	1727	2460	1920
R^2 (Overall)	0.6324	0.7252	0.5536	0.6092	0.6199

Notes: All specifications include household-level fixed effects and time trends. All specifications include dummies for income deciles, and controls for social group and religion. *,** and *** respectively denote significance at 10%, 5% and 1% levels. Standard errors are clustered at the village/ urban-block level. The coefficient of the constant has not been reported.

5.2 Policy Implications

Interesting policy implications can be derived from our results. We find that there may be 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 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 employ different strategies to correct for possible endogeneity in these

Table 7: NSS and IHDS Results on Rural and Urban Households

Data	NSS								IHDS	
Year	1987-88		1999-00		2004-05		2009-10		2005-06 and 2011-12	
Round	43		55		61		66		Panel Data	
Dep.Var.: Whether prim. cooking fuel of HH i is LPG (NSS) or whether HH i spent on LPG in the last 30 days (IHDS)	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Average LPG use rate (Village/ Urban Block)	0.425***	0.506***	0.538***	0.431***	0.368***	0.270***	0.400***	0.239***	1.006***	1.00***
	(0.031)	(0.009)	(0.013)	(0.009)	(0.011)	(0.011)	(0.011)	(0.011)	(0.017)	(0.094)
Average LPG use rate * 2011 indicator									-0.057**	0.058
									(0.030)	(0.134)
Whether bordering an urban centre?	-0.038	-0.006	-0.002	-0.023***	-0.138	-0.054	-0.045	-0.169		
	(110.147)	(0.024)	(0.002)	(0.006)	(0.100)	(0.053)	(0.074)	(0.146)		
Whether HH has access to electricity?	0.018***	0.053***	0.031***	0.088***	0.029***	0.076***	0.022***	0.097***	0.028**	0.171*
	(0.0014)	(0.004)	(0.002)	(0.006)	(0.003)	(0.007)	(0.004)	(0.010)	(0.014)	(0.103)
Whether HH purchased a cookstove in last 30/365 days?	0.007	-0.058***	-0.034***	-0.135***	0.003	-0.046	0.001	0.016		
	(0.007)	(0.007)	(0.005)	(0.007)	(0.027)	(0.036)	(0.020)	(0.029)		
Household size	0.001***	0.028***	0.006***	0.039***	0.009***	0.030***	0.015***	0.027***	-0.005*	-0.010
	(0.0002)	(0.0009)	(0.0004)	(0.001)	(0.0005)	(0.001)	(0.0007)	(0.001)	(0.003)	(0.011)
Age of head of household	0.0001***	0.004***	0.0004***	0.002***	0.0004***	0.001***	0.0008***	0.001***		
	(0.00004)	(0.0002)	(0.00008)	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.0002)		
Whether head of HH is female	0.0007	0.014***	0.009***	0.040***	0.005	0.029***	0.005	0.026***		
	(0.001)	(0.006)	(0.003)	(0.006)	(0.003)	(0.006)	(0.005)	(0.006)		
Whether head of HH is educated	0.013***	0.088***	0.057***	0.138***	0.066***	0.121***	0.093***	0.116***		
	(0.001)	(0.004)	(0.002)	(0.005)	(0.003)	(0.005)	(0.004)	(0.006)		
Price of LPG	-0.0002	0.00001	-0.001**	-0.0126***	-0.0002***	-0.002**	-0.00003	-0.007***		
	(0.002)	(0.0005)	(0.0007)	(0.002)	(0.00008)	(0.0009)	(0.002)	(0.001)		
Price of kerosene	0.002*	0.012***	0.00003	-0.0004	-0.000005	-0.00001	0.002	0.003		
	(0.001)	(0.003)	(0.00008)	(0.0005)	(0.00001)	(0.000004)	(0.002)	(0.004)		
Whether HH has access to firewood	-0.050***	-0.194***	-0.162***	-0.248***	-0.372***	-0.413***	-0.440***	-0.435***		
	(0.004)	(0.005)	(0.006)	(0.006)	(0.0072)	(0.007)	(0.008)	(0.008)		
Number of years of education of household head									0.003***	0.006
									(0.001)	(0.007)
Whether household has non-biomass cookstove?									0.061***	0.196***
									(0.012)	(0.050)
Hours of cookstove use (/ Day)									-0.012***	-0.020
									(0.003)	(0.014)
Time spent in collecting firewood (/ Trip)									-0.00002	0.001***
									(0.00008)	(0.0006)
Whether household has vent in kitchen?									0.001	0.080*
									(0.008)	(0.049)
Time trend									0.031**	-0.020
									(0.017)	(0.116)
Observations	65307	39567	61097	41897	62937	35026	39915	27459	13730	3342

Notes: The specifications using NSS data include dummies for MPCEs, and for belonging to districts, religion and social groups. Columns (1) and (2) represent an exception, where religion dummies are not included, as they would prevent the convergence of the model. 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 expenditure on repairs in the 55th round.

estimations.

Additionally, we find that these spillovers vary across rural and urban areas: we find evidence that social spillovers exist for both sets of households, and that they weaken over time for both. However, this weakening is steeper for urban households, which had an advantage in terms of initial access to LPG, while rural households experience a prolonged duration of positive spillover effects.

This paper finds that social spillovers are weaker for households residing in states which have higher LPG adoption rates, in line with the S-shaped diffusion model of technology. States with higher initial rates of adoption experience weaker spillover effects with time, in contrast to states that are at the beginning of the adoption curve. This finding should be relevant to policy-makers, especially if they are looking at policies to target certain states or regions. If these spillover effects are stronger in states with relatively low rates of adoption, it might make sense to target subsidies to certain households in these states, rather than providing them to the entire population.

We also attempt to investigate whether these social spillovers are stronger amongst households belonging to a social network. We find that these effects are stronger for members of a caste association, which confirms the importance of these social institutions in India. This can be very useful information to policy-makers and practitioners, who can benefit from organising informational campaigns for members of particular types of social groups, or targeting subsidies to them.

Given that India has very sharp variation in terms of rural and urban adoption of LPG, we look at the differences in social spillovers between rural and urban areas. The results of this paper suggest the presence of spillovers in urban areas, which are decreasing in magnitude relatively quickly over time, and more persistent rural spillovers. A possible implication of this finding is that targeting "pivotal" households, such as rural households residing in states which have not begun using LPG on a large scale, may lead to quick and wide adoption of LPG. Given the nature of data that we use, we are not able to identify whether these spillovers are purely informational, or whether they may be related to imitation, health externalities, and learning-by-doing involved in following the example of other households (cf. Kremer and Miguel (2007)). Nevertheless, we believe that information provision may hasten the adoption process, especially in rural areas. For instance, policy-makers may learn from the use of randomised control trials to provide information on the benefits of LPG, in rural areas in particular, and compare its effectiveness to that of other policies aimed at addressing barriers to the adoption of LPG, such as lack of affordability, or problems of access.

Finally, this paper does not provide any direct evidence either supporting or refuting the effectiveness of subsidies in encouraging Indian households to adopt LPG. Given that the Indian government has been looking to phase out these subsidies for a while, it remains to be seen whether spillovers would still exist, in the absence of subsidies. However, if social spillovers are a factor in determining a households's choice of cooking fuel, subsidies to certain households in the early phases of the adoption process may actually be beneficial in ensuring that more households switch to the cleaner fuel.

6 Conclusion

Greater adoption of clean cooking fuels like LPG by the Indian population is vital for achieving a sustained reduction in indoor air pollution, and thus ensuring the consequent improvement of respiratory health. Lower greenhouse gas emissions are also associated with the use of clean cookstoves. This paper analyses whether there are social spillovers in the adoption of LPG in India, and if these exist, how they vary in strength in different parts of the country. In this paper, we use two sources of data from a widely heterogenous population, a repeated cross-section and a panel data, which enable us to provide a broad scope in addressing this research question. We obtain multiple pieces of evidence suggesting that social spillovers may be present in the Indian LPG context. We find differences between rural and urban households in the persistence of these effects over time. We control for several household-level characteristics of LPG adoption which have been shown to be important determinants in the literature, and address potential threats to identification. Our results may have strong implications for policy-makers looking to encourage consumers to switch to cleaner sources of energy in developing countries. We provide evidence suggesting that social learning amongst consumers of energy products may be present in a developing-country context, and could be used as a policy measure by governments looking to hasten the switch to cleaner sources of energy.

References

Bandiera, Oriana and Imran Rasul (2006) "Social networks and technology adoption in northern Mozambique," *The Economic Journal*, Vol. 116, 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*.

Bass, Frank M (1969) "A new product growth for model consumer durables," *Management Science*, Vol. 15, pp. 215–227.

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, pp. 55–66.

Bollinger, Bryan and Kenneth Gillingham (2012) "Peer effects in the diffusion of solar photovoltaic panels," *Marketing Science*, Vol. 31, pp. 900–912.

Bond, Tami C, Ekta Bhardwaj, Rong Dong, Rahil Jogani, Soonkyu Jung, Christoph Roden, David G Streets, and Nina M Trautmann (2007) "Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850–2000," *Global Biogeochemical Cycles*, Vol. 21.

Bouma, Jetske, Erwin Bulte, and Daan van Soest (2008) "Trust and cooperation: Social capital and community resource management," *Journal of Environmental Economics and Management*, Vol. 56, pp. 155–166.

Boy, Erick, Nigel Bruce, Kirk R Smith, and Ruben Hernandez (2000) "Fuel efficiency of an improved wood-burning stove in rural Guatemala: implications for health, environment and development," *Energy for Sustainable Development*, Vol. 4, pp. 23–31.

Carattini, Stefano (2015) "Green consumers and climate policy: Reconciling Ostrom and Nyborg, Howarth and Brekke," *Haute ecole de gestion de Geneve: Cahier de recherche C–15/2/1*.

Case, Anne C and Lawrence F Katz (1991) "The company you keep: The effects of family and neighborhood on disadvantaged youths."

Cheng, Chao-yo and Johannes Urpelainen (2014) "Fuel stacking in India: Changes in the cooking and lighting mix, 1987–2010," *Energy*, Vol. 76, pp. 306–317.

De Groot, Morris H (1970) *Optimal Statistical Decisions*: McGraw Hill.

Desai, Sonalde and Reeve Vanneman (2009) "India Human Development Survey-I (IHDS-I), 2005-06. ICPSR22626-v5," *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*, Vol. 622.

——— (2015) "India Human Development Survey-II (IHDS-II), 2011-12. ICPSR36151-v2," *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*, pp. 07–31.

Duflo, Esther, Michael Greenstone, and Rema Hanna (2008) "Cooking stoves, indoor air pollution and respiratory health in rural Orissa," *Economic and Political Weekly*, 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, 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, pp. 757–774.

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, 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, 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, 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, pp. 80–114.

IHME (2013) "The global burden of disease: generating evidence, guiding policy," Institute for Health Metrics and Evaluation (Seattle).

IISD (2014) "Subsidies to Liquefied Petroleum Gas in India," *Global Subsidies Initiative*.

Knack, Stephen and Philip Keefer (1997) "Does social capital have an economic payoff? A cross-country investigation," *The Quarterly Journal of Economics*, pp. 1251–1288.

Kremer, M and E Miguel (2007) "The Illusion of Sustainability," *The Quarterly Journal of Economics*, Vol. 122(3), pp. 1007–1065.

Kumar, KS Kavi and Brinda Viswanathan (2007) "Changing structure of income indoor air pollution relationship in India," *Energy Policy*, Vol. 35, 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, p. 637.

Manski, Charles F (1993) "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, Vol. 60, pp. 531–542.

——— (2000) "Economic Analysis of Social Interactions," *The Journal of Economic Perspectives*, Vol. 14, pp. 115–136.

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, pp. 10815–10820.

Moffitt, Robert A et al. (2001) "Policy interventions, low-level equilibria, and social interactions," *Social Dynamics*, Vol. 4, pp. 6–17.

Munshi, Kaivan (2004) "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of Development Economics*, Vol. 73, pp. 185–213.

National Sample Survey Office, Ministry of Statistics and Government of India Programme Implementation, "Consumer Expenditure Survey," Accessed 1st June 2015.

OECD/IEA (2006) "World Energy Outlook."

PPAC (2015) "Total Subsidy on PDS Kerosene and Domestic LPG to Customers."

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, pp. 143–153.

Reddy, B Sudhakara (1995) "A multilogit model for fuel shifts in the domestic sector," *Energy*, Vol. 20, pp. 929–936.

Romieu, Isabelle, Horacio Riojas-Rodriguez, Adriana Teresa Marron-Mares, Astrid Schilmann, Rogelio Perez-Padilla, and Omar Masera (2009) "Improved biomass stove intervention in rural Mexico: impact on the respiratory health of women," *American Journal of Respiratory and Critical Care Medicine*, Vol. 180, pp. 649–656.

Singh, Punam and Haripriya Gundimeda (2014) "Life cycle energy analysis (LCEA) of cooking fuel sources used in India households," *Energy and Environmental Engineering*, Vol. 2, pp. 20–30.

Staiger, D and JH Stock (1997) "Instrumental variables regression with weak instruments," *Econometrica*, Vol. 65, pp. 557–586.

Stock, James H and Motohiro Yogo (2005) "Testing for Weak Instruments in Linear IV Regression," *Identification and Inference for Econometric Models*, pp. 80–108.

UNFCCC (2015) "India's Intended Nationally Determined Contribution."

Verbeek, Marno (2008) "Pseudo-panels and repeated cross-sections," in *The Econometrics of Panel Data: A Handbook of the Theory with Applications*: Springer, pp. 369–383.

WHO (2016) "Indoor Air Pollution and Health. Factsheet No. 292. WHO, Geneva."

Young, H Peyton (2009) "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning," *The American Economic Review*, Vol. 99, pp. 1899–1924.

Zhang, Junfeng and Kirk R Smith (2007) "Household air pollution from coal and biomass fuels in China: measurements, health impacts, and interventions," *Environmental Health Perspectives*, pp. 848–855.

Appendix A Figures

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

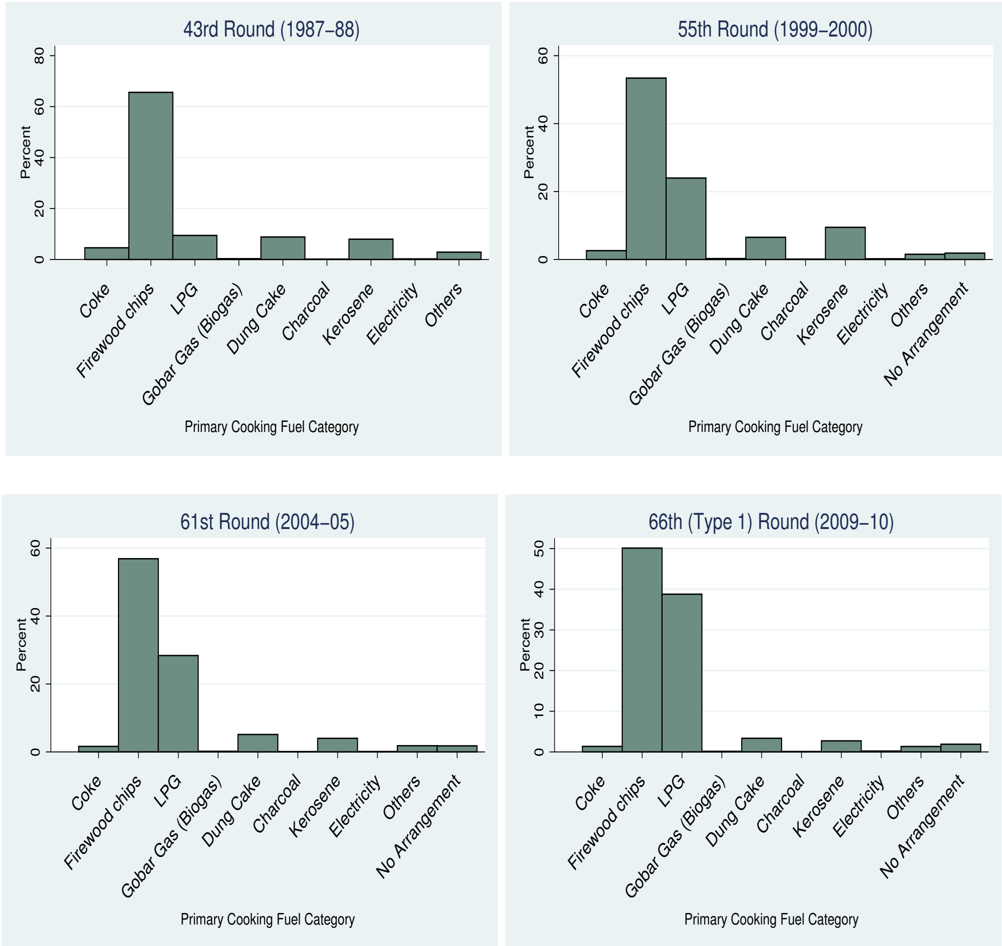


Figure A2: Population Share (%) Using LPG as the Primary Cooking Fuel: 1983 to 2011-12 (Source:NSS)

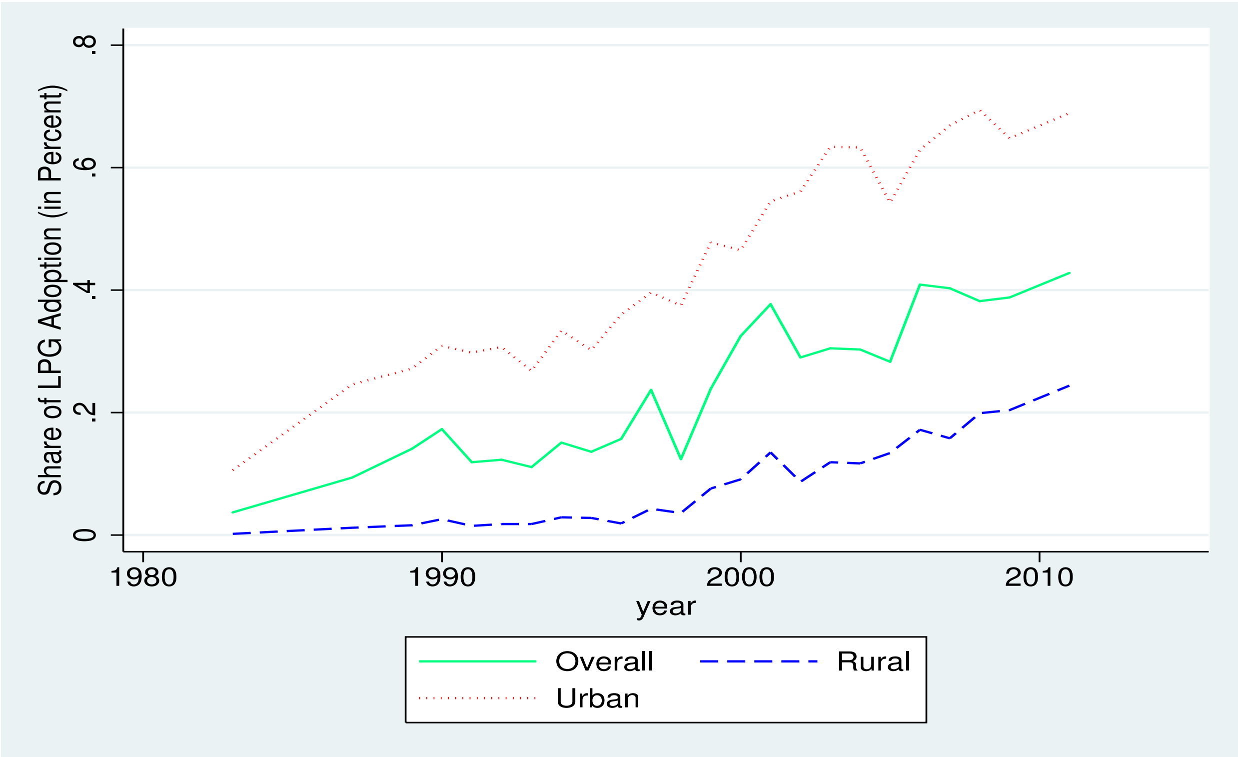
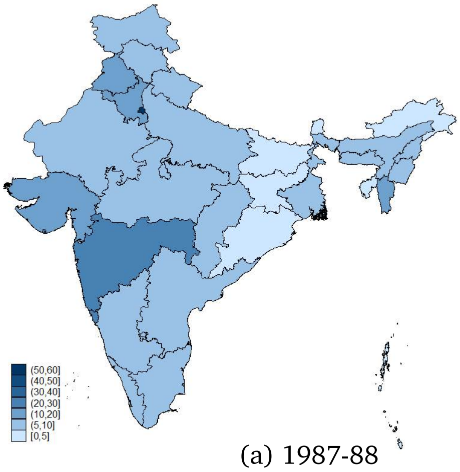
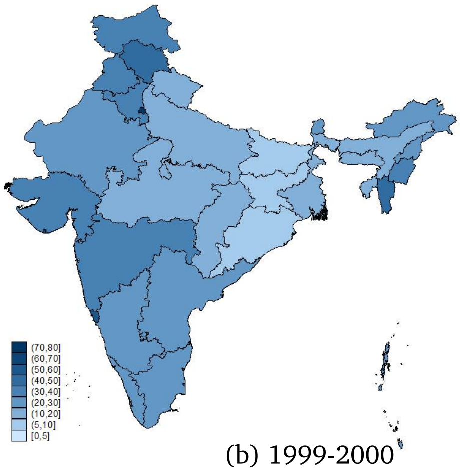


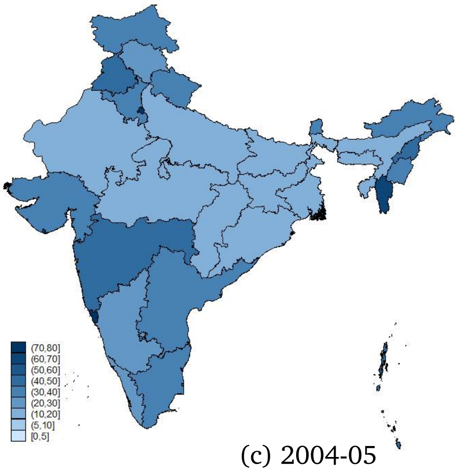
Figure A3: Share of Households Using LPG as the Primary Cooking Fuel (By State) (Source:NSS)



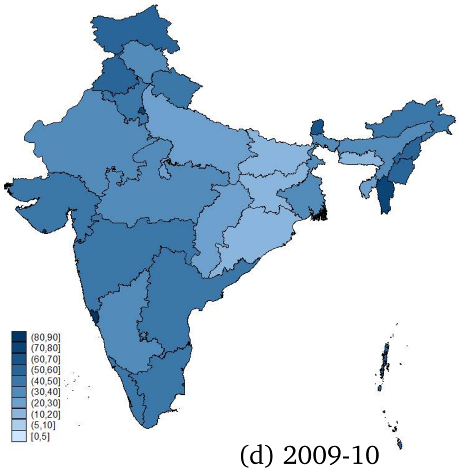
(a) 1987-88



(b) 1999-2000



(c) 2004-05



(d) 2009-10

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

Appendix B Tables

Table B1: NSS Data: Logit Estimations

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)
Average LPG use rate (Village/ Urban Block)	0.170***	0.280***	0.343***	0.365***
	(0.003)	(0.005)	(0.007)	(0.007)
Whether bordering an urban centre?	0.298	-0.298**	-0.247**	-0.092
	(0.195)	(0.157)	(0.131)	(0.146)
Whether HH has access to electricity?	0.441***	0.528***	0.557***	0.498***
	(0.019)	(0.020)	(0.024)	(0.032)
Whether HH lives in a rural area?	-0.572***	-0.159***	-0.080***	-0.053***
	(0.037)	(0.020)	(0.019)	(0.018)
Whether HH purchased a cookstove in last 30/365 days?	-0.010***	-0.051***	-0.0002	0.0002
	(0.002)	(0.003)	(0.0004)	(0.0005)
Household size	1.140***	0.950***	0.710***	0.656***
	(0.034)	(0.024)	(0.021)	(0.021)
Age of head of household	1.327***	0.578***	0.380***	0.293***
	(0.060)	(0.041)	(0.042)	(0.049)
Whether head of HH is female	0.023***	0.027***	0.020***	0.012
	(0.006)	(0.003)	(0.003)	(0.003)
Whether head of HH is educated	0.529***	0.338***	0.357***	0.187***
	(0.018)	(0.009)	(0.010)	(0.006)
Price of LPG	0.014	-0.680***	-0.187**	-0.463***
	(0.040)	(0.134)	(0.092)	(0.121)
Price of kerosene	-0.052	-0.005	-0.001	-0.013
	(0.082)	(0.013)	(0.001)	(0.117)
Whether HH has access to firewood	-1.579***	-1.199***	-1.685***	-1.663***
	(0.037)	(0.023)	(0.025)	(0.026)
Observations	104177	102728	97930	67372
Pseudo R^2	0.6093	0.6046	0.6022	0.6234

Notes: Values reported are marginal effects. All specifications include dummy variables for districts, MPCE deciles, religion and social group (except for the 43rd round, where the religion and social group dummies are are not included, as they would prevent the convergence of the model. 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 Data: First-Stage Estimations

Round	43	55	61	66
Year	1987-88	1999-00	2004-05	2009-10
Corresponding Second-Stage Results (in Table 4)	Column (2)	Column (4)	Column (6)	Column (8)
Dependent Variable: Average Village/ Urban Block LPG Use Rate	(1)	(2)	(3)	(4)
Proportion of population in 10^{th} income decile	0.508***	0.539***	0.429***	0.303***
	(0.028)	(0.019)	(0.004)	(0.019)
Whether bordering an urban centre?	-0.113***	0.023***	-0.009	0.051
	(0.022)	(0.004)	(0.067)	(0.058)
Whether HH has access to electricity?	0.030***	0.046***	0.038***	0.052***
	(0.002)	(0.003)	(0.003)	(0.004)
Whether HH lives in a rural area?	-0.130***	-0.265***	-0.284***	-0.300***
	(0.003)	(0.005)	(0.005)	(0.005)
Whether HH purchased a cookstove in last 30/365 days?	0.00008	-0.027***	0.006	0.037***
	(0.008)	(0.006)	(0.011)	(0.017)
Household size	0.0009***	0.002***	-0.001***	-0.0004
	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Age of head of household	0.0009***	0.0004***	0.0004***	0.0003***
	(0.00006)	(0.00007)	(0.00006)	(0.0001)
Whether head of HH is female	0.008***	0.018***	0.016***	0.025***
	(0.002)	(0.003)	(0.002)	(0.003)
Whether head of HH is educated	0.027***	0.035***	0.025***	0.029***
	(0.001)	(0.002)	(0.002)	(0.003)
Price of LPG	-0.0003	-0.005***	0.006	-0.002**
	(0.0004)	(0.0008)	(0.071)	(0.0009)
Price of kerosene	-0.005***	0.0002	0.006***	-0.002
	(0.001)	(0.0002)	(0.003)	(0.002)
Whether HH has access to firewood	-0.084***	-0.110***	-0.134***	-0.174***
	(0.003)	(0.004)	(0.004)	(0.005)
Observations	104148	102994	97933	67372
5% maximal IV relative bias (Stock and Yogo, 2005)	19.28	19.28	19.28	19.28
10% maximal IV size (Stock and Yogo, 2005)	29.18	29.18	29.18	29.18
Cragg Donald F-Statistic	6654.707	3806.454	2363.815	597.317
P-value	0	0	0	0

Notes: All specifications include MPCE, district, religion and social group 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 B3: IHDS Data: Village FE and Village-by-Year Time Trends

Dependent Variable: Whether HH i spent on LPG in last 30 days Column	Village FE (1)	Village-By-Year Time Trends (2)
Average LPG use rate (Village/ Urban Block)	1.430***	1.434***
	(0.069)	(0.075)
Whether HH has access to electricity?	-0.138***	-0.145***
	(0.039)	(0.041)
Household size	0.008***	0.008***
	(0.002)	(0.002)
Number of years of education of household head	0.011***	0.011***
	(0.001)	(0.0008)
Whether household has non-biomass cookstove?	0.069***	0.067***
	(0.019)	(0.020)
Hours of cookstove use (/ Day)	-0.004	-0.004
	(0.003)	(0.004)
Time spent in collecting firewood (/ Trip)	-0.001***	-0.001***
	(0.0002)	(0.0002)
Whether household has vent in kitchen?	0.077***	0.078***
	(0.010)	(0.010)
Observations	16875	16495
Hansen J-statistic	9.705	8.706
P-value	0.084	0.1214

Notes: Results in column (1) include household-specific time trends. All specifications include income decile dummies ($8^{\text{th}} to 10^{\text{th}}$) and controls for religion and social group. *,** and *** respectively denote significance at 10%, 5% and 1% levels. The coefficient of the constant has not been reported.

Table B4: IHDS Data: Random Effects and Population-Averaged Models

Dependent Variable: Whether HH *i* spent on LPG in last 30 days Column	Random Effects (1)	Population-Averaged Model (2)
Average LPG use rate (Village/ Urban Block)	0.887***	0.865***
	(0.006)	(0.006)
Whether HH has access to electricity?	0.077***	0.071***
	(0.011)	(0.008)
Household size	-0.004***	-0.004***
	(0.001)	(0.0008)
Number of years of education of household head	0.006***	0.006***
	(0.0005)	(0.0004)
Whether household has non-biomass cookstove?	0.112***	0.114***
	(0.007)	(0.005)
Hours of cookstove use (/ Day)	-0.009***	-0.009***
	(0.002)	(0.002)
Time spent in collecting firewood (/ Trip)	0.0002***	0.0002***
	(0.00005)	(0.00004)
Whether household has vent in kitchen?	0.017***	0.017***
	(0.005)	(0.004)
Observations	17072	17072
R^2 (Overall)	0.7207	
Wald χ^2		101652.01
P-Value		0

Notes: Both specifications include household-level individual effects and household-specific time trends. Income decile dummies and controls for religion and social group are included in both models. Standard errors are clustered at the village/urban block level in the random effects model, while robust standard errors are estimated for the population-averaged model. *,** and *** respectively denote significance at 10%, 5% and 1% levels. The coefficient of the constant has not been reported.

Table B5: IHDS Data: First-Stage Estimations

Corresponding Second-Stage Results Column	Table 5 Column (2) (1)	Table B3 Column (1) (2)	Table B3 Column (2) (3)
Income: 5^{th} Decile	1.802***	0.623***	0.628***
	(0.202)	(0.078)	(0.083)
Income: 6^{th} Decile	2.149***	0.816***	0.826***
	(0.201)	(0.071)	(0.077)
Income: 7^{th} Decile	-1.072***	0.737***	0.753***
	(0.298)	(0.063)	(0.066)
Income: 8^{th} Decile	-0.668*	-0.426***	-0.394***
	(0.396)	(0.119)	(0.129)
Income: 9^{th} Decile	-0.830**	-0.317***	-0.292***
	(0.424)	(0.110)	(0.125)
Income:10^{th} Decile	-1.073***	-0.191	-0.153
	(0.391)	(0.119)	(0.127)
Whether HH lives in a rural area?	0.146**		
	(0.070)		
Whether HH has access to electricity?	0.372***	0.457***	0.453***
	(0.031)	(0.030)	(0.028)
Household size	-0.066***	-0.022***	-0.022***
	(0.004)	(0.002)	(0.002)
Number of years of education of household head	0.022***	-0.003***	-0.003***
	(0.002)	(0.0007)	(0.0007)
Whether household has non-biomass cookstove?	0.362***	0.184***	0.187***
	(0.024)	(0.014)	(0.015)
Hours of cookstove use (/ Day)	-0.021***	-0.018***	-0.017***
	(0.008)	(0.005)	(0.005)
Time spent in collecting firewood (/ Trip)	0.002***	0.002***	0.002***
	(0.0003)	(0.0002)	(0.0002)
Whether household has vent in kitchen?	-0.037	-0.056***	-0.062***
	(0.023)	(0.010)	(0.011)
Observations	9350	16875	16495
5% maximal IV relative bias (Stock and Yogo, 2005)	19.28	19.28	19.28
10% maximal IV size (Stock and Yogo, 2005)	29.18	29.18	29.18
Cragg Donald F-Statistic	32.75	36.98	33.53
P-value	0	0	0

Notes: Dependent variable for the models in columns (1), (2) and (3) is the average village/urban block level LPG use rate for all households other than household *i*. Results in column (1) includes household fixed effects, the results in column (2) include village-level fixed effects, and the results in column (3) include village-by-year time trends. Exogenous instruments for the results in all columns are the proportion of population (by village) belonging from the 5^{th} to the 10^{th} income deciles. The specification in column (1) includes income decile dummies (from the 7^{th} to the 10^{th}), whereas the specifications in columns (2) and (3) include income decile dummies for the 8^{th} to the 10^{th} income deciles. Controls for religion and social group are included in all specifications. *,** and *** respectively denote significance at 10%, 5% and 1% levels. The coefficient of the constant has not been reported.