

Environment, Population, and Health Series

No. 8, May 2017

## **Household Cooking Fuel Use in Rural and Peri-urban Viet Nam: A Multilevel Longitudinal Analysis of Supply Side Factors**

Sumeet Saksena, Chinh Cong Tran, and Jefferson Fox

## Household Cooking Fuel Use in Rural and Peri-urban Viet Nam: A Multilevel Longitudinal Analysis of Supply Side Factors

Sumeet Saksena, Chinh Cong Tran, and Jefferson Fox

**East-West Center Working Papers** is an unreviewed and unedited prepublication series reporting on research in progress. The views expressed are those of the author and not necessarily those of the Center. East-West Center Working Papers are circulated for comment and to inform interested colleagues about work in progress at the Center.

Working Papers are available online for free at [EastWestCenter.org/ewcworkingpapers](http://EastWestCenter.org/ewcworkingpapers).

The East-West Center promotes better relations and understanding among the people and nations of the United States, Asia, and the Pacific through cooperative study, research, and dialogue. Established by the US Congress in 1960, the Center serves as a resource for information and analysis on critical issues of common concern, bringing people together to exchange views, build expertise, and develop policy options.

The Center's 21-acre Honolulu campus, adjacent to the University of Hawai'i at Mānoa, is located midway between Asia and the US mainland and features research, residential, and international conference facilities. The Center's Washington, DC, office focuses on preparing the United States for an era of growing Asia Pacific prominence.

The Center is an independent, public, nonprofit organization with funding from the US government, and additional support provided by private agencies, individuals, foundations, corporations, and governments in the region.

### [EastWestCenter.org/publications](http://EastWestCenter.org/publications)

Publications Office | East-West Center  
1601 East-West Road | Honolulu, Hawai'i 96848-1601

Tel: 808.944.7145 | Fax: 808.944.7376  
[EWCBooks@EastWestCenter.org](mailto:EWCBooks@EastWestCenter.org)

*Sumeet Saksena* is a senior fellow in the Research Program at the East-West Center. His interests include health risk assessments, peri-urban issues and environmental justice. He holds a Ph.D. in environmental science and engineering from the Indian Institute of Technology, Mumbai, India.

*Chinh Cong Tran* is vice director, Center for Community Support Development Studies Hanoi, Viet Nam. His interests include natural resources economics and modeling. He holds a Ph.D. from the Hawai'i at Mānoa, College of Tropical Agriculture and Human Resources.

*Jefferson Fox* is a senior fellow in the Research Program at the East-West Center. He conducts research on land-use and land-cover change in Asia and the impact of these changes on the region and the global environment. Other areas of study include resource-management systems and land-cover transitions in Montane Mainland Southeast Asia—their role in altering regional hydrological processes under a changing climate; the ethics, values, and practice of spatial information technology and society; and natural resources and violent ethnic conflict in the Asia Pacific region. He holds a Ph.D. from the University of Wisconsin-Madison.

This study is a part of the East-West Center's program on peri-urbanization and health.

# Household cooking fuel use in rural and peri-urban Viet Nam: A multilevel longitudinal analysis of supply side factors

Sumeet Saksena, Chinh C. Tran, Jefferson Fox  
The East-West Center, Honolulu, USA.

## **Abstract**

We analyzed the 2006 and 2011 agriculture census data of approximately 9500 communes to examine the role of supply and demand side factors in the choice of main cooking fuel. The use of fuel wood was found to be significantly associated with all forest types - whether they were natural or plantations; whether they were owned by households or other entities. We have provided evidence that officially designated rural places are not homogenous in terms of fuel use. There is a distinct difference between traditional rural communes and the more modern rural communes. The latter we term as peri-urban communes. In peri-urban communes wood usage decreased over the years. In rural communes wood use however increased. This may partly be due to a natural transition of moving up the energy ladder from crop residues to wood. We find some evidence that given Viet Nam's highly successful and well documented small-holder plantation programs that have led to significant reforestation there may be rural pockets in the country where movement up the energy ladder from wood to cleaner fuels has slowed down considerably. Peri-urban areas have a greater fraction of gas users and also have a greater diversity of fuel use patterns.

In the future household factors and behaviors are likely to continue being the dominant drivers of fuel switching. However, our study indicates that external interventions aimed at improving community conditions such as access to markets, roads and highways can also facilitate households moving up the energy ladder. These factors may affect modern fuel access directly or may also work at modernizing communes to create a broader wealth effect over a period of time. Our findings are of much relevance to the newly proposed policy paradigm of 'making the clean available' as opposed to 'making the available clean' (Smith 2015). Promoting and marketing top of the ladder clean fuels would require satisfactory marketing and transportation networks, which our study has shown to be associated with fuel use. Our community level analysis is also relevant to those agencies that wish to transform an entire community at a time rather than targeting only a few households within it.

# 1. Introduction

The use of open fires and traditional cook stoves and fuels is one of the world's most pressing health and environmental problems. Globally, three billion people rely on solid fuels to cook, causing serious environmental and health impacts that disproportionately affect women and children. According to the World Health Organization, household air pollution from cooking kills over 4 million people every year and sickens millions more. Burning solid fuels releases emissions of some of the most important contributors to global climate change: carbon dioxide, methane and other ozone producing gases such as carbon monoxide, as well as short-lived but very efficient sunlight-absorbing particles like black carbon and brown carbon. Unsustainable wood harvesting also contributes to deforestation, reducing carbon uptake by forests. Residential solid fuel burning accounts for 25% of global black carbon emissions, about 84% of which is from households in developing countries. Researchers, practitioners and policy makers need information on future projections of health and climate risks associated with cooking fuels in terms of both spontaneous change (business-as-usual scenario) or change in risks owing to impact of interventions and new policies. For both these purposes there is a need to know which factors influence the choice of household cooking fuel.

Most studies that have attempted to identify the determinants and correlates of household fuel use have focused mainly on demand side factors such as household income, wealth, etc. (Lewis and Pattanayak 2012; Rehfuess, Briggs, Joffe, *et al.* 2010) (Farsi, Filippini and Pachauri 2007; Heltberg 2004; Heltberg 2005; Heltberg, Arndt and Sekhar 2000; Mainali, Pachauri and Nagai 2012; Mehta, Gore, Prüss-Üstün, *et al.* 2006). Some studies have also considered the monetary cost of the fuel. There have been very few studies that have also considered supply side factors such as natural resources availability and access; community level infrastructure, etc. (Gundimeda and Kohlin 2008; Rehfuess, Briggs, Joffe, *et al.* 2010; Vahlne 2015).

There is a well-established theory of how urbanization influences household energy choice (Barnes, Krutilla and Hyde 2005). It is well known that fuel switching occurs more rapidly in urban places than in rural places (Gundimeda and Kohlin 2008; Heltberg 2004; Heltberg 2005). Previous studies have provided possible explanations for this, including a lack of infrastructure for modern fuels, lower income and wealth, traditional lifestyle, lower opportunity cost and higher availability of collectable fuels in rural areas. The availability of biomass has been found to also have a strong influence on the process of urban fuel switching (Barnes, Krutilla and Hyde 2005). The empirical relationship between the process of urbanization and fuel choice has been studied mainly by considering rural-urban differentials in usage patterns. This has been done either by modeling urban and rural populations separately or by including a binary term in single models. Such modeling approaches and the theories they are based on do not reflect the nuances of different levels of 'rurality'. That is, the prevalent view is that all places that are officially designated as rural are similar. However, recent literature

provides evidence that even within the rural sphere there exists different levels of ruralness or ‘rurality’. A recent study in Viet Nam found that the likeliness that household is or will start using LPG increases with increased wealth if the household resides in a less rural environment (Vahlne 2017). They based their definition of rurality on area level aspects such as distance to town, village average land, household density and commune mean income. The study therefore recommended that future models incorporate indicators of rurality.

We address the following research questions in our study:

1. What is the level of association between supply side factors such forests, croplands, community infrastructure, etc. and fuel use?
2. To what extent do relatively more modern rural areas differ in fuel usage compared to less modern rural areas? We use the concept of ‘peri-urban’ places to answer this question. Those rural places that in a relative sense have a few modern urban characteristics are classified by us as peri-urban places.

We wish to simultaneously estimate the effect of demand side factors that operate at the household level and supply side factors that operate at the community level. Multilevel modeling approaches are best suited to this task (Merlo, Chaix, Yang, *et al.* 2005). Only a few studies of household cooking fuel use have employed this technique (Bonjour, Adair-Rohani, Wolf, *et al.* 2013; Rehfuess, Briggs, Joffe, *et al.* 2010). None of the multilevel studies studied the time factor. We have used a longitudinal multilevel modeling approach to examine how temporal changes in the predictor variables effect the temporal change in the use of cooking fuel.

## **2. Methods**

### *2.1 Dependent variable and unit of analysis*

It was our intention to carry out a national analysis using large databases. Viet Nam’s Rural, Agriculture and Fisheries (RAF) census (described in detail under sources of data) contains an item on the main fuel used for cooking, similar to the item in population and housing censuses conducted in many other developing countries including Viet Nam (Bonjour, Adair-Rohani, Wolf, *et al.* 2013; Mehta, Gore, Prüss-Üstün, *et al.* 2006). The fuel choices in the RAF census are: wood, coal, gas, electricity and ‘other fuels’. ‘Other fuels’ for cooking relate mainly to crop residues in rural areas (Dollar, Glewwe and Litvack 1998; Vahlne 2015; Vahlne and Ahlgren 2014). Henceforth, we will use the term ‘miscellaneous’ for ‘other fuels’. As with most population census surveys, the respondent in the RAF census is only allowed to mention a single choice of main cooking fuel. Our unit of analysis was the commune. A commune in Viet Nam is the third level in the administrative structure – below the province (or centrally controlled city) and district. A commune is the lowest level at which local leaders are elected and it consists of a few villages/hamlets. Household level variables aggregated to the commune level was the only database format available to us. Thus the dependent variable is in the form of fraction of households in a commune that use a certain type of cooking fuel.

We measured fuel-use diversity using the Gini-Simpson Diversity Index (Hill 1973). The Gini-Simpson Diversity Index is given by  $1 - \lambda$ , where  $\lambda$  equals the probability that two entities taken at random from the dataset of interest represent the same type. In situations with only one class (complete homogeneity) the Gini-Simpson index would have a value equal to zero. Such diversity indices have been used to measure land-use diversity (Forman 1995).

## 2.2 Predictor variables

It is self-evident that households mainly obtain wood from forests and possibly also from scattered trees and shrubs on croplands and homesteads. Certain perennial crops (cashew, rubber, coffee, tea and fruits) may also yield fuel wood (FSIV 2009). We therefore include percentage of commune land under forests and crops in our models for wood. Furthermore, access to forests is determined also by ownership status. We therefore distinguish between forests owned by households and those that are owned by the government, cooperatives or enterprises. Plantation forestry by farm households (small holders) is common in Viet Nam (Meyfroidt and Lambin 2008; Meyfroidt and Lambin 2008; Sandewall, Ohlsson, Sandewall, *et al.* 2010). Both farmers and foresters have focused on non-native fast growing genera primarily *Acacia*, *Eucalyptus*, and *Pinus*. More recently there has been an effort to promote slower growing native species. Branches, tops and bark are often sold as fuelwood from such smallholder plantations (Tan 2011). It is also self-evident that crop residues are sourced from crops lands. Therefore land under annual and perennial crops reflects the potential supply of crop residues (termed as miscellaneous in our analysis). As there is no a priori reason to believe that there is an association between coal or gas use with forest and croplands, we did not include land use classes for these two fuels. To avoid the error that arises out of including a set of closed variables (land use categories in our study) in a model, we excluded two major land use categories - built up land and land under aquaculture. The access to both free and purchased fuels is partly dependent on commune infrastructure such as presence of markets and road network. The transportation of such fuels from the source or market to the house is also dependent on the road infrastructure.

On the demand side we included factors that are indicative of family income and wealth. The fraction of households that derive their main income from agriculture is the main surrogate indicator for income. As measures for wealth we included type of toilet; electrification; and ownership of TV, fridge and motorbike.

## 2.3 Data sources

GSO Viet Nam conducts a comprehensive Rural, Agriculture and Fisheries (RAF) census every five years. All households, farms and enterprises in officially designated rural areas are included in the survey. In rural areas data are also collected at the commune and enterprise level. In

urban areas, GSO only includes households and farms which are known, on an a priori basis, to be active in relevant activities. In urban areas the RAF census does not collect commune/city- or enterprise-level data. For our analysis, we excluded all officially designated urban households that were part of the agriculture survey. Thus our final database is reflective of the conditions of the approximately 60 million rural population in 2011. We further classified the official rural communes into rural and peri-urban areas. Peri-urban communes are defined as places that have mixed rural and urban characteristics. GSO does not have an official classification for peri-urban areas. We used the classification method developed recently for Viet Nam (Saksena, Fox, Spencer, *et al.* 2014) to identify peri-urban communes. Their method is based on cluster analysis of key factors that are indicative of agricultural transitions and urbanization.

The RAF includes basic demographic and housing information in addition to agricultural information. The household questionnaire includes questions of agricultural and forestry land owned by the household. The commune level questionnaire includes questions on basic infrastructure and on croplands and forests owned by the government, cooperatives, enterprises, etc. Our database consists of over 9400 communes that were surveyed in 2006 and 2011. Viet Nam's GIS commune level shapefile was made available to us through another project along with other common geographical layers (roads, rivers, etc.) (Saksena, Fox, Epprecht, *et al.* 2015).

#### ***2.4 Data coding, pre-processing and collinearity***

We assigned a code of 0 to the Year 2006 and 5 to the Year 2011. Rural places were assigned a code 1 and peri-urban places a code of 0. Places with a market were coded as 0 and where markets were absent the code was 1. A dummy variable was created for each of Viet Nam's eight agro-ecological zone – 0 for a commune being in that zone and 1 for not being in that zone. These coding schemes were designed given the peculiarities of the chosen statistical software. Prior to modeling we cleaned the data by identifying illogical values for all variables and then either assigning a missing value to them or adjusting the values. Illogical values occurred mainly (less than 1% of the cases) for land-related variables such as percentage of commune land under a particular type of land use. Next we tested each variable for normality using the BestFit software (Palisade Corporation). A few of the variables were found to follow a log-normal distribution and a log-transform was used on them. We then examined the bi-variate correlations between all the independent factors (or their log-transform, as the case may be). We also performed a factor analysis. At the commune level we found a high degree of correlation between agriculture income, household electrification, type of toilet and ownership of TVs and fridges. So we chose to only retain the agriculture income variable. Household ownership of motorbikes was weakly correlated with the other measures of wealth, so we included it in our model. We also included dummy variables for Viet Nam's eight agro-ecological zone.

We initially ran an ordinary least squares (OLS) regression model using forward elimination. The results of OLS regression indicated collinearity among the agro-ecological dummy variables. Hence for each fuel model we omitted one dummy variable separately. This does not mean that the final models did not contain data of the omitted agro-ecological zone. Only the dummy variable for that zone was excluded.

## ***2.5 Statistical modeling***

When the structure of the data is naturally hierarchical or nested, multilevel models can separate the variation in the dependent variable into individual and higher level (group or area) components (Merlo, Chaix, Yang, *et al.* 2005). We used a repeated measures three-level Linear Mixed Model (LMM). Time (year as the repeated measure) was treated as Level 1, with communes (Level 2) nested under districts (Level 3) to take account of random effects for an area influenced by its neighbors (Heck, Thomas and Tabata 2014). Longitudinal analysis using mixed effects allows for including time varying (e.g., croplands) and invariant (e.g., distance to nearest town) independent variables. We used the maximum likelihood estimation method. Considering that we had only two time intervals, we chose a scaled identity for the repeated measures covariance metric. We also chose scaled identity for measuring covariance of the random effects (levels 2 and 3). LMM is a ‘place’ oriented approach that is well suited to analyzing the effect of administrative grouping. The phenomenon we are studying is embedded in the natural context where physical proximity is also an important reason for clustering effects. A spatially derived autoregressive term is a ‘space’ oriented approach that accounts for the effects of physical proximity (Arcaya, Brewster, Zigler, *et al.* 2012). We used a hybrid modeling approach of using a LMM with an autoregressive term (Rehfuess, Briggs, Joffe, *et al.* 2010). The autocovariate term was calculated by using spatial weights matrix, weighted by the inverse of the Euclidean distance (Augustin, Muggleston and Buckland 1996). The limit of the autocorrelation of the response variable was obtained from the range of the spatial correlogram  $\rho(h)$  (Pebesma 2004). We calculated the autocorrelation term separately for 2006 and 2011. To compare models we used the Akaike Information Criterion (AIC). The lower the value of AIC the better suited the model. We used the MIXED module of the IBM SPSS version 23 software (IBM, New York).

## ***2.6 Building models***

We began by running simple ordinary least squares regression models including the spatial autocorrelation term. Next we ran null LMM models using only random effects (three level nested structure) with no predictors. To the null model we then added all predictors other than the spatial autocorrelation term. The full and final model had all the predictors and the autocorrelation term. Across the LMM models we compared suitability using the AIC.

### 3. Results

#### 3.1 Trends in fuel usage

At the national level (excluding urban areas) the use of wood as the main cooking fuel decreased in 2011 compared to 2006 (**Table 1**) by six percentage points. However, in rural areas the use of wood increased in 2011. We speculate that this is because households shifted from using miscellaneous fuels and coal to wood. Examining the spatial and temporal trends across agro-ecological zones, we found that in all zones except the Red River Delta, the use of wood decreased (**Figure 1**). We believe it is the same shift from coal and miscellaneous fuels to wood that explains this pattern. The use of coal as the main cooking fuel was found to be significant only in the RRD and to some extent in Northeast, as the country's main coal mines are in these regions. The areas around Ha Noi (Red River Delta) and Ho Chi Minh City (South East South) showed the highest use of cooking gas (**Figure 2**).

The use of coal and gas was significantly higher in peri-urban communes compared to rural communes. Wood usage was lower in peri-urban areas compared to rural areas. In 2006 the fraction of houses using miscellaneous fuels in rural areas was higher than in peri-urban areas. However, in 2011 the fraction of houses using miscellaneous fuels was higher in peri-urban areas. A similar situation has been observed in RRD by another study (Vahlne et al.). We speculate this might be occurring because households that were using coal instead of being able to switch to using wood found it easier and less expensive to switch to miscellaneous fuels. Less access to wood from forests combined with surplus of crop residues from increased agricultural activities would facilitate this transition.

As with many other characteristics, peri-urban communes were found to be highly heterogeneous in terms of mix of fuels used for cooking, compared to rural communes (**Figure 3**). In 2011 almost half the peri-urban residents used gas; one-third used wood; coal and miscellaneous fuel users accounted for slightly less than one-tenth of the households each. In contrast in rural communes, approximately three-quarters of the households used wood and almost one-fourth of the households used gas.

#### 3.2 Trends in predictor variables

Households whose main income is from agriculture decreased by 13 percentage points in 2011 compared to 2006 (**Table 2**). In the same period, household ownership of motorbikes increased by 32 percentage points. Fraction of land under natural forests and plantations owned by the government, cooperatives, etc. increased by about 1 percentage point each. While the fraction of land under plantations owned by households increased by 1 percentage point, the land under household owned natural forests decreased by 1 percentage point. In rural communes almost 4% of the commune land was under household owned plantations, whereas in peri-urban communes only 1% of land was under similar forests. The change in land under all forest categories was small but statistically significant. There was a slight increase in land under

perennial crops accompanied by a slight decrease in annual crops. Both these trends were statistically significant. The changes in land under annual and perennial crops were not correlated. It was therefore not a situation where land under annual crops shifted to perennial crops or vice versa. It is more likely that land under crops changed to built-up land.

### **3.3 Models**

#### **3.3.1 Model performance**

For all four fuels we found that the full LMM model had a vastly superior goodness of fit compared to the fixed effects only OLS regression model (**Tables 3-6**). So for example in the case of fuel wood the full LMM model had an accuracy of 95%, while the OLS regression model had an accuracy of 76%. Thus the clustering effects of nested administrative units and spatial autocorrelation do need to be incorporated to construct accurate models as well as determine realistic coefficients of the fixed effects. By comparing the AIC we found that for all four fuels the full model with spatial autocorrelation was superior to the similar model without the autocorrelation term, which in turn was superior to the null model (empty model with no predictors).

#### **3.3.2 Model findings**

All four categories of forests were found to be significantly associated with fuel wood use (**Table 3**). A significant positive association was found between land under perennial crops and wood use, supporting the belief that certain perennial crop species provide fuelwood as a byproduct, or that agroforestry practices and scattered trees on crop lands provide wood. The significant negative association between land under annual crops (such as paddy) and wood use is not easily explained. It may reflect competing land uses. The model predicted less wood fuel use in peri-urban areas as compared to rural areas. This is consistent with the theory that increasing levels of urbanization and modernization would enable a shift to cleaner fuels. Places with markets had less use of wood compared with those that did not have a market. The distance of the commune to the national highway and to the nearest town was found to have a positive association with wood use. That is, places which are remote and poorly connected have a higher use of fuelwood. These last three factors indicate that communes with less developed infrastructure are more likely to have higher fraction of fuelwood users. As expected, a positive association was found between the fraction of houses whose main income is from agriculture and fuelwood use. The household wealth indicator, ownership of a motorbike, was not significantly associated with fuelwood use. This may indicate that within poor communes, supply side constraints prevent even wealthier families from shifting away from wood. The year variable in our wood model was found to be insignificant, indicating that the time variations in the other variables included in the model were sufficient to explain the time variation in the use of fuelwood. The variance across communes within a district accounted for 32% of the total

variance in the outcome; the variance across districts accounted for 28% and across time the variance accounted for was 40%.

Land under annual crops was found to be positively associated with the use of miscellaneous fuels, which we believe are mainly crop residues (**Table 4**). A negative association was found between household owned plantations and miscellaneous fuels use. This may reflect competing land uses affecting supply of crop residues. The miscellaneous fuels' association with distance to highway and towns, presence of a market and peri-urban status was similar to that of fuelwood. While we found a positive association between use of miscellaneous fuels and agriculture income, as in the case of wood, we found a negative association with ownership of motorbikes. This indicates that though for wood wealth cannot overcome supply side related constraints, but wealth does facilitate shift from miscellaneous fuels to cleaner fuels up the energy ladder. The variance across communes within a district accounted for 25% of the total variance in the outcome; the variance across districts accounted for 23% and across time the variance accounted for was 52%.

Our model for coal predicts higher usage in peri-urban areas and in places that have a market (**Table 5**). Coal being a fuel that cannot be harvested by households needs a market system to be available. Places that were closer to highways had a higher usage of coal. An expected negative association was found between agriculture income and coal use. The same patterns were observed for the use of gas (**Table 6**). Additionally for gas a positive association was found between use and ownership of motorbikes. For coal, the variance across communes within a district accounted for 37% of the total variance in the outcome; the variance across districts accounted for 32% and across time the variance accounted for was 31%. For gas, the variance across communes within a district accounted for 37% of the total variance in the outcome; the variance across districts accounted for 13% and across time the variance accounted for was 50%.

## 4 Conclusions and Policy Relevance

We have provided evidence that cooking fuel use in officially designated rural areas is simultaneously affected by household level demand side factors and community level supply side factors. Our study is one of the few community level studies that have quantified the association between cooking fuel use and supply side factors, especially those related to land uses such as forestry and agriculture. Most previous studies measured the relationship between wood use and forests in terms of distance of a village to the nearest forest. Important as this metric is, especially from a gender and labor perspective, it does not provide information on what happens if the size of the forest increases or decreases. Just one previous study accounted for fraction of land under forests at much larger regional levels (Gundimeda and Kohlin 2008). Another study experimented with including annual fuelwood production at the national level but did not find statistical significance at their scale of analysis (Mehta, Gore, Prüss-Üstün, *et al.*

2006). Our study also improves on the categorization of forests, in terms of household vs. government ownership and natural vs. plantation forests. A notable finding is related to the strong association between fuelwood use and plantations owned by households. In modern rural places (peri-urban ) the use of wood decreased between 2006-2011. In contrast, in traditional rural places the use of wood actually increased between 2006 and 2011. Some of this change may have happened because users of miscellaneous fuels (such as crop residues) upgraded to wood. Nevertheless, we speculate that the success of the small holder plantation program in Viet Nam may have at the minimum slowed down the transition from wood to gas. Or the program may have contributed to the ‘stacking’ effect where households cannot identify the ‘main’ fuel they use because they use multiple types of fuels for different tasks, in different seasons, etc. . A clean fuel such as gas is used along with fuelwood. The official small holder plantation programs were designed with an aim to alleviate poverty and there is no documentation that there was any clean energy related goals in these programs (Sam and Trung 2001). Thus it is likely that in the short term wood surplus at the farm level or community level discourages households from investing in cleaner cooking systems. Over the long term, if the small holder plantation program does succeed in poverty alleviation then it is likely that households would shift on their own to costlier cleaner fuels. Many countries such as Malaysia, Myanmar and parts of Indonesia have been traditionally practicing small holder plantations (FAO 1982). In others such as China, Viet Nam, India, Laos, the Philippines, Ethiopia, South Africa, Ghana and Tanzania recent government programs have expanded considerably (Sandewall, Kassa, Wu, *et al.* 2015). We therefore recommend more studies on the impacts of such plantation programs on the rural cooking energy ladder.

Our study has also quantified the association between market access and transport infrastructure with the use of all the main fuels. Places with less developed markets and road infrastructure were found to have a higher fraction of users of dirtier fuels and a lesser fraction of users of cleaner fuels. These insight are relevant to the new paradigm of ‘make the clean available’ that has been proposed. Promoting and marketing top of the ladder clean fuels would require satisfactory marketing and transportation networks.

The validity of treating all rural places as the same in models or in practice has been brought into serious question by our work. Distinguishing between the traditional rural places and the more modern rural places highlights not just the diversity in fuel use practices but also the rate and direction of change in fuel use. We used the concept of ‘peri-urban’ places and found similarities with the ‘rurality’ classification that has been recently proposed (Vahlne 2017). The peri-urban approach and the rurality approach essentially urge researchers and practitioners to discard the rural-urban dichotomy concept of modeling and planning. By extension of this logic, we recommend that even future studies of urban fuel use should aim to differentiate between modern urban places and the less modern urban places (Saksena, Fox, Spencer, *et al.* 2014). For a few health outcomes previous studies have shown that the rural-

urban dichotomy paradigm is either not valid or is not a refined way of examining the true effects of urbanization and modernization (Allender, Foster, Hutchinson, *et al.* 2008; Allender, Wickramasinghe, Goldacre, *et al.* 2011; Allender, Lacey, Webster, *et al.* 2010; Coogan, White, Evans, *et al.* 2011; Dahly and Adair 2007; Jones-Smith and Popkin 2010; McDade and Adair 2001; Novak, Allender, Scarborough, *et al.* 2012; Riha, Karabarinde, Ssenyomo, *et al.* 2014; Saksena, Fox, Epprecht, *et al.* 2015). We believe then that for all the health effects too associated with the use of cooking fuels the same would hold.

One of the limitations of our study is that the unit of analysis was a commune instead of a household. This may have created errors known as the ‘ecological fallacy’, where interpretations about households are being made based on community data. Our primary intention was to identify statistically significant predictors and direction of their influence (positive or negative). We speculate that our findings would not change drastically even if an analogous model is created at the household level. On the other hand, our community level analysis would be highly appropriate under the newly proposed ‘it takes a village’ paradigm (Smith 2015). This paradigm assumes that the most effective interventions are likely to occur at the community level, reducing the air pollution exposure of the entire community in one go. There are mass scale economic, social and cultural benefits of this paradigm. We argue then that our community level models would help practitioners, intervention agencies, community based organizations to identify best candidate communities for their programs. Another limitation of our study is that it did not consider the use of more than one cooking fuel by a household, known as the stacking effect. Official global risk assessments still do not incorporate the stacking effect but future studies should address this knowledge gap.

Our study also highlights the usefulness of additional sources of data, including government databases. Till recently household energy modeling and health risk modeling have relied mainly on national population and housing censuses (PHC); and official sample surveys such as Demographic and Health Surveys (DHS), Living Standard Surveys (LSS) and Consumer Expenditure Surveys (CES) . We encourage researchers and practitioners to also consider, as we did, the Agriculture, Forestry and Fisheries (AFF) census, that many developing countries conduct periodically. AFF censuses usually have land use information at the household and community level in addition to community infrastructure information. At the least, researchers should fully utilize the information available in the community level questionnaires and forms associated with the DHS, LSS and CES. We also see an important role in the future for databases on land-use and land-cover obtained through remote sensing. We acknowledge though that downscaling remote sensing data to the village level is not a simple task. Further refinements in terms of data on forest cover, tree species, crop types, types of rural markets and roads, etc. would provide an even more nuanced understanding of the spontaneous changes to occur in fuel consumption as well enable intervention agencies to cost-effectively target candidate

communities. The modeling of household cooking energy would benefit by becoming far more multidisciplinary in nature.

## Acknowledgements

The authors would like to thank the General Statistics Office, Ha Noi, Viet Nam for providing the commune level data from the agriculture censuses. We also wish to thank Prof. Ronald H Heck, University of Hawaii, Honolulu, USA for advising us on mixed effects modeling. Finally, we wish to thank Prof. Douglas Barnes, The World Bank, Washington D.C., USA for suggestions regarding modeling strategies.

## References

- Allender S, Foster C, Hutchinson L, and Arambepola C. 2008  
**Quantification of urbanization in relation to chronic diseases in developing countries: a systematic review**  
*Journal of urban health : bulletin of the New York Academy of Medicine* **85**(6): 938-951
- Allender S, Wickramasinghe K, Goldacre M, Matthews D, and Katulanda P. 2011  
**Quantifying urbanization as a risk factor for noncommunicable disease**  
*Journal of urban health : bulletin of the New York Academy of Medicine* **88**(5): 906-918
- Allender S, Lacey B, Webster P, Rayner M, Deepa M, Scarborough P, Arambepola C, Datta M, and Mohan V. 2010  
**Level of urbanization and noncommunicable disease risk factors in Tamil Nadu, India**  
*Bull World Health Organ* **88**(4): 297-304
- Arcaya M, Brewster M, Zigler C M, and Subramanian S V. 2012  
**Area variations in health: a spatial multilevel modeling approach**  
*Health & place* **18**(4): 824-831
- Augustin N H, Muggleston M A, and Buckland S T. 1996  
**An autologistic model for the spatial distribution of wildlife**  
*Journal of Applied Ecology* **33**(2): 339-347
- Barnes D F, Krutilla K, and Hyde W F. 2005  
**The urban household energy transition : social and environmental impacts in the developing world**  
Washington, DC: Resources for the Future : Energy Sector Management Assistance Program. xiii, 141 p. pp.
- Bonjour S, Adair-Rohani H, Wolf J, Bruce N G, Mehta S, Pruss-Ustun A, Lahiff M, Rehfuess E A, Mishra V, and Smith K R. 2013  
**Solid Fuel Use for Household Cooking: Country and Regional Estimates for 1980-2010**  
*Environ Health Perspect* **121**(7): 784-790
- Coogan P F, White L F, Evans S R, Adler T J, Hathaway K M, Palmer J R, and Rosenberg L. 2011  
**Longitudinal assessment of urban form and weight gain in African-American women**  
*Am J Prev Med* **40**(4): 411-418
- Dahly D L and Adair L S. 2007  
**Quantifying the urban environment: a scale measure of urbanicity outperforms the urban-rural dichotomy**  
*Soc Sci Med* **64**(7): 1407-1419
- Dollar D, Glewwe P, and Litvack J I. 1998  
**Household welfare and Vietnam's transition**  
Washington, D.C.: World Bank. xvii, 330 p. pp.

- FAO. 1982  
**FAO Village Forestry Inventory: Project Findings and Recommendations. Report Number DP/BGD/78/020**  
Food and Agriculture Organization. Rome, Italy
- Farsi M, Filippini M, and Pachauri S. 2007  
**Fuel choices in urban Indian households**  
*Environment and Development Economics* **12**: 757-774
- Forman R T T. 1995  
**Land mosaics : the ecology of landscapes and regions**  
Cambridge ; New York: Cambridge University Press. xx, 632 p. pp.
- FSIV. 2009  
**Vietnam Forestry Outlook Study**  
Working Paper No. APFSOS II/WP/2009/09  
Food and Agriculture Organization. Bangkok, Thailand
- Gundimeda H and Kohlin G. 2008  
**Fuel demand elasticities for energy and environmental policies: Indian sample survey evidence**  
*Energy Econ* **30**(2): 517-546
- Heck R H, Thomas S L, and Tabata L N. 2014  
**Multilevel and longitudinal modeling with IBM SPSS**  
New York: Routledge, Taylor & Francis Group. xvii, 443 pages pp.
- Heltberg R. 2004  
**Fuel switching: evidence from eight developing countries**  
*Energy Econ* **26**(5): 869-887
- Heltberg R. 2005  
**Factors determining household fuel choice in Guatemala**  
*Environment and Development Economics* **10**: 337-361
- Heltberg R, Arndt T C, and Sekhar N U. 2000  
**Fuelwood consumption and forest degradation: A household model for domestic energy substitution in rural India**  
*Land Economics* **76**(2): 213-232
- Hill M O. 1973  
**Diversity and evenness: A unifying notation and its consequences**  
*Ecology* **54**(2): 427
- Jones-Smith J C and Popkin B M. 2010  
**Understanding community context and adult health changes in China: development of an urbanicity scale**  
*Soc Sci Med* **71**(8): 1436-1446

- Lewis J J and Pattanayak S K. 2012  
**Who Adopts Improved Fuels and Cookstoves? A Systematic Review**  
*Environ Health Perspect* **120**(5): 637-645
- Mainali B, Pachauri S, and Nagai Y. 2012  
**Analyzing cooking fuel and stove choices in China till 2030**  
*Journal of Renewable & Sustainable Energy* **4**(3): 031805-031805-031815
- McDade T W and Adair L S. 2001  
**Defining the "urban" in urbanization and health: a factor analysis approach**  
*Soc Sci Med* **53**(1): 55-70
- Mehta S, Gore F, Prüss-Üstün A, Rehfuess E, and Smith K. 2006  
**Modeling household solid fuel use towards reporting of the Millennium Development Goal indicator**  
*Energy for Sustainable Development* **10**(3): 36-45
- Merlo J, Chaix B, Yang M, Lynch J, and Rastam L. 2005  
**A brief conceptual tutorial of multilevel analysis in social epidemiology: linking the statistical concept of clustering to the idea of contextual phenomenon**  
*J Epidemiol Community Health* **59**(6): 443-449
- Meyfroidt P and Lambin E F. 2008  
**The causes of the reforestation in Vietnam**  
*Land Use Pol* **25**(2): 182-197
- Meyfroidt P and Lambin E F. 2008  
**Forest transition in Vietnam and its environmental impacts**  
*Glob Change Biol* **14**(6): 1319-1336
- Novak N L, Allender S, Scarborough P, and West D. 2012  
**The development and validation of an urbanicity scale in a multi-country study**  
*BMC Public Health* **12**: 530
- Pebesma E J. 2004  
**Multivariable geostatistics in S: the gstat package**  
*Comput Geosci* **30**(7): 683-691
- Rehfuess E A, Briggs D J, Joffe M, and Best N. 2010  
**Bayesian modelling of household solid fuel use: Insights towards designing effective interventions to promote fuel switching in Africa**  
*Environ Res* **110**(7): 725-732
- Riha J, Karabarinde A, Ssenyomo G, Allender S, Asiki G, Kamali A, Young E H, Sandhu M S, and Seeley J. 2014  
**Urbanicity and lifestyle risk factors for cardiometabolic diseases in rural Uganda: a cross-sectional study**  
*PLoS Med* **11**(7): e1001683

Saksena S, Fox J, Spencer J, Castrence M, DiGregorio M, Epprecht M, Sultana N, Finucane M, Nguyen L, and Vien T D. 2014

**Classifying and mapping the urban transition in Vietnam**

*Applied Geography* **50**: 80-89

Saksena S, Fox J, Epprecht M, Tran C C, Nong D H, Spencer J H, Nguyen L, Finucane M L, Tran V D, and Wilcox B A. 2015

**Evidence for the Convergence Model: The Emergence of Highly Pathogenic Avian Influenza (H5N1) in Viet Nam**

*PLoS One* **10**(9): e0138138

Sam D D and Trung L Q. 2001

**Forest Policy Trends in Vietnam**

*Policy Trend Report 2001*: 69-73

Sandewall M, Ohlsson B, Sandewall R K, and Viet L S. 2010

**The Expansion of Farm-Based Plantation Forestry in Vietnam**

*Ambio* **39**(8): 567-579

Sandewall M, Kassa H, Wu S, Khoa P V, He Y, and Ohlsson B. 2015

**Policies to promote household based plantation forestry and their impacts on livelihoods and the environment: cases from Ethiopia, China, Vietnam and Sweden**

*Int For Rev* **17**(1): 98-111

Smith K R. 2015

**Changing paradigms in clean cooking**

*Ecohealth* **12**(1): 196-199

Tan N Q. 2011

**Chopping for chips: an analysis of wood flows from smallholder plantations in Vietnam**

Working Paper 65

Center for International Forestry Research. Bogor, Indonesia

Vahlne N. 2015

**Fuel choice, fuel switching and improved cookstoves in Vietnamese households: Analysis, models and proposals for new solutions**

Gothenburg, Sweden: Chalmers University of Technology. 72 pp. [Doctoral thesis submitted to the Department of Energy and Environment]

Vahlne N. 2017

**On LPG usage in rural Vietnamese households**

*Development Engineering* **2**: 1-11

Vahlne N and Ahlgren E O. 2014

**Policy implications for improved cook stove programs—A case study of the importance of village fuel use variations**

*Energy Policy* **66**: 484-495

**Table 1: Fraction of households (mean) by cooking fuel type, type of place and year**

Fuel	Type of place	Year		Paired t-test t, p value	
		2006	2011		
Wood	Rural	0.66 (0.33)	0.71 (0.24)	1412, p < 0.0001	
	Peri-urban	0.34 (0.27)	0.30 (0.25)		
	National	0.60 (0.34)	0.54 (0.32)		
Miscellaneous	Rural	0.14 (0.25)	0.03 (0.10)		2014, p < 0.0001
	Peri-urban	0.12 (0.20)	0.08 (0.14)		
	National	0.14 (0.24)	0.05 (0.12)		
Coal	Rural	0.07 (0.15)	0.02 (0.15)	1454, p < 0.0001	
	Peri-urban	0.15 (0.20)	0.08 (0.13)		
	National	0.08 (0.17)	0.05 (0.11)		
Gas	Rural	0.12 (0.12)	0.22 (0.18)		-4389, p < 0.0001
	Peri-urban	0.38 (0.24)	0.53 (0.24)		
	National	0.17 (0.18)	0.35 (0.25)		

n = 9497 communes; value in parenthesis is the standard deviation; note: National does not include officially designated urban places

**Table 2: Trends (mean) in demand and supply side predictors**

Independent variable	Year		Paired t-test t value, p
	2006	2011	
Fraction of houses with main income from agriculture	0.774 (0.194)	0.647 (0.250)	-92, p<0.001
Fraction of houses that own a motorbike	0.557 (0.260)	0.873 (0.158)	161, p<0.001
<i>Fraction of commune land under:</i>			
Annual crops	0.274 (0.220)	0.263 (0.201)	-14, p<0.001
Perennial crops	0.067 (0.126)	0.070 (0.138)	6, p<0.000
Natural forests owned by households	0.033 (0.067)	0.026 (0.060)	-16, p<0.001
Plantations owned by households	0.033 (0.087)	0.046 (0.091)	16, p <0.001
Natural forests owned by government, etc.	0.058 (0.124)	0.067 (0.138)	23, p<0.001
Plantations owned by the government, etc.	0.120 (0.227)	0.127 (0.231)	14, p<0.001

n = 9497 communes; value in parenthesis is the standard deviation

**Table 3: Longitudinal Linear Mixed Model with Spatial Autocorrelation for Wood**

Parameter	Estimate	Standard error	Significance, p<
Intercept	-0.525	0.066	0.001
Year	0.003	0.003	0.409
Spatial autocorrelation	0.803	0.025	0.001
<i>Household level demand side parameters</i>			
Main income from agriculture <sup>1</sup>	0.280	0.009	0.001
Motorbike ownership <sup>1</sup>	0.003	0.009	0.766
<i>Community level supply side parameters</i>			
Peri-urban class	-0.027	0.003	0.001
Market present	-0.015	0.002	0.001
Distance to national highway <sup>*</sup>	0.006	0.001	0.001
Distance to town <sup>*</sup>	0.008	0.002	0.001
Plantation forest owned by household <sup>2</sup>	0.068	0.015	0.001
Natural forest owned by household <sup>2</sup>	0.180	0.022	0.001
Plantation forest owned by government, etc. <sup>2</sup>	0.079	0.009	0.001
Natural forest owned by government, etc. <sup>2</sup>	0.069	0.007	0.001
Annual crops land <sup>2</sup>	-0.116	0.010	0.001
Perennial crops land <sup>2</sup>	0.050	0.014	0.001
<i>Agro-ecological zone</i>			
Red River Delta	0.185	0.015	0.001
North East	0.010	0.012	0.436
North West	-0.006	0.017	0.737
North Central Coast	omitted		
South Central Coast	0.060	0.015	0.001
Central Highlands	0.081	0.015	0.001
Southeast South	0.145	0.015	0.001
Mekong River Delta	0.025	0.012	0.047
<i>Model performance</i>			
Goodness of fit, r	0.978		0.01
AIC full model	-21968		
AIC without autocorrelation	-21077		
AIC null model	-16844		
Goodness of fit of OLS regression model, r	0.869		0.01

n = 17239

<sup>1</sup>Fraction of households; <sup>2</sup>Fraction of commune land under land use class; <sup>\*</sup>Transform of the type  $\log_{10}(1+x)$  was used; AIC = Akaike Information Criterion

**Table 4: Longitudinal Linear Mixed Model with Spatial Autocorrelation for Miscellaneous Fuels**

Parameter	Estimate	Standard error	Significance, p<
Intercept	0.321	0.040	0.001
Year	0.008	0.002	0.001
Spatial autocorrelation	0.938	0.032	0.001
<i>Household level demand side parameters</i>			
Main income from agriculture <sup>1</sup>	0.081	0.007	0.001
Motorbike ownership <sup>1</sup>	-0.072	0.007	0.001
<i>Community level supply side parameters</i>			
Peri-urban class	-0.026	0.003	0.001
Market present	-0.001	0.002	0.669
Distance to national highway*	0.001	0.001	0.276
Distance to town*	0.000	0.001	0.850
Plantation forest owned by household <sup>2</sup>	-0.008	0.012	0.536
Natural forest owned by household <sup>2</sup>	-0.077	0.017	0.001
Plantation forest owned by government, etc. <sup>2</sup>	-0.015	0.008	0.043
Natural forest owned by government, etc. <sup>2</sup>	-0.012	0.005	0.019
Annual crops land <sup>2</sup>	0.152	0.007	0.001
Perennial crops land <sup>2</sup>	-0.029	0.011	0.009
<i>Agro-ecological zone</i>			
Red River Delta	-0.126	0.011	0.001
North East	omitted		
North West	0.012	0.011	0.273
North Central Coast	-0.011	0.009	0.213
South Central Coast	-0.079	0.010	0.001
Central Highlands	-0.076	0.010	0.001
Southeast South	-0.120	0.010	0.001
Mekong River Delta	-0.032	0.009	0.001
<i>Model performance</i>			
Goodness of fit, r	0.941		0.01
AIC full model	-30093		
AIC without autocorrelation	-29331		
AIC null model	-27600		
Goodness of fit of OLS regression model, r	0.777		0.01

n = 17239

<sup>1</sup>Fraction of households; <sup>2</sup>Fraction of commune land under land use class; \*Transform of the type log<sub>10</sub>(1+x) was used; AIC = Akaike Information Criterion

**Table 5: Longitudinal Linear Mixed Model with Spatial Autocorrelation for Coal**

Parameter	Estimate	Standard error	Significance, p<
Intercept	0.079	0.041	0.055
Year	0.002	0.002	0.405
Spatial autocorrelation	1.596	0.041	0.001
<i>Household level demand side parameters</i>			
Main income from agriculture <sup>1</sup>	-0.070	0.005	0.001
Motorbike ownership <sup>1</sup>	-0.004	0.005	0.472
<i>Community level supply side parameters</i>			
Peri-urban class	0.013	0.002	0.001
Market present	0.002	0.001	0.207
Distance to national highway*	-0.001	0.000	0.004
Distance to town*	0.000	0.001	0.707
<i>Agro-ecological zone</i>			
Red River Delta	0.040	0.010	0.001
North East	0.020	0.008	0.011
North West	0.007	0.011	0.477
North Central Coast	omitted		
South Central Coast	-0.034	0.009	0.001
Central Highlands	-0.028	0.009	0.003
Southeast South	-0.032	0.009	0.001
Mekong River Delta	-0.037	0.008	0.001
<i>Model performance</i>			
Goodness of fit, r	0.960		0.01
AIC full model	-40911		
AIC without autocorrelation	-40211		
AIC null model	-38625		
Goodness of fit of OLS regression model, r	0.681		0.01

n = 17239

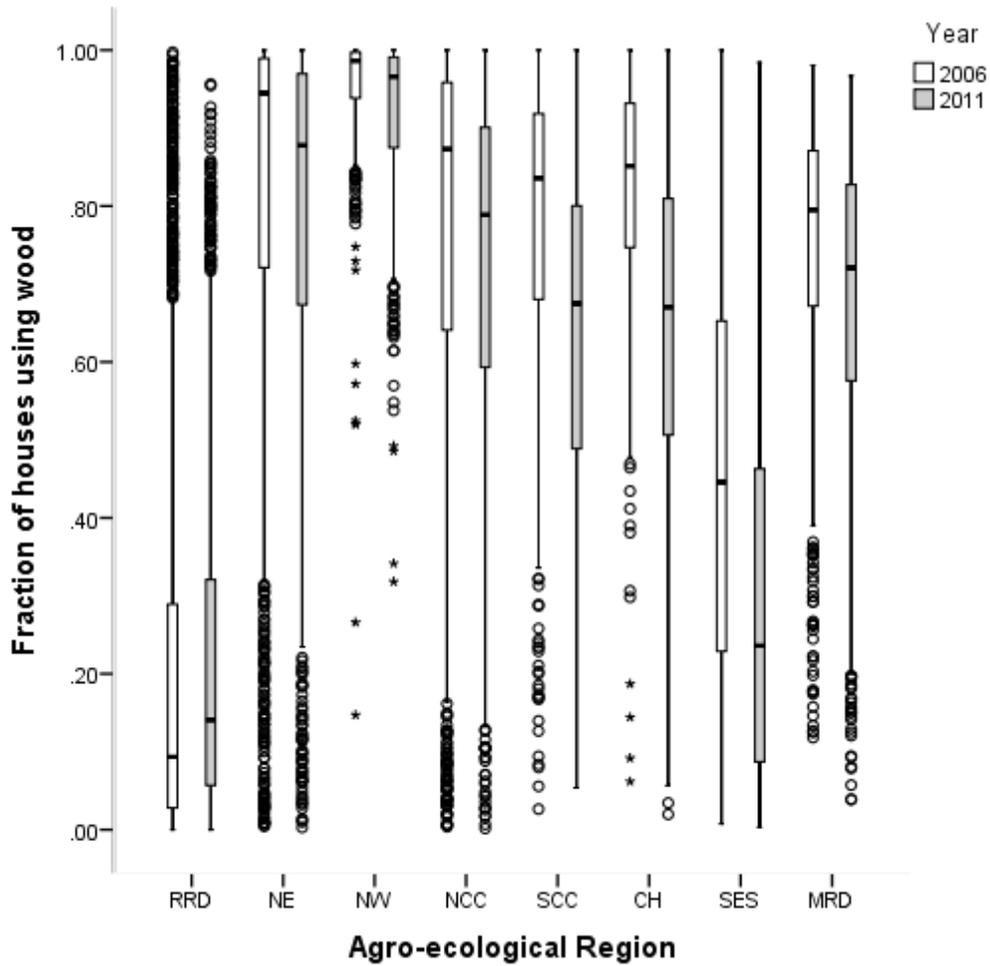
<sup>1</sup>Fraction of households; \* Transform of the type  $\log_{10}(1+x)$  was used; AIC = Akaike Information Criterion

**Table 6: Longitudinal Linear Mixed Model with Spatial Autocorrelation for Gas**

Parameter	Estimate	Standard error	Significance, p<
Intercept	0.067	0.038	0.077
Year	-0.014	0.002	0.001
Spatial autocorrelation	1.053	0.036	0.001
<i>Household level demand side parameters</i>			
Main income from agriculture <sup>1</sup>	-0.298	0.006	0.001
Motorbike ownership <sup>1</sup>	0.062	0.006	0.001
<i>Community level supply side parameters</i>			
Peri-urban class	0.036	0.002	0.001
Market present	0.013	0.002	0.001
Distance to national highway*	-0.006	0.001	0.001
Distance to town*	-0.008	0.001	0.001
<i>Agro-ecological zone</i>			
Red River Delta	0.039	0.006	0.001
North East	omitted		
North West	-0.013	0.008	0.101
North Central Coast	0.019	0.006	0.002
South Central Coast	0.029	0.007	0.001
Central Highlands	0.009	0.008	0.265
Southeast South	0.040	0.012	0.001
Mekong River Delta	0.048	0.008	0.001
<i>Model performance</i>			
Goodness of fit, r	0.970		0.01
AIC full model	-34524		
AIC without autocorrelation	-29842		
AIC null model	-27563		
Goodness of fit of OLS regression model, r	0.835		0.01

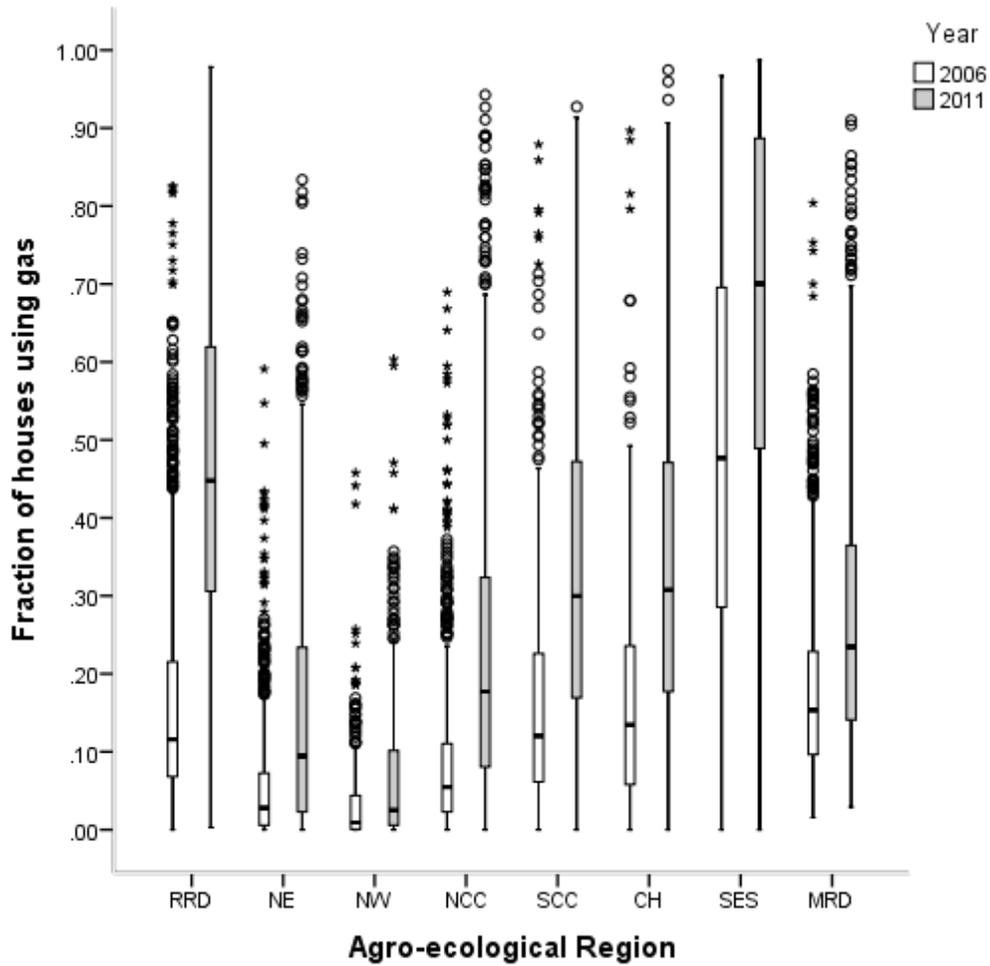
n = 17236

<sup>1</sup>Fraction of households; \* Transform of the type  $\log_{10}(1+x)$  was used; AIC = Akaike Information Criterion



**Figure 1: Variation of fuelwood use across agro-ecological zones and time**

RRD = Red River Delta, NE = North East, NW = North West, NCC = North Central Coast, SCC = South Central Coast, CH = Central Highlands, SES = South East South, MRD = Mekong River Delta



**Figure 2: Variation of gas use across agro-ecological zones and time**

RRD = Red River Delta, NE = North East, NW = North West, NCC = North Central Coast, SCC = South Central Coast, CH = Central Highlands, SES = South East South, MRD = Mekong River Delta

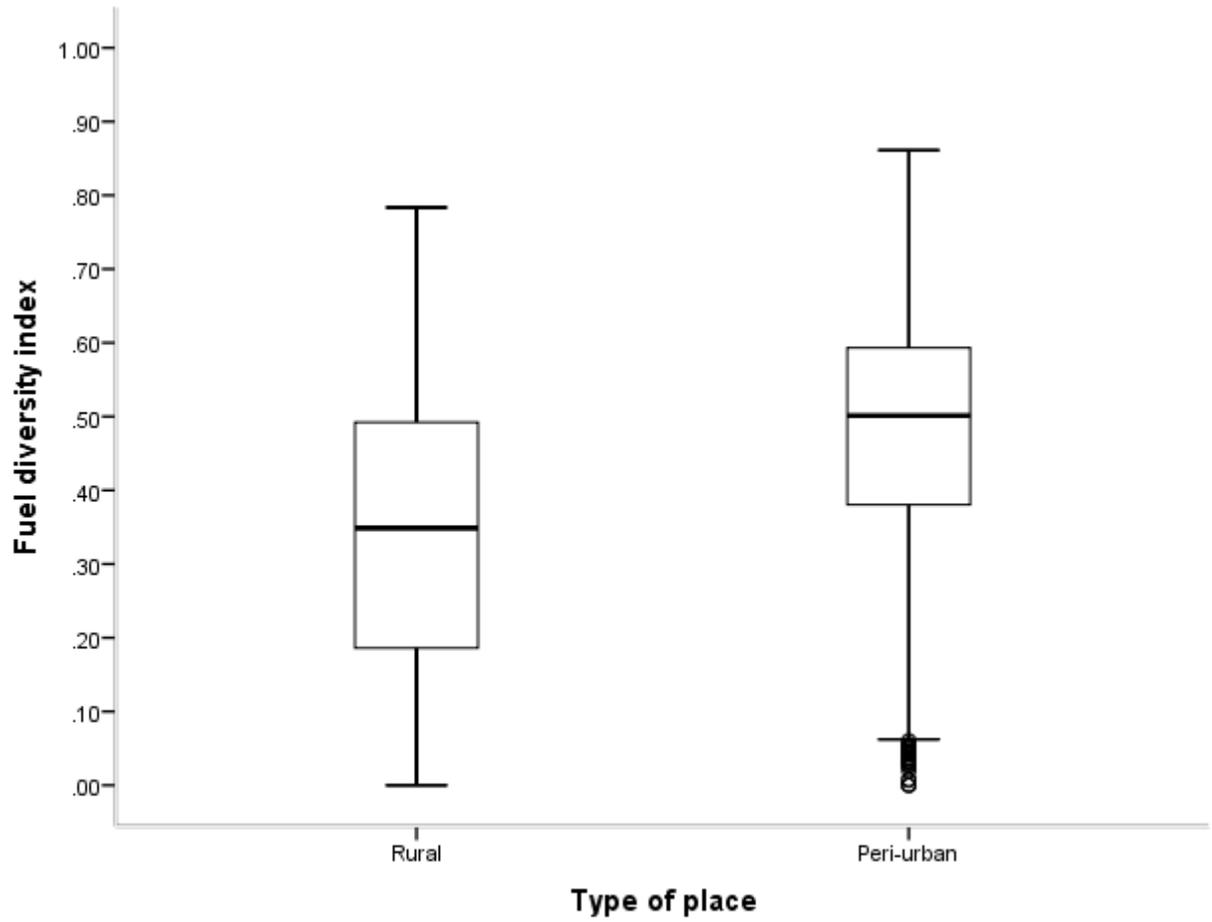


Figure 3: Cooking fuel diversity index across type of place