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

Suchita Srinivasan* Stefano Carattini†

August 13, 2016

Abstract

The Indian population is still heavily reliant on solid biomass as 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, and whether these spillovers might explain the spatial disparities that exist in its use by Indian households. 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, i.e. a household is more likely to adopt LPG if other households residing in the same village or urban block do so. We find that this effect is positive for urban households, but insignificant for rural households. We also find that spillovers may be stronger in states that have previously had low rates of adoption, and for households that are members in associations or social networks. Our results have strong implications for policy-makers who can utilise lessons from social learning to encourage consumers to switch to cleaner sources of energy in developing countries.

Keywords: Spatial Disparities; Social Learning ;Clean Energy Adoption; Public Policy.

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

---

*This version: August 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.

†Haute école de gestion de Genève, HES-SO// University of Applied Sciences Western Switzerland 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: s.carattini@lse.ac.uk.

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.[1] 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). IHME (2013) estimates that 2.9 million deaths were caused by ambient air pollution due to PM 2.5 (particulate matter with diameter less than or equal to 2.5 micrometres). Causes of death range from pneumonia, stroke, heart disease, chronic obstructive pulmonary disease (COPD), 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 local (and global) environment. Smoke generated from burning wood contains harmful pollutants such as carbon monoxide and particulate matter. 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, which is often arable (OECD/IEA, 2006). These environmental concerns are particularly pressing in the light of India's commitments in the recently concluded Paris Agreement, where India 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 the energy emissions intensity by 30-35% by 2030 compared to 2005 levels (UNFCCC, 2015).

The 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. [3] Cleaner cooking alternatives,

[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.

[3] .There is also evidence to suggest that providing clean cookstoves to households may not lead to improvements in health quality in the long-term: Hanna et al. (2016) find this from a randomised control trial in India. However, the reasons for decline in health outcomes in this study may be attributed to the lack of maintenance of cookstoves, and their inadequate and inefficient use.

such as liquefied petroleum gas (LPG) have grown in popularity over time, but rather slowly. From the beginning of its use, LPG was supplied to urban areas of the country, whereas the rural areas always lagged behind in terms of adoption. The Indian government, in order to incentivise consumers to switch to LPG, subsidised the cooking fuel considerably, but these subsidies have been mis-targeted to urban areas.[4]These factors have resulted in urban households, who are richer, more educated, and more likely to have access to LPG, using it more than than poorer, rural households.

The literature also abounds on the role of socioeconomic factors in determining which households use clean cooking fuels in developing countries.[5] However, one aspect of this decision which has not been extensively studied is whether households learn from other 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 hypothesise that social learning behaviour (or social spillovers) could also affect a household's choice of cooking fuel, and thus contribute to the wide variations we observe in its adoption.

The objective of this paper is to investigate whether social spillovers exist in the adoption of LPG in India, and if they do, how they vary across rural and urban areas, states, and based on the strength of the interaction. 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 some 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/urban block do so. In addition, we find that these effects are stronger for urban households than for rural households, and 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, suggesting that social networks might play an important role in encouraging households to switch to clean energy sources.

---

[4] .Despite significant reductions in subsidies over the years, as of 2015-16, there is still an LPG subsidy of Rs.150.82 per litre per cylinder (PPAC, 2199)

[5] .Lewis and Pattanayak (2012) provide a summary of this literature, and the types of factors that have been studied as important determinants of the adoption of clean cooking fuels and improved cookstoves (ICS)

This paper uses sample survey datasets on household-level consumer expenditure, which are nationally representative and large-scale. The biggest benefits are that the large scale of the dataset 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 into consideration in conjunction, may suggest the presence of peer effects in the adoption of LPG. For the NSS data, we use the non-linear (probit) methodology to estimate the dependence of a household's decision to use LPG as the primary cooking fuel on the corresponding decisions of other households in the same village, or urban block, while controlling for socio-economic factors that the literature suggests influence a household's fuel choice. In addition, we also employ the instrumental variable two-stage least squares (IV-2SLS) methodology to control for endogeneity in this estimation. For the IHDS data, we use fixed effects estimations, again using the IV-2SLS methodology. Our results are robust across specifications.

The main contribution of this paper lies in its investigation of social spillovers among users of a clean fuel in a developing country, where many households have access to cheaper alternatives. While sources of fuel such as firewood are available freely, and at significantly lower costs, we provide evidence that households may still be influenced to purchase clean energy alternatives if others households residing in the same village or urban block do so. Our objective is to suggest the possibility that these social interactions might explain the spatial disparities that are observed in the adoption of LPG in India. These findings are relevant to policy-makers, given that in many countries such as India, clean fuels are subsidised. It may be beneficial for policy-makers to target certain "influential" households in rural areas, for example, if there is evidence that rural households do indeed learn from the experiences of other rural households.

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, while 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 as sources of cooking fuel. For cooking, rural households have relied heavily on firewood use over time, and continue to do so, whereas there is sufficient evidence of an energy transition for urban households. 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% for 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.

In this paper, we choose to restrict our attention to adoption of LPG as the clean cooking fuel alternative. This is because it is the most widely used clean cooking fuel amongst Indian households, and it is also the most affordable (despite still being expensive for many consumers, especially rural ones). LPG is currently being used by most urban households, and increasingly by many rural ones. Income and awareness are obvious determinants of the choice of a household to consume cleaner fuels such as LPG. However, the shift to cleaner fuels does not follow the energy-ladder model, where households switch to cleaner cooking fuels in a linear way as the level of income increases. [6]

Acquiring an LPG connection requires an initial expenditure to purchase the stove, and install the equipment. Additionally, households are required to purchase cylinders of the gas as and

[6] . 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)

when required. Rural households are often not able to afford the recurrent expenditures required to purchase 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 their access to poor or homeless people, even in urban areas (Gupta and Köhlin, 2006). LPG is marketed by state-owned petroleum distribution companies (Indian Oil Corporation (IOC), Bharat Petroleum Corporation Limited (BPCL) and Hindustan Petroleum Corporation Limited (HPCL)) 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 is 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 (IISD, 2014). 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 (IISD, 2014). There are also disparities in the distribution of 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; there are also wide disparities in the distribution of subsidies. The same five states, for instance, receive almost 50% of the subsidies as well, and even within these states, it is the urban areas that benefit more than rural areas (IISD, 2014).

Recent reforms have been undertaken by governments to improve the accessibility of LPG to Indian consumers, both rural and urban. Other than the spread of the LPG distribution network, which has expanded considerably in rural areas, efforts have also been made to reduce the distorting subsidies on LPG. For instance, in September 2012, the central government capped the number of subsidised cylinders that a household can acquire at six per year, in order to reduce the expenditure on LPG. However, they backtracked within a few months and on January 2013, the limit was increased to nine per household per annum (which was further increased to 12 by 2014). In 2013, the central government also initiated the Direct Benefit Transfer (DBT) scheme with the intention of developing an electronic payment system to ensure that subsidies were directly transferred to the bank accounts of consumers.

However, governments have found it politically infeasible to initiate a phase-out of the subsidies, even though efforts are being made to divert more resources towards poorer, rural households. In March 2015, the central government initiated a policy encouraging rich, urban consumers of LPG to voluntarily renounce their subsidies, following which almost two million households surrendered their rights to receive subsidies on LPG cylinders. Such measures were taken in order to ameliorate the disparities that currently exist in LPG access, and try to ensure greater equity in its access.

2.2 Literature Review

There is ample literature which has looked 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 cleaner fuels, and improper maintenance of ICS is prevalent in developing countries, and this limits the health benefits that can be reaped from their use. Mobarak et al. (2012), for instance, find from surveys in Bangladesh that households' willingness-to-pay for improved cookstoves is low, and that their perceived level of risk of ill-health from burning solid biomass is also low.[7]This leads households to only use these stoves if they are provided for free, and limits their regular use. Duflo et al. (2008) provide evidence from literature on the the importance of maintenance in leading to improved health outcomes.

A recent paper (Hanna et al., 2016) used experimental data from India to find that clean cookstoves led 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. In this paper, we hypothesise that a household's decision to use LPG as the primary cooking fuel may depend on actual use of the fuel by other households: by considering the determinants of the use decision of a household, we are, in effect, analysing factors which may play a role in determining health outcomes of households.

[7] .This is an empirical fact which has been found to be true for many environmental quality improvements: Greenstone and Jack (2015) build a framework to understand why agents have low willingness to pay for environmental improvements in developing countries, and find that it can be driven by low income, steep marginal costs of environmental improvements, inefficient policy-making and market failures

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 cooing 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 commonly used determinants of the choice to adopt clean cooking fuels in the literature, even in the Indian context (Reddy (1995), Kumar and Viswanathan (2007), Farsi et al. (2007))[8]. However, access to cleaner cooking fuels, also plays an important role in the Indian context. Gupta and Köhlin (2006) find that there are several factors, other than the prices of fuels themselves, which may determine the fuel choice of a household-access to cleaner cooking fuel perhaps being one of the most important. This finding is pertinent from a policy perspective, given that most cooking fuels such as kerosene and LPG are subsidised by the Indian government.

Another aspect of the transition to clean cooking fuels which has been studied in the Indian context is the behaviour of fuel stacking: Cheng and Urpelainen (2014), for instance, find that from 1987 to 2010, many Indian rural households began using LPG, but continued to use firewood as well. They find that most socio-economic characteristics that explained the adoption of LPG in 1987 are no longer significant determinants of the behaviour of households to switch from firewood to LPG in 2010, when LPG use becomes widespread: only income and the urban nature of a households continue to explain the behaviour of households. Thus, it appears that households have not followed the "energy-ladder" model, where consumers switch to cleaner cooking fuels monotonically, as their levels of income increase. In this paper, we control for factors which have previously been found to be important such as income, and investigate whether social spillovers may also explain a household's decision to adopt LPG.

Thus, socio-economic determinants need not be the only factor influencing household adoption decisions of clean fuels, or even clean technologies. There is also some literature on the role of

[8] .Reddy (1995) uses the "energy ladder" concept to explain the transition of a household from dirty to cleaner cooking fuels in the city of Bangalore in India, given the importance that income plays in determining cooking fuel choice. Kumar and Viswanathan (2007) find that there are non-monotonicities in the adoption of clean cooking fuels in India, with the use of cleaner fuels increasing with income in urban areas (where income levels are on average are higher than in rural areas), but the use of dirty fuels (such as coke and coal) increasing with income for rural households (this is confirmed by Rao and Reddy (2007), who also find income (and household size) affect the probability of choosing a fuel non-linearly). Farsi et al. (2007) also find that low levels of income are one of the significant barriers towards greater adoption of cleaner cooking fuels in the Indian context (not only are the fuels more expensive than alternatives such as firewood, but so is the initial equipment).

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. Literature on developed economics has looked at the role of spillovers, or "peer-effects" in explaining the adoption of solar panels, for instance. Bollinger and Gillingham (2012) study the presence of peer effects in the diffusion of solar PV panels in California, and find that one addition to the installed base of the panels in a zip code is likely to increase the probability of adoption by households in the same zip code by 0.78%. They 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, but in Connecticut, and find that the number of previously installed systems have a role in determining the adoption decisions of households. They use a rich set of controls related to the built environment, and socio-demographic factors. They also find that these peer effects decrease over time, and that factors such as housing density actually decrease the adoption of PV panels.

In this paper, we use pan-Indian data to study whether such social spillovers in LPG use exist for households that reside in the same geographical area (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 can expect these spillovers to be more pronounced for households with memberships in these social networks.[9]Carattini (2015) provides a summary of the literature on the role of social norms and trust in the evolution of beliefs amongst consumers about the effectiveness of different environmental policy instruments, when cooperation is needed to achieve goals of climate change mitigation, for instance.

The literature on the role of social spillovers in developing countries is dominated by agricul-

[9] .There is a broad literature on whether membership in organisations could play a role in strengthening social institutions, or not. Putnam et al. (1994) and Knack and Keefer (1997) provide the two contrasting views on this issue.

ture.[10]There is scant literature from developing countries on social spillovers in the decision to adopt clean energy technologies. Beltramo et al. (2015), for instance, study the adoption of efficient cookstoves using data from a randomised control trial in Uganda; their results are sobering. They find no evidence to suggest that households located near early buyers of clean cookstoves are more likely to purchase it themselves. They also do not find any evidence to suggest that peer-effects are stronger for the geographically and socially close households. In this paper, we do not use experimental or quasi-experimental data, which is used to prove the existence of "peer-effects" in the literature. Thus, our approach is different: we use multiple datasets to provide cross-country 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. Our results, while not proving causality, are useful to policy-makers interested in expediting the adoption of clean energy by households in developing countries.

3 Theoretical Framework

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 the model developed in that paper, 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.

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

[10] .Bandiera and Rasul (2006), for instance, study the adoption decisions of farmers in Mozambique, and they find that there are non-linearities in the dependence of a farmer's crop choice decision, and those of his family and friends: farmers are more likely to adopt a new crop if a few farmers in their network adopt, but they are not that likely to adopt if too many farmers in their network adopt. Munshi (2004) finds that social learning amongst farmers for high-yield variety technology for rice in India was much weaker than it was for wheat, given that there is much more heterogeneity amongst rice growing conditions in India, compared to wheat growing conditions.

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 doesn't. 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 it's "rigidity" of beliefs, such that low values of τ_i indicate that household i is very amenable to changing its beliefs.

The essence of the model which will be useful for the framework developed in this paper is how much "social interaction" household i has with other households (denote as β_i) . 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. However, unlike in that model where it was a function of a continuous variable denoting cumulative information generated up to time t, we modify the expression for this variable 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 on 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 enough.

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 certain propositions, which we will test in the empirical section of the paper.

3.2 Derivation of Expression for Expected Benefit of Adoption

While the objective of Young (2009)'s model was to derive the dynamics of the diffusion process of a technology in the presence of informational spillovers, our objective is to focus on the spillovers themselves, and to pinpoint how these spillovers may vary across different categories of households. We denote these spillovers as "social spillovers", to remain agnostic about their nature. Since the decision to adopt the technology (or LPG) is driven by the expected benefits from its use, and using the assumption of myopic consumers. 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

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

In order to derive the propositions in the next subsection, we 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 Propositions

Given the higher population density in urban areas compared to rural, we assume that for all households j residing in urban areas and k residing in rural areas, β_j is larger than β_k, i.e. in urban areas, the extent of social interactions are higher than in rural areas. Define the share of adoption up to time t, λ_t, as

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

Proposition 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, 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.

Proposition 2: The magnitude of the spillovers are higher in areas with higher population density, i.e. in urban areas compared to rural areas.

Proof: The expression for $\frac{\partial E[B_{it}]}{\partial Y^{\mathrm{A}}{}_t}$ is

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

Using this derivative, it is straightforward to show that the mixed partial derivative $\frac{\partial E[B_{it}]}{\partial Y^{\mathrm{A}}{}_t \beta_i} \geq 0$, i.e. that the magnitude of the spillover is higher when household i has more social interactions (given the regularity condition on τ_i which we assumed in the previous subsection). This implies that for households residing in urban areas, where due to greater population density, β_i can be expected to higher, the social spillover will be stronger than for households in rural areas, ceteris paribus.

Corollary 1: Informational spillovers are stronger amongst households interacting in tighter social networks.

Proof: This follows from the proof of Proposition 2, i.e. the social spillover is stronger for households that have higher β_i. This would imply that households in a closer-knit network (where they have the possibility of interacting with more households) will experience stronger spillovers.

Proposition 3: 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 $\frac{\partial E[B_{it}]}{\partial {Y^{\mathrm{A}}}_t \lambda_t} \leq 0$, i.e. the magnitude of the spillover deteriorates 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 provides evidence in support of the S-shaped curve commonly used to explain the diffusion of new technologies.

4 Data and Methodology

4.1 Data

The objective of the empirical estimation in this paper is to provide multiple strands of evidence on the role of social spillovers in incentivising Indian consumers to adopt LPG, and to prove the hypotheses that we derived in the previous section. We use two sets of data for the empirical analysis. The first type of data employed is that of the National Sample Survey (NSS) of India, which is published by the National Sample Survey Organisation (NSSO), a subdivision of the Ministry of Statistics and 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 giving it a repeated cross-sectional nature. Each sample frame is 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 conducted "thick" rounds of the NSS at a frequency of approximately every five years, whereas in the interim, "thin" rounds are conducted, where a smaller sample of households is surveyed. The thick rounds that are included in the data used for the analysis in this paper is from 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 the NSS data because of the large sample sizes in these rounds relative to those of the thin rounds; this is particularly useful when using survey data, where it is imperative to correct for the possibility of errors being correlated across geographical units (such as villages or urban blocks), by clustering the standard errors. The sampling of over 100000 households in each of these rounds provides ample geographical heterogeneity in the data, which is an asset, given the focus of the paper on spatial disparities. [11]

The NSS data does not provide details on the exact address or locations (names of villages, or urban blocks, to be precise) of the households , due to data privacy concerns. However, we are able to obtain the names of the district (and state) of each household's residence. In addition, villages and urban blocks are uniquely numbered, to enable us to identify households that may be "neighbours". This is the finest level of geographical identification possible, given the data.

The second database used for the analysis in this paper is the IHDS (India Human Development Survey) data compiled by the University if Maryland and the National Council of Applied Economic Research (Desai and Vanneman (2009), Desai and Vanneman (2015)). It is a panel dataset, with 83% of the households sampled in the first round (2005-06) being sampled again in 2011-12. While this survey is conducted on a smaller scale than the NSS, both in terms of the number of households and regions sampled, the panel nature of the data enables us to track changes in adoption over time.

In both datasets, households are asked detailed questions about their expenditure on items over a "reference period", which is the period of time over which the household is asked to provide information about expenditure. The reference period often varies across items. For fuel-related expenditures, most of the rounds ask the households for expenditure over the previous 30

[11] .The other alternative would be to create a "pseudo-panel" with the repeated cross-sections of data available, based on some characteristics of the households that do not change over time (such as year of birth, location, gender, etc.). There are two main reasons why we choose not to adopt this approach. Firstly, the clumping of observations together in creating the panel would make it very difficult to study peer-effects in the adoption of LPG. The second reason for not creating a pseudo panel is more practical: in order for estimates to be consistent, it is required that each pseudo panel have a minimum of 100 observations per unit of time (Verbeek, 2008), which may be problematic in the thin rounds.

days.[12]Households are also asked about expenditure on durable goods (such as cookstoves) in the last 365 days. Along with the information on expenditure, both datasets compile information on the demographic characteristics of all the members of the households surveyed, such as the age, gender, marital status, industry of occupation, and level of education. Information is also provided on land ownership (total land possessed, 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.)

In this paper, we use data 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 from the NSS data. 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 (Cheng and Urpelainen, 2014). In the IHDS data, we restrict the sample to those 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.

Figure 1 in the appendix shows the distribution of households by primary fuel-type (for cooking purposes) in the four thick rounds of the NSS which this paper uses for analysis. As is clear from these graphs, firewood was, and still is, the primary cooking fuel for a majority of the households. It is also clear that 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, and it is one of the chief cooking fuels adopted in primarily urban areas, whereas dung cakes have gained popularity in rural areas.

Figure 2, also in the appendix, plots the evolution of the proportion of households for whom LPG was the main cooking fuel over all the years in the sample, and shows that it has gained popularity amongst households, especially in recent rounds of the survey. The importance of urban areas in contributing to this trend is apparent by looking at the increase in the share of its adoption in urban areas: while the share of LPG adoption amongst rural households has also increased over

[12] .The 66th 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 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 66th round used 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. The first type of data is utilised for this analysis in this paper, where the recall period is 30 days for items of food and fuel, to enable better comparison with results from other thick rounds.

the period, the pace of increase has been much slower than that of urban households, suggesting that rural households have perhaps not been able to "catch-up" to their urban counterparts.

Figure 3 shows the regions which have contributed to the increase in popularity of LPG over time. The maps show the proportion of households in each state for whom LPG was the primary cooking fuel in the four periods of the thick rounds. In 1987, the highest proportion of LPG users were in Delhi and the" union territories" of Goa, Chandigarh and Daman and Diu, which are primarily urban areas. Over time, some of the bigger states (Maharashtra, Tamil Nadu and Karnataka) experience an increase in the share of LPG adopters.

Table 1 and Table 2 provide summary statistics on some of the demographic characteristics of the households, using NSS and IHDS data respectively. Table 1 suggests that while the households have not changed considerably in terms of size, or the characteristics of the head of the household, it appears that many households have begun using LPG as the primary cooking fuel, have acquired access to electricity, and have incurred higher monthly per capita expenditures as time has passed. These observations can be confirmed from Table 2 as well. By 2011-12, most households in the sample have access to electricity, and spend on LPG. However, it does not appear that the proportion of households using non-biomass cookstoves is increasing, suggesting that households might spend on LPG, but continue using other fuels as the primary cooking fuels, which is an indication of fuel-stacking.

The measure of LPG adoption that we use from the NSS data is a binary variable for whether LPG is the primary cooking fuel of a household or not; while the data allows us to create a variable including other clean fuel options such as biogas and dung cakes, we prefer to focus on the differences in the adoption of LPG in this paper, and the reasons for the same. The IHDS data does not have a corresponding variable, thus our measure of adoption for this data becomes whether the household spent on LPG in the last 30 days.

4.2 Methodology

In order to investigate the presence of spillovers in the adoption of LPG, we adopt a multi-pronged approach by providing evidence from both cross-sectional and panel datasets to test the hypotheses developed in the theoretical framework. The objective of this section is to highlight the econometric models that we estimate, and to justify their choice based on both the existing literature, and

Table 1: 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 At Least Has 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 Spend on 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 had Access to Firewood (%)	73	44	128,039	60.84	48.81	120,127	68	46	124,637	63	48	100,845

Table 2: 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 Spends on LPG (%)	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

Table 3: Sample Size and LPG Adoption Rate by Village/Urban Block

Statistic	NSS				IHDS
Round	43rd	55th	61st	66th (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.

on facts specific to the Indian cooking fuel-use context.

As previously mentioned, our primary model tests for the presence of informational spillovers across households located in the same geographical region. India comprises 29 states and seven union territories (UTs), and each state/UT is further divided into districts, which are further divided into villages (rural areas) or blocks (urban areas). It is possible to ascertain the names of the state and districts of residence, but neither NSS nor IHDS data provide the names of the villages or urban blocks. It is, however, possible to identify households that live in the same village or urban block, in both datasets. Thus, the geographical entity we use to study for the presence of spillovers is the village/urban block.

Table 3 provides information on the mean, maximum and minimum number of households by village or urban block in both datasets, and the mean LPG adoption rate (at the village/urban block level). While most papers studying the presence of peer-effects focus on the entire population belonging either to a certain geographical region, or a network, in this paper we try to analyse whether informational spillovers may explain the adoption of LPG across India, and how these effects vary across regions in the country. For this reason, we choose to use survey data and investigate the presence of these geographical spillovers at a national level. As Table 3 reveals, the IHDS data samples more households per village than the NSS, but given the geographical breadth of the NSS sample frame, we use both sets of data for the analysis in the paper.

Using data from the four thick rounds of the NSS, we first estimate a linear probability model (LPM) [13] of the form

$$A_i = \alpha_0 + \alpha_1 A_{-ij} + \alpha_2 E_i + \alpha_3 U_i + \alpha_4 X_i + \mu_i \tag{10}$$

for each round, where the dependent variable is denoted by A_i, a binary variable indicating whether LPG is the primary cooking fuel adopted by 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.

In this basic specification, we do not control for access to LPG (or presence of a local suppler of LPG, for instance), due to lack of availability of this information. As proxies for access to LPG, we use a variable to indicate whether household i spent on electricity in the last 30 days (E_i), and another variable to indicate whether household i resides in a district which is adjoining one or more urban centres (S_i).[14] 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 had access to firewood, monthly per capita expenditure (MPCE) dummies, and prices of LPG and kerosene.[15] μ_i denotes the stochastic error term. Standard errors are clustered at the village/urban block level.

It is immediately clear that there are endogeneity concerns in carrying out this estimation us-

[13] .The results using the non-linear logit model are included in Table 1 in the appendix.

[14] .While electricity is not required for using a cookstove with an LPG cylinder, we find that this variable is very highly correlated with the use of LPG as a primary cooking fuel for the households in the sample. The pairwise correlation coefficient is found to be 33.35%, 35.74%, 33.86% and 31.92% respectively in the four rounds.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)

[15] .The NSS data does not include 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 would only be populated 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 take the average price in the district as a measure of price for the households that did not actually purchase LPG in a given year (following the procedure adopted by Kumar and Viswanathan (2007)).

ing OLS, which has been extensively detailed in the social learning literature as the "reflection" problem (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 (termed "simultaneity"). Moreover, both a household's choice, and that of its neighbours, may be influenced by a set of common (unobservable) factors, such as whether there is an LPG retailer located in that village/urban block, the subsidies provided by the retailers, etc. Lastly, there may be endogenous group formation (or self-selection of peers).

The identification problems associated with peer effects have been extensively studied in the literature. Papers have adopted different empirical approaches to deal with these issues, from the use of fixed effects in the availability of panel data, to deal with the first two problems, to the use of instruments to deal with the reflection problem (Bollinger and Gillingham, 2012). In the estimations carried out using NSS data, given that we have repeated cross-sections, we try to mitigate the concerns about correlated unobservables at the district level, by including district-level dummies. In order to correct for simultaneity, we estimate an instrumental variable probit model.

We follow the approach adopted by Duflo and Saez (2002) and Case and Katz (1991): Duflo and Saez (2002) attempt to study whether there are peer effects in participation in retirement plans amongst colleagues, and find that the choice of employees at a university to enrol in a retirement plan, and the choice of vendor, were influenced by the decisions made by other employees in the same department. 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 MPCE decile as an instrument for average LPG adoption at the village-level. [16]

By using the proportion of population in each village or urban block belonging to the highest MPCE decile as an instrument for the average village/urban-block level LPG adoption rate, we try to handle the weak-instrument problem: 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

[16] .This methodology was detailed by Case and Katz (1991), who studied the effect of neighbourhood peers on the behaviour of youths in poor neighbourhoods in Boston. They find significant neighbourhood effects for many kinds of criminal activity among young persons.

highest MPCE decile. The model that is estimated is thus the same as above, except A_{-ij} is treated as an endogenous variable, and this is done for each of the four thick rounds of the NSS data.

However, it is not certain that the use of repeated cross-sectional data can mitigate that endogeneity concerns that were mentioned above, especially those related to correlated unobservables. The results obtained using NSS data are still illustrative, given the geographical scope of the database, and the large sample that it uses. In order to substantiate these results though, we choose to use the panel IHDS data. The linear probability model that we estimate is:

$$A_{it} = \alpha_1 A_{-ijt} + \alpha_2 X_{it} + \alpha_3 \delta_i + \alpha_4 \epsilon_t + \mu_{it} \tag{11}$$

where, as before, A_i is a binary variable indicating whether LPG is the primary cooking fuel adopted by household (i) (in time period t), and the main independent variable 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, hours a day that the cook-stove is used, and whether the household has a kitchen with a vent.

In our baseline specification, we use household-level fixed effects, and a time trend variable. This is adequate for controlling for time-invariant unobservables which could affect households: for instance, access to LPG could be controlled for, at least for the population that either had access in both years, or didn't. However, this may still be insufficient to capture the effect of time-varying unobservables such as improved supply of LPG to the village/urban block from 2005 to 2011. In order to be able to better capture these effects, we also estimate models using village-level fixed effects and village-by-year time trends.[17]

Since the endogeneity concerns still remain, we also estimate the IV-2SLS model using household-level fixed effects and a time trend. In choosing instruments, we adopt the same methodology as with the NSS estimations, using the proportion of population in the same village or urban block belonging to the highest income deciles (following Duflo and Saez (2002) and Case and Katz

[17] .These results are included in Table 3 in the appendix, along with the results using random effects, and the population-averaged model in Table 4

(1991)).

In order to test Corollary 1, we also estimate model (11) above, but only for households that belong 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. We do not use the IV-2SLS methodology for this estimation, because of a reduction in sample size. Similarly, to test Hypotheses 2, we estimate the model (11) for rural and urban sub-populations.

Finally, Hypothesis 3 of the theoretical framework suggested that spillovers, if they exist, 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 referring the the 55th round of the NSS data (from 2004-05), and categorising the states on the basis of their adoption rates in that time period (less than 20%, between 20-30 %, between 30-40% and more than 40%). We then use the IHDS panel data and create binary variables for these four categories, and interact these dummies with the village/urban block level LPG adoption rate to analyse how the spillover effect varies across states. We estimate a linear probability model with fixed effects to study this.

5 Results and Policy Implications

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 data from the four thick rounds of the survey. This table includes the results of the linear probability models and the instrumental variable probit models, whereas the results of the logit model are included in Table 1 of the appendix, and the first-stage results are included in Table 2 of the appendix.

The results of Table 4 suggest the possibility of spillovers, which appear to be larger in the 43rd round, but deteriorate in magnitude by the 66th round (this is true of both types of models). These effects are positive, and significant at the 1% level (except in the IVM of the 66th round, where this variable is insignificant). The magnitude of the coefficient is higher in the IVM than in the LPM, which suggests the possibility that the linear estimates are underestimating the strength of this effect. For instance, 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 0.62 units in the LPM, and by 1.75 units according to the IVM. For the IVM results, the Cragg-Donald F-Statistics (which are used to test for weak identification of the instrument) are consistently high, and surpass the bounds proposed by Stock and Yogo (2005) to identify weak instruments (first-stage results are provided in Table 2 in the appendix).

These results show that households which have access to electricity are more likely to adopt LPG, as our initial hypothesis suggested. The other variable which we use to try and control for supply using the NSS data, whether a household lives in a district adjoining a big urban centre, is insignificant in many specifications. Households with older heads of household, female heads, and more educated heads 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.

The results also seem to suggest that larger households are more likely to adopt LPG; the existing literature has been inconclusive on the nature of the relationship between the size of the household, and the decision of a household to adopt clean cooking fuels. Lewis and Pattanayak (2012) found in their summary of studies related to adoption of improved cookstoves and fuel choices by households in developing countries that 89% of the studies considered household size as a potential determinant, but find that they were inconclusive on the nature of the relationship between the two factors. On the other hand, they find that the relationship with income is quite robust across studies: richer households are more likely to adopt cleaner fuels. In this model, we control for income using dummies for monthly per capita expenditure deciles, which are significant in every round. This corroborates the findings of the literature 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. It includes the results of the LPM in column 1, the second-stage IV-2SLS results in the second column, and the results which are relevant for testing Propositions 2 and 3 of the theoretical framework in columns 3 and 4 respectively. The first-stage results of the IV-2SLS model are included in Table 5 of the appendix, the results obtained using village/urban block level fixed effects and village-by-year time trends are included in Table 3, and the results of the random effects and population-averaged models are included in Table 4 of the appendix.

The results of columns 1 and 2 in Table 5 indicate that the variable for average LPG adoption

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 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 had 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 had 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-squared	0.4729		0.5775		0.6004		0.6381	
Wald test of endogeneity (Chi Square)		0.21		31.64		10.47		21.28
P-value		0.6438		0		0.0012		0

Notes: Values reported are marginal effects. Proportion of population in the same village or urban block in the highest income decile is used as an instrument. All specifications include dummies for MPCE, districts, religion and social group. The dummy for monthly per capita expenditure of the 10th decile has been omitted. Standard errors are clustered at the village/urban block level (reported in parentheses). *,** and *** respectively denote significance at 10%, 5% and 1% levels. Coefficient of constant has not been reproduced. The variable "Whether HH purchased a cookstove in last 30/365 days" includes expenditure on repairs in the 55th round. The IVM in the 61st round does not include controls for religion.

at the village/urban-block level has a positive coefficient, and it is significant at the 1% level of significance, providing support for Proposition 1 derived in the theoretical framework. 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 village LPG adoption rate leads to a one unit increase in the probability of household i spending on LPG (LPM), whereas it leads to a 1.16 unit increase in the probability of household i spending on LPG according to the IV-2SLS model (again, 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). Households that use a non-biomass cookstove, have a vent in the kitchen, and that use the stove for fewer hours per day are also more likely to spend on LPG. The results also suggest that larger and richer households are more likely to spend on LPG, as are households with more educated heads (these results corroborate with those obtained using the NSS data).

The results presented in column 3 of Table 5 are helpful in ascertaining whether households in urban areas experience stronger spillover effects than households in rural areas. The model is estimated using IV-2SLS, and the results suggest that there are positive spillover effects in urban areas (significant at the 1% level), whereas these effects are absent in the rural areas (the interaction term between the average adoption rate and the rural indicator is insignificant). This supports our hunch that households residing in more densely-populated areas experience stronger spillover effects.

Column 4 of Table 5 presents the results of the estimation by including indicator variables for the levels of LPG adoption before the start of the time period of the panel data(which is calculated using the 55th round of the NSS data). Due to the complexity created by using multiple endogenous variable in an IV-2SLS estimation, a LPM is estimated. The results suggest that the spillover effects are significant at the 1% level, and of the highest magnitude, for those states which start out 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 average adoption rate variable is insignificant in the states that started out with higher rates of adoption, namely those where the adoption rate was between 30-40%, or greater than 40%. This suggests the possibility that the

LPG use in India follows the S-shaped diffusion curve found to be relevant for many technologies.

Table 6 below includes the results of the estimation of the LPM on smaller subsamples, namely amongst households belonging to certain types of social networks. According to Corollary 1 derived in the theoretical framework, the strength of the spillovers may be stronger amongst households belonging to social networks, given that the extent of social interaction increases. This is not immediately borne by the results presented in this table: the coefficient on the average LPG use rate is positive and significant at the 1% level, however on comparing the values of these coefficients across network-types with the coefficient derived in column 1 of Table 5, there is found to be no significant difference between the two. This is the case for the women's groups, self-help groups, credit and savings organisations, and religious and social organisations. The only type of network where the spillover effect was stronger than the general spillover effect derived in column 1 of Table 5 is in caste associations. Due to decline in the number of observations, we do not estimate a IVM in this case, but these results may be driven by the endogeneity concerns that we mentioned before.

Table 3 in the appendix provides the results of the 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. Using 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 (column 1). Village-by-year time trends are also used (in column 2) to capture the effect of time-varying unobservables at the village/urban-block level. The IV-2SLS methodology is used for both estimations, and the first-stage results are provided in Table 5 in the appendix. 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 the possibility of the true effect being underestimated in the absence of supply data.

Table 5: IHDS Data Baseline Results

Dependent Variable: Whether HH I spent on LPG in the last 30 days	LPM	IV-2SLS (Overall) 2nd-Stage	IV-2SLS (Rural vs. Urban) 2nd-Stage	LPM (1999-00 LPG Adoption Rates)
Column	(1)	(2)	(3)	(4)
Average LPG use rate (Village/ Urban Block)	1.009***	1.157***	1.406***	
	(0.006)	(0.032)	(0.178)	
Avg. LPG Use Rate * Rural Indicator			-0.277	
			(0.178)	
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 than 40%				-0.720
				(1.247)
Whether HH lives in a rural area?	-0.089	-0.109**	0.167	-0.050
	(0.062)	(0.060)	(0.187)	(0.061)
Whether HH had access to electricity?	0.032***	-0.028**	-0.018	0.048***
	(0.008)	(0.017)	(0.013)	(0.015)
Household size	-0.002	0.007***	0.006***	-0.005**
	(0.002)	(0.003)	(0.002)	(0.003)
Number of Years of Education of Household Head	0.004***	-0.00007	0.0005	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Whether Household has Non-Biomass Cookstove?	0.074***	0.016	0.025*	0.075***
	(0.011)	(0.016)	(0.014)	(0.012)
Hours of Cookstove Use (/ Day)	-0.011***	-0.007***	-0.008***	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Time Spent in Collecting Firewood (/ Trip)	0.00004	-0.0002***	-0.0002***	0.00001
	(0.00007)	(0.0001)	(0.0001)	(0.00009)
Whether Household Has Vent in Kitchen?	0.0134**	0.018***	0.018***	0.013*
	(0.007)	(0.008)	(0.007)	(0.008)
Observations	18590	9350	9350	17072
Hansen J-statistic		7.93	16.544	
P-value		0.1601	0.1676	

Notes: Values reported are marginal effects. All specifications include household-level fixed effects and time trends. Results in columns 1,2 and 3 include dummies for whether the household belongs to the 7th, 8th, 9th or 10th income deciles, whereas the result in column 4 includes all income decile dummies. Controls are included for social group and religion in all specifications. *,** and *** respectively denote significance at 10%, 5% and 1% levels. Coefficient of constant for models in columns 1 and 4 has not been reproduced.

Table 6: IHDS Data 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 had 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

Notes: Values reported are marginal effects. All specifications include household-level fixed effects and time trends. All specifications include all 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. Coefficient of constant for models has not been reproduced.

5.2 Policy Implications

The paper's initial results offer interesting policy implications. 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, and this is has been found using four rounds of cross-sectional data, 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 attempt to correct for possible endogeneity in these estimations. While we do not have data on supply of LPG, we utilise the information that we are given, and econometric methodologies to correct for the lack of this data the best we can, and still find evidence to suggest that households adopt LPG if other households in the same village or urban block do so.

Additionally, the magnitude of these spillovers varies across rural and urban areas: we find evidence that social spillovers are strong in urban areas, but absent in rural areas. We hypothesise that this is driven by higher population density in urban areas, which would lead to more social interactions for urban households. This may also be driven by the fact that urban households have better supply of LPG than rural households do, but we obtain the same result when we use village fixed effects and village-by-year time trends with the IHDS data, in an attempt to correct for changing access to LPG (results not provided). Given that there is already a marked difference in adoption of LPG (and supply) between rural and urban areas in India, this result suggests that policy-makers in rural areas may need to use additional instruments to encourage households to switch to LPG use.

This paper also finds that these social spillovers are weaker for households residing in states which had higher LPG adoption rates to begin with, which provides some evidence in favour of the s-shaped diffusion curve. States which begin with higher rates of adoption experience weaker spillover effects, whereas states that are just beginning adoption might experience stronger spillover effects. This 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 that relatively "new" to the technology, it might make sense to target subsidies to certain households in these states, rather than disbursing them amongst the entire population.

We also attempt to investigate whether these social spillovers are stronger amongst households that belong to a social network. Our hypothesis was that these effects, when they exist, should be of higher magnitude than the overall social spillover effect of living in the same village/urban block. We are able to identify households belonging to five different kinds of organisations or social networks in the IHDS dataset, which allows us to get closer to the traditional definition of "peer-effects" used in the literature. While we do not find conclusive evidence for this claim in the empirical section, our hunch is that this may be driven by small sample size and the inability to control for endogeneity bias in these estimations, rather than a refutation of the hypothesis itself. These spillovers are also of use to policy-makers, who can benefit from organising informational campaigns for members of social groups, or targeting subsidies to them.

Given that India has very sharp variation in terms of rural and urban adoption of LPG, the results of this paper suggest that targeting "pivotal" households such as urban households residing in states which have not begun using LPG on a large scale may, by the presence of these spillovers, 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 there are factors such as imitation, health externalities, and learning-by-doing involved in the decision to mimic other households [18] Nevertheless, information provision may hasten the adoption process, especially in rural areas. However, by no means does this undermine the importance of other barriers towards greater adoption of LPG, such as lack of affordability, lack of supply, along with easy access to solid biomass such as firewood in rural areas.

This paper does not provide any evidence either supporting or refuting the effectiveness of

[18] .Kremer and Miguel (2007) suggest that these four factors may determine the causes of spillovers, or peer-effects, in the adoption of technologies.

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 these subsidies. However, as already mention, if social spillovers are indeed a factor in determining a households's choice of cooking fuel, subsidies to certain households in the beginning of the adoption process may actually be beneficial in ensuring that more households switch to the cleaner fuel.

6 Conclusion

The objective of this paper was to analyse whether there social spillovers in the adoption of a clean cooking fuel (LPG) in India, and if these exist, how they vary in strength in different parts of the country. In this paper we utilise two kinds of sample survey data from a widely heterogeneous population, which enables us to provide a broad scope in answering this research question. We obtain multiple pieces of evidence in this paper, that when analysed together suggest that social spillovers may be present in the Indian LPG context. We find differences between rural and urban households in the strength of these effects, and we also find differences over time. We control for several household-level characteristics of LPG adoption, that have already been proved to be important determinants in the literature. Our results have strong implications for policy-makers looking to encourage consumers to switch to cleaner sources of energy in developing countries. We provide evidence to suggest that social learning amongst consumers of energy products may be prevalent in the developing-country context, and could be utilised 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.

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.

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," *Genève: Haute école de gestion de Genève*.

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

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.

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

Desai, Sonalde and Reeve Vanneman (2009) "National Council of Applied Economic Research.(2005),"India Human Development Survey (IHDS)," Computer file, ICPSR22626-v5," *Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]*, Vol. 622.

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," *Journal of Economic Literature*, Vol. 53, pp. 5–42.

Gupta, Gautam and Gunnar Köhlin (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. Quarterly Journal of Economics. Vol. CXXII No. 3. August."

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, "Total Subsidy on PDS Kerosene and Domestic LPG to Customers," Government of India, May 2015.

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-Rodríguez, Adriana Teresa Marrón-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.

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: Essays in honor of Thomas Rothenberg*.

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

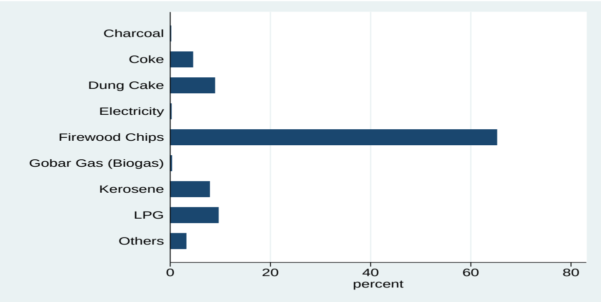
Verbeek, Marno (2008) "Pseudo-panels and repeated cross-sections," in *the econometrics of panel data*: Springer, pp. 369–383.

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

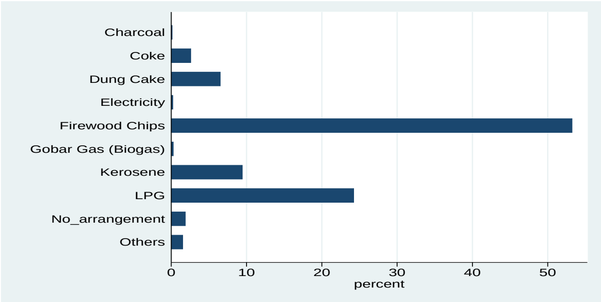
Young, H Peyton (2009) "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning," *The American economic review*, Vol. 99, pp. 1899–1924.

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.

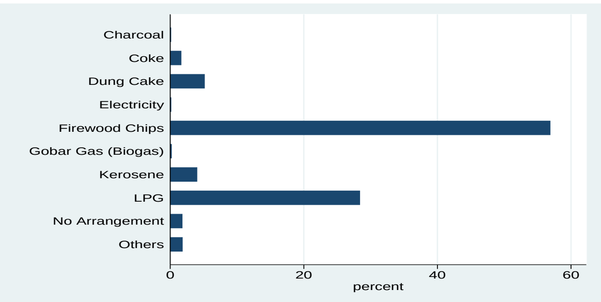
Appendix A Figures



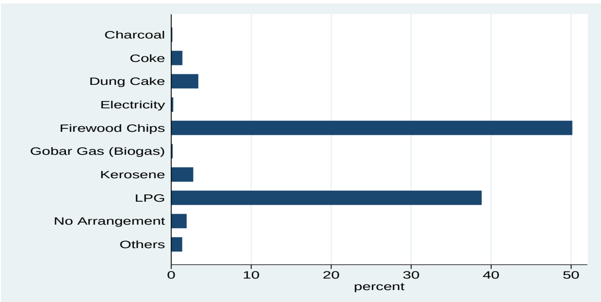
(a) 1987-88



(b) 1999-2000



(c) 2004-05



(d) 2009-10

Figure 1: Distribution of Households by Primary Fuel Type (Cooking) in Thick Rounds of the NSS

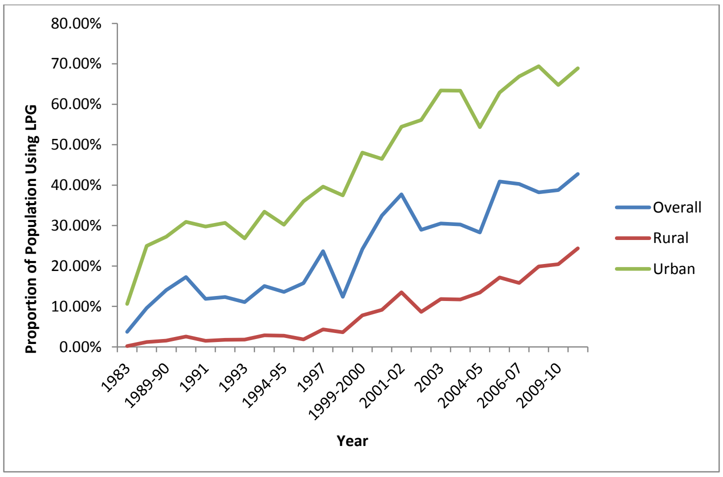
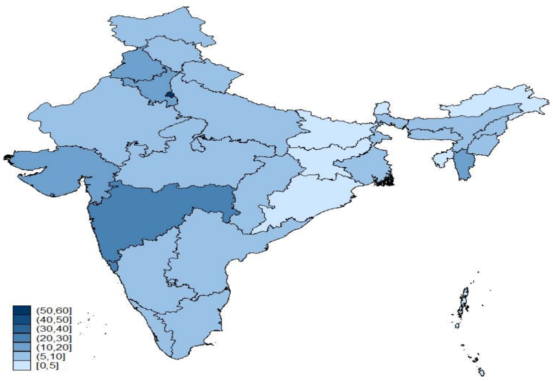
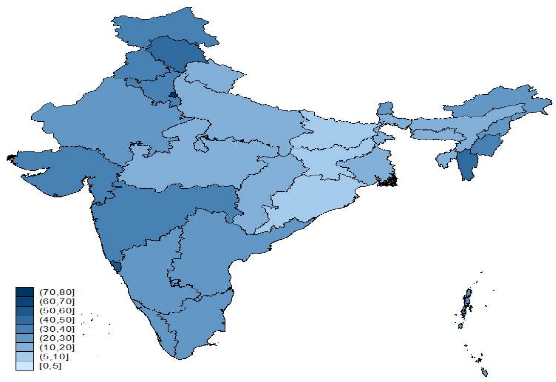


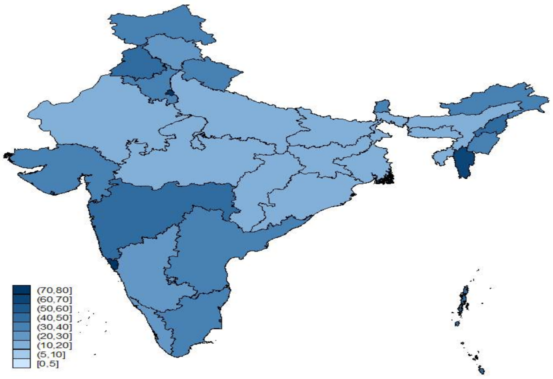
Figure 2: Proportion of Population Using LPG as the Primary Cooking Fuel: 1983 to 2011-12



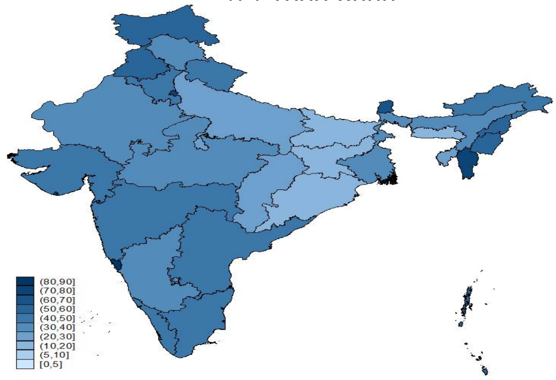
(a) 1997-98



(b) 1999-2000



(c) 2004-05



(d) 2009-10

Figure 3: Proportion of Population Using LPG as the Primary Cooking Fuel in the Thick Rounds of the NSS

Appendix B Tables

Table 1: 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 had 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 had 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-squared	0.6093	0.6046	0.6022	0.6234

Notes: Values reported are marginal effects. All specifications include dummy variables for districts, MPCE deciles and religion and social group (except for the 43rd round, where the religion and social group dummies are omitted due to lack of convergence). Standard errors are clustered at the village/urban block level (reported in parentheses). *,** and *** respectively denote significance at 10%, 5% and 1% levels. Coefficient of constant has not been reproduced.

Table 2: NSS Data: First-Stage Estimations

Round	43	55	61	66
Year	1987-88	1999-00	2004-05	2009-10
Corresponding Second-Stage Results	Table 4 (Text) Column 2	Table 4 (Text) Column 4	Table 4 (Text) Column 6	Table 4 (Text) Column 8
Dependent Variable: Average Village/ Urban Block LPG Use Rate	(1)	(2)	(3)	(4)
Proportion of population in 10th 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 had 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 had 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-Yogo (2005))	19.28	19.28	19.28	19.28
10% maximal IV size (Stock-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: Values reported are marginal effects. 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. Coefficient of constant has not been reproduced.

Table 3: 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 had 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: Values reported are marginal effects. Results in column (1) include household-specific time trends. All specifications include income decile dummies (8th to 10th) and controls for religion and social group. *,** and *** respectively denote significance at 10%, 5% and 1% levels.

Table 4: 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)	6.778***	0.360***
	(0.114)	(0.003)
Whether HH had access to electricity?	2.300***	0.122***
	(0.295)	(0.018)
Household size	-0.086***	-0.005***
	(0.013)	(0.0007)
Number of Years of Education of Household Head	0.101***	0.005***
	(0.008)	(0.000)
Whether Household has Non-Biomass Cookstove?	2.264***	0.120***
	(0.106)	(0.006)
Hours of Cookstove Use (/ Day)	-0.083***	-0.004***
	(0.021)	(0.001)
Time Spent in Collecting Firewood (/ Trip)	0.004***	0.0002***
	(0.001)	(0.00005)
Whether Household Has Vent in Kitchen?	0.259***	0.014***
	(0.072)	(0.004)
Observations	17072	17072

Notes: Values reported are marginal effects. 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. Robust standard errors are estimated for the population-averaged model . *,** and *** respectively denote significance at 10%, 5% and 1% levels.

Table 5: IHDS Data: First-Stage Estimations

Corresponding Second-Stage Results	Table 5 (Text) Column 2	Table 5 (Text) Column 3	Table 5 (Text) Column 3	Table 3 (Appendix) Column 1	Table 3 (Appendix) Column 2
Dependent Variable	Avg. Village LPG Use Rate (HH FE)	Avg. Village LPG Use Rate (HH FE)	Avg. Village LPG Use Rate *Rural indicator (HH FE)	Avg. Village LPG Use Rate (Village FE)	Avg. Village LPG Use Rate (Village by year time trends)
Column	(1)	(2)	(3)	(4)	(5)
Income : 4th Decile		1.036***	0.197		
		(0.345)	(0.411)		
Income: 5th Decile	1.802***	0.969***	0.531*	0.623***	0.628***
	(0.202)	(0.333)	(0.310)	(0.078)	(0.083)
Income: 6th Decile	2.149***	1.519***	0.4	0.816***	0.826***
	(0.201)	(0.531)	(0.413)	(0.071)	(0.077)
Income: 7th Decile	-1.072***	-1.344***	-1.803***	0.737***	0.753***
	(0.298)	(0.393)	(0.435)	(0.063)	(0.066)
Income: 8th Decile	-0.668*	-1.643***	-2.594***	-0.426***	-0.394***
	(0.396)	(0.465)	(0.531)	(0.119)	(0.129)
Income: 9th Decile	-0.830**	-1.310***	-1.868***	-0.317***	-0.292***
	(0.424)	(0.581)	(0.781)	(0.110)	(0.125)
Income:10th Decile	-1.073***	-2.137***	-3.351***	-0.191	-0.153
	(0.391)	(0.685)	(0.787)	(0.119)	(0.127)
Whether HH lives in a rural area?	0.146**	0.05	0.924***		
	(0.070)	(0.08)	(0.083)		
Whether HH had access to electricity?	0.372***	0.350***	0.345***	0.457***	0.453***
	(0.031)	(0.030)	(0.031)	(0.030)	(0.028)
Household size	-0.066***	-0.070***	-0.068***	-0.022***	-0.022***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)
Number of Years of Education of Household Head	0.022***	0.021***	0.021***	-0.003***	-0.003***
	(0.002)	(0.002)	(0.002)	(0.0007)	(0.0007)
Whether Household has Non-Biomass Cookstove?	0.362***	0.348***	0.342***	0.184***	0.187***
	(0.024)	(0.023)	(0.023)	(0.014)	(0.015)
Hours of Cookstove Use (/ Day)	-0.021***	-0.0170**	-0.019**	-0.018***	-0.017***
	(0.008)	(0.008)	(0.008)	(0.005)	(0.005)
Time Spent in Collecting Firewood (/ Trip)	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Whether Household Has Vent in Kitchen?	-0.037	-0.041**	-0.041***	-0.056***	-0.062***
	(0.023)	(0.022)	(0.022)	(0.010)	(0.011)
Observations	9350	9350	9350	16875	16495
5% maximal IV relative bias (Stock-Yogo (2005))	19.28	19.83	19.83	19.28	19.28
10% maximal IV size	29.18	36.36	36.36	29.18	29.18
Cragg Donald F-Statistic	32.75	22.69	21.81	36.98	33.53
P-value	0	0	0	0	0

Notes: Values reported are marginal effects. Exogenous instruments for the results in column 1 are the proportion of population (by village) belonging 4th to the 10th income deciles; in columns 2 and 3, they are the proportion of population belonging from the 5th to the 10th income deciles, along with interactions of rural indicator with these proportions. All specifications include income decile dummies (from the 7th to the 10th) and controls for religion and social group. *,** and *** respectively denote significance at 10%, 5% and 1% levels.