UNIVERSITY OF WAIKATO

Hamilton
New Zealand

Households' Energy Mix Selection in Pakistan

Muhammad Irfan, Michael P. Cameron and Gazi Hassan

Working Paper in Economics 28/17

December 2017

Corresponding Author

Muhammad Irfan
School of Accounting, Finance and Economics
University of Waikato
Private Bag 3105
Hamilton
NEW ZEALAND, 3240
Email: irfaneconomist786@gmail.com

Michael P. Cameron
michael.cameron@waikato.ac.nz

Gazi Hassan
gazi.hassan@waikato.ac.nz

Abstract

Despite the adverse effects of biomass fuels on health and the environment, the use of solid fuels at the household level for cooking, lighting and heating purposes is very common in developing countries such as Pakistan. Globally almost 3 billion people depend on traditional or conventional solid energy sources for cooking. These solid fuels are a major cause of indoor air pollution and can severely damage health and the environment, so there is a need to better understand the factors that lead to the consumption of solid fuels. This study analyzes data from the Pakistan Social and Living Standards Measurement (PSLM) Survey 2013-14 to establish the non-price factors associated with the fuel mix selection of households. A novel aspect of the study is that, rather than treating fuel choices as independent, we first group household fuel mix choices into categories using cluster analysis. We then apply multinomial logit models to investigate the factors associated with households' fuel mix selection. We find that income, education, agricultural occupation and urban location are strongest factors associated with the selection of a mix of fuels that is substantially made up of clean fuels, while agricultural occupation, large family size, and having cattle are associated with fuel mixes that are more heavily based on solid fuels. Moreover, we show that income growth is unlikely to lead to substantial uptake of cleaner fuels in rural areas. Our results suggest that the government, if concerned about indoor air pollution, should rapidly increase the availability of natural gas and electricity connections to support a shift to cleaner fuel mixes.

Keywords

fuel choice
energy mix
Pakistan

JEL Classification

D71; P18; P28

1. INTRODUCTION

Energy and health are both very important for any household, and are key resources for economic growth and development. However, globally 3 billion people depend upon solid fuels such as coal, charcoal, firewood, animal dung, and crop residues[1] for heating and cooking purposes (Landrigan *et al.* 2017). The combustion of these solid fuels emits a multitude of complex chemicals including carbon monoxide, nitrogen dioxide, polycyclic aromatic hydrocarbons (PAH), formaldehyde, and others inhalable particulates, damaging people's health and the environment (Cooper 1980 andTorres-Duque, Maldonado, Pérez-Padilla, Ezzati, and Viegi 2008). As a result of solid fuel use, almost 1.6 million people around the world die prematurely each year due to indoor air pollution, and millions of people face serious diseases such as asthma, lung infections, eye infections, sinus problems, tuberculosis (TB), and cardiovascular diseases (Mishra 2003, Kim, Jahan and Kabir 2011, Kim *et al.* 2011, Lakshmi *et al.* 2012 and Sehgal, Rizwan and Krishnan 2014). The number of annual deaths attributed to acute respiratory infections (ARI) among children under age five in Pakistan has been estimated to be 51,760, with a further 18,980 annual deaths due to chronic obstructive pulmonary disease (Colbeck, Nasir and Ali 2010).[2]

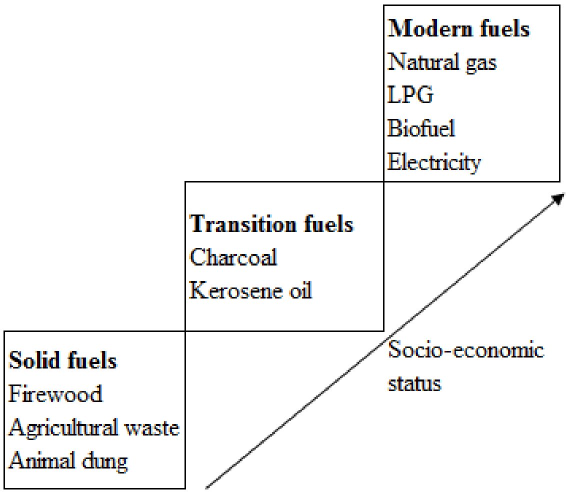
The trade-offs associated with solid fuel as a source of energy not only occur for human health, but also for the environment. The forests of developing countries are progressively depleting due to wood usage as a household cooking fuel (Arnold, Köhlin and Persson 2006 and Bhatt and Sachan 2004). Forests are necessary for economic, ecological, social, environmental, and health benefits, and provide food, medicines, forest products, and social resources, as well as helping to reduce global warming (Bonan 2008).

Despite the adverse effects of biomass fuels on health and the environment, the use of solid fuels for cooking, lighting and heating purposes remains very common in developing and middle-income countries. Household fuel selection is associated with many socio-economic factors. Household income is one important factor. The energy ladder model contends that households will switch from biomass to modern fuels such as natural gas and electricity as their income (or socio-economic status) rises (see Figure 1). The energy ladder model shows a three-stage fuel switching process. In the first stage, households use traditional solid fuel sources, such as agricultural waste, animal waste, and firewood. As their socio-economic status improves, households move upwards along the energy ladder, and use somewhat-cleaner fuels such as charcoal, kerosene, and coal. At the highest level of the energy ladder, households switch to using advanced 'clean' fuels like natural gas, LPG, biofuels, and electricity (Hosier and Dowd 1987 and Leach 1992).

[1] These residues include cotton sticks, bagasse, husks, wheat straw, roots, corn stalks, stubble, leaves and seed pods, etc.

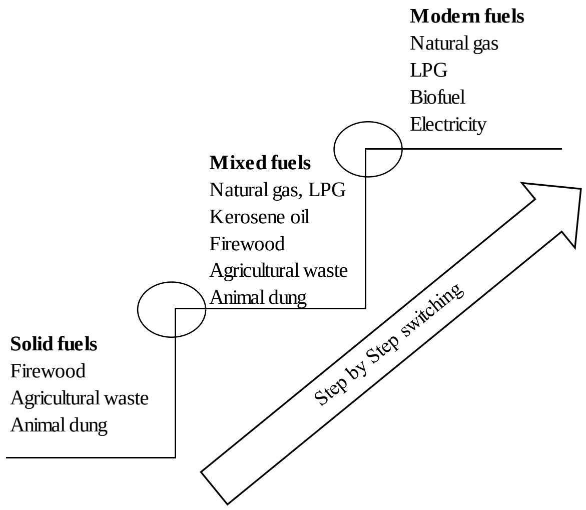
[2] http://www.who.int/indoorair/publications/indoor_air_national_burden_estimate_revised.pdf?ua=1

Figure 1: Energy Ladder



However, for most household's energy switching does not occur on a series of simple discrete steps as suggested by the linear energy ladder model (Campbell *et al.* 2003) . Instead, use of multiple fuels concurrently is common. Moreover, as household income increases, many households continue to use some amount of the fuels from the lower steps on the ladder. This is referred to as fuel stacking or energy stacking (see Figure 2). The energy transition shown in Figure 2 is a bi-directional process, as users can go up or down the ladder, while some continue to use traditional fuels alongside more advanced fuels. However, once most households achieve the highest socio-economic status they only use modern fuels such as natural gas, LPG, biofuel, and electricity (Campbell *et al.* 2003, Heltberg 2004 and Pachauri and Jiang 2008).

Figure 2: Energy Stacking



In Pakistan, like many other middle-income countries, evidence of fuel stacking is observable as electricity, firewood, natural gas, crop residues, animal dung, and Liquefied Petroleum Gas (LPG) are the main fuels used for cooking, heating and lighting. In contrast, electricity is used for lighting. In rural areas, the consumption of solid fuels such as firewood, dry animal dung, and crop residues is higher than in urban areas. On the other hand, the consumption of clean energy sources such as natural gas is higher in urban areas than rural areas. Accessibility to fuels may be one of the main causes of consuming any specific fuel.

The objective of this study is to identify the non-price factors associated with fuel selection in Pakistan. A novel aspect of the study is that, rather than treating fuel choices as independent, we first group household fuel mix choices into categories using cluster analysis. Cluster analysis groups households by their choices of fuels; hence the clusters will represent the different fuel mix options selected by the households in the sample. This data-driven approach to defining household fuel mix selection recognizes that the decision to use a particular fuel type is not made independent of the other fuels that households already use. Moreover, our approach makes optimal use of the actual fuel mixes observed in the dataset, and avoids arbitrary decisions about which fuels make up a given fuel mix. This cluster analysis approach distinguishes our study from previous studies on fuel selection, where fuel selections are either treated independently (Osiolo 2009, Farsi, Filippini and Pachauri 2007, Ouedraogo 2006 and Pundo and Fraser 2006) or where fuel mixes are arbitrarily determined by the researchers (Lee 2013, Narasimha Rao and Reddy 2007 and Heltberg 2005).

Our approach is more appropriate for designing policy to encourage the use of cleaner fuels and discourage the use of dirty fuels, because it better reflects actual fuel mix decisions of households, who likely do not make decisions about use of each fuel independently of their use of other fuels, and most often use a mix of fuels for cooking and heating (World energy outlook, 2006)[3]. The remainder of the paper is organized as follows. Section 2 discusses relevant literature, and in Section 3 we discuss the data and methodology. Section 4 presents the results, Section 5 summarizes our findings and concludes, and suggests some policy implications.

2. LITERATURE REVIEW

Many studies have investigated the determinants of households' fuel selection in developing countries. However, few of these studies investigate households' selection of *fuel mix*, and those that do use arbitrarily determined fuel mixes (that is, where the fuel mixes to be investigated are identified *a priori* by the researchers). Those research studies that investigate fuel combinations at the household level have been limited to descriptive analyses of the fuel combinations. Overall, the extant literature on the determinants of fuel selection mainly focuses on households' use of individual fuels, often their most commonly or extensively used fuel.

[3] https://www.iea.org/publications/freepublications/publication/cooking.pdf

These studies identify many socio-economic and demographic variables that are associated with the selection of household fuels including income, education, gender, household location (rural or urban), family size, land holding, and livestock holding.

Most studies have investigated the factors associated with household's fuel selection by only considering individual fuels, such as Farsi *et al.* (2007), who investigated the factors associated with fuel choice (firewood, kerosene, and LPG) for cooking in urban India. Based on the energy ladder as a theoretical framework, they applied an ordered probit model. They found that income and education were positively related to the use of LPG and that female-headed households were also more likely to adopt LPG. Similarly, Ouedraogo (2006) studied the determinants of fuel choice in Burkina Faso. A multinomial logit model was applied and they concluded that lower income was a significant constraint for the adoption of LPG, in comparison to firewood. In rural Kenya, Pundo and Fraser (2006) found that education of the wives of household heads played a vital role in fuel choice; as education of the wives increases the use of fuelwood decreases and the use of comparatively cleaner fuel like kerosene increases.

In Kenya, Osiolo (2009) examined fuel selection among firewood, charcoal, kerosene, LPG, and electricity. Total expenditure (a proxy of income) and education were positively associated with the adoption of kerosene, LPG, and electricity. On the other hand, they found that kerosene and LPG were less likely to be used in rural areas. Likewise, Jumbe and Angelsen (2011) applied a multinomial probit model on data from 404 households in 31 villages surrounding two forests in Malawi. They found that the distance to the firewood source, firewood species, and area of the firewood source were important determinants of selecting firewood as a fuel source. In Madhya Pradesh state in India, Sehjpal, Ramji, Soni, and Kumar (2014) collected data from 200 rural households and applied binary logit models to investigate households' selection between traditional and modern fuels. They found that if women got engaged in income generating activities, the chances of selecting cleaner fuels would increase. Moreover, other factors such as education, price of the fuels, and the availability of electricity connections were also important factors in fuel selection. Similarly, Rahut, Das, De Groote, and Behera (2014) applied a multinomial logit model on 2007 Bhutan Living Standard Survey data and found that female-headed households were more likely to adopt clean fuels such as LPG and electricity. In another study conducted in Bhutan, Rahut *et al.* (2014) found that higher education, higher income and female headed households were more likely to adopt cleaner fuels such as LPG and electricity.

Unlike the choice among individual fuels, fuel combinations are rarely examined in the literature. Studies using fuel combinations have generally not applied any statistical or econometric techniques to form the fuel mix combinations. For example, Heltberg (2005) made the fuel combinations on his own (LPG only; wood only; LPG and wood; and LPG and charcoal) and using data from Guatemala he found that the prices of fuels play a vital role. Especially the price of solid fuels such as firewood significantly affects the quantity demanded. Furthermore, he found that households with higher education tend to consume cleaner fuels such as LPG and electricity. Similarly, Narasimha Rao and Reddy (2007) investigated the factors associated with fuel selection in India. They found that the education of the household head and household income played significant roles in the selection of cleaner fuels, such as biofuels and LPG.

Lee (2013) found that education and income are the key factors associated with fuel consumption in Uganda. They found that, as income and education increases the consumption of solid fuels decreases. Similarly, Jan, Khan, and Hayat (2012) explored the determinants of rural household energy choice in Pakistan by collecting data from 100 randomly selected households. They found that income was not the only factor associated with the cleaner fuel selection, the preference of the consumer and access to the alternative sources also played important role in fuel selection. In another study, Nasir, Murtaza, and Colbeck (2015) examined fuel choice in Pakistan and found that household location (urban or rural), availability of natural gas and electricity, and poverty were the main factors associated with fuel selection, while poverty was the main hindrance to the selection of clean fuels such as natural gas (piped gas) and LPG.

Some other studies have also used descriptive statistics to form fuel mixes, such as Brouwer and Falcão (2004) in Mozambique, Joon, Chandra, and Bhattacharya (2009) in India, Miah, Foysal, Koike, and Kobayashi (2011) in Bangladesh, and Peng, Hisham, and Pan (2010) in Hubei, China. Most of the findings are similar to those previously cited above, where households are more likely to use cleaner fuel mixes if they have higher income, more education, and better access to the cleaner fuels. Across most studies, income has consistently been an important factor associated with modern and cleaner fuel use, as predicted by the energy ladder model. However, to date no study has been specific about the levels of income that would induce households to select cleaner fuels or fuel mixes.

The extant literature suffers from some substantial shortcomings. Studies that use logistic models to investigate the odds of a household using individual fuels rely on the assumption that each household makes its decision about whether to use a given fuel or not independently of whether they are also using other fuels. A multinomial logit model exacerbates this problem, because it further assumes that the use of fuels are mutually exclusive. This may be appropriate in the context of determining the 'main' fuel used by households, but in so doing a great deal of the nuance of households' fuel mix choices is lost. Households rarely rely on a single fuel and, as demonstrated in this paper, often use many fuels in addition to their 'main' fuel. Investigating the *fuel mix* choice of households is therefore preferable. However, extant studies that have looked at fuel mix selection have done so using fuel mixes that were arbitrarily determined by the researchers (for example, Heltberg 2005, Lee 2013 and Narasimha Rao and Reddy 2007), and therefore often suffer from the same problems of independence and mutual exclusion noted above.

To avoid these problems, in this paper we adopt a data-driven approach to identifying the household fuel mixes actually represented within our sample. To achieve this, we use cluster analysis to determine the fuel mix selections of households. Cluster analysis allows us to identify mutually exclusive fuel mixes that households select, based on the observed mixes of fuels that households in our sample consume. To our knowledge, this is the first study to derive household fuel mixes and investigate the factors associated with them in this way.

3. DATA AND VARIABLES

3.1 Data

Data from the Pakistan Social and Living Standards Measurement Survey 2013-14 (PSLM) was used for this study. The Federal Bureau of Statistics (FBS) developed the data collection frame for the PSLM. Each city was divided into enumeration blocks consisting of 200-250 households. Each enumeration block was then classified into three strata based on household incomes, i.e. low, medium, and high. A two-stage stratified sample design was adopted to collect the data. Each primary sampling unit (PSU) from a stratum was selected through a probability proportional to size (PPS) method, and within each PSU 12 rural and 16 urban households were selected. Initially, 19,620 households from 1368 PSUs were selected. However, due to ongoing conflict in some areas 61 PSUs were dropped and finally 17,989 households were interviewed from 1307 PSUs. Thus, the data can be considered to be reasonably representative of households in both rural and urban areas in Pakistan.

Generally, households in Pakistan use natural gas, LPG, firewood, agricultural waste, animal dung, and kerosene oil at household level for cooking and heating purposes. The mean consumption of the fuels at household level in Pakistan and mean expenditures on these fuels are shown in Table 1. Households are spending the greatest proportion of their energy budget on natural gas and the least proportion of their energy budget on kerosene oil.

Table 1: Mean Consumption and Expenditures of the Household

Fuels	Acronyms	Mean Consumption	Mean Expenditures (PKR)
Natural gas (MMBTU)	ng	0.709	197.88
LPG (Kg)	lpg	0.548	79.108
Firewood (Kg)	fw	53.925	43.622
Agricultural waste (Kg)	aw	29.574	141.867
Animal dung (Kg)	ad	27.542	105.719
Kerosene oil (Litre)	ko	0.124	9.154

Note: 1 USD = 100 PKR, 2014

3.2 Variables of the Models

Table 2 summarizes the key independent variables in the sample, with mean values weighted to account for the stratified nature of the sample. Urban households (36.4 percent of the weighted sample) have greater accessibility to natural gas, and would therefore be expected to be more likely to use that fuel source. Female-headed households (3.7 percent of the weighted sample) have been shown in the literature to be more likely to use cleaner fuels than male-headed households. Agricultural households (24.3 percent of the weighted sample included at least once household member with an agricultural occupation, and 7.1 percent of the weighted sample had cattle) can be expected to be more likely to use agricultural waste or animal dung as fuel sources. Annual expenditure is used as a proxy of income due to substantial missing income data, and was linearized through taking its natural log. The energy ladder suggests that households with higher income are expected to be more likely to use cleaner fuels.

Table 2: Description of the Variables

Variables	Description	Mean
Urban	Urban =1; Rural =0	0.364
Age	Age of the household head in years	43.95
Gender	Household head gender, Male=1; Female=0	0.963
Household size	Number of family members in the household	6.345
Education	Number of schooling years of household head	4.964
Agri. Occupation	Any member of the household's occupation was agricultural in last year=1; otherwise=0	0.243
Cattle	Household has one or more cattle=1; otherwise=0	0.071
No. of rooms	Number of rooms in the household's dwelling	2.287
Elt. connection	Household has electricity connection=1; otherwise=0	0.915
Ln of Expenditures	Natural log of total yearly household expenditure	12.25

Some other variables such as cooking habits, taste preferences, and other cultural factors could also affect household fuel selection. However, these are difficult to measure quantitatively and our data set did not had these variables. Secondly, the market price of the fuels was not available in the dataset, and therefore could not be included. However, as noted by Irfan, Cameron, and Hassan (2018), there is little cross-sectional variation in fuel prices in Pakistan and thus, even if price data were available, it is unlikely that it would have been able to be included in our cross-sectional model.

3.3 Cluster Analysis

Before applying the fuel choice model, it is essential to create groups of households that use similar fuel mixes according to the actual fuel consumption of households in the sample. Cluster analysis is an appropriate technique for recognizing groups with similar attributes. We have a large dataset and therefore partitioning is the most suitable method to create the clusters. *K*-means cluster analysis aims at dividing the data into different segments in such a way that within cluster variation is minimised. The clustering/segmenting procedure starts by randomly allocating entities to a number of clusters; then the entities are reallocated to other clusters to decrease the variation within cluster, which is measured as the squared distance from each observation to the centre of the related cluster (Romesburg 2004).

Initially, we normalized our fuel variables (natural gas [ng], liquefied petroleum gas [lpg], firewood [fw], agricultural waste [aw], animal dung [ad], and kerosene [ko]) to avoid scaling problems (Scott and Knott 1974). The optimal number of clusters was determined by considering the Calinski/Harabasz pseudo-F (C-H F) values for different numbers of clusters (Caliński and Harabasz, 1974). The C-H F values were 3501 for eight clusters, 3536 for nine clusters, and 3328 for ten clusters, suggesting that nine clusters was the optimal solution. Each household was then allocated to one of the seven clusters. To achieve this, first each cluster's geometric centre (that is, its centroid) was calculated, by calculating the mean values of the households contained in the clusters regarding given variables (ng, lpg, fw, aw, ad, and ko). Then the distances from each household to the newly located cluster centres were calculated and households were again allocated to a specific cluster on the basis of their least distance to other cluster centres. This process iterated until the sum of the squared Euclidean distances was minimised (Romesburg 2004). We then merged the two clusters that had the smallest numbers of households (467 and 97 respectively) into their nearest neighbours, in order to avoid problems for the analysis related to having small cell sizes in the multinomial logit model. This was preferred to reducing the number of clusters directly in the cluster analyses, where smaller numbers of clusters still led to some clusters with only a small number of group members.

4. METHODS

4.1 Multinomial Logit Model

The Multinomial Logit Model (MLM) shows the behavior of consumers with a common consumption objective when they are faced with the choice between many mutually exclusive options. In our case, this is the choice between consuming different fuel mixes (represented by the fuel mix clusters determined using the method described previously). The MLM is based on the random utility model. Individuals make decisions by comparing the levels of utility associated with each possible alternative. In classical demand theory the problem of consumer choice is usually described as a problem of utility maximization under a limited budget, with a utility function characterizing the consumer's preferences for consuming varying amounts of each type of commodities.

The fuel mix selection model is based on the rule that a household selects that fuel mix that maximizes their utility. Let a household p from n total households in the sample select a fuel type j from m mutually exclusive fuel mixes (clusters). The utility function U_p of a fuel mix type X_j can be written as:

$$U_p = f(X_j) + e_{jp} \quad (1)$$

where:

$j=1, 2, 3, \ldots\ldots\ldots, m$
$p= 1, 2, 3, \ldots\ldots\ldots\ldots, n$

and e_{jp} is the error term following an i.i.d extreme valued distribution. The CDF of each error term is given by $[F(e_{jp}) = \exp\{-e^{-ejp}\}]$.

Finally, we have:

$$\Pr[Cl = \mathrm{j}] = \frac{\exp^{\beta_j X_i'}}{1 + \sum_{j=0}^{m} \exp^{\beta_j X_i'}} \quad (2)$$

where:

$\Pr[Cl = j]$ is the probability of choosing fuel mix j, with one of the fuel mixes as a reference category.

j = number of fuel mixes (total seven) in the choice set.

j = 0 for the reference fuel mix.

X_i = explanatory variables.

β_j = vector of the estimated parameters (so that β_j shows the effect of X_j on the likelihood of choosing *jth* fuel mix).

The models are weighted to account for the stratified nature of the sample.

5. RESULTS AND DISCUSSION

5.1 Cluster Analysis Results

Figure 3 depicts the fuel combinations in each fuel mix cluster, and Table 3 shows number of households in, and the fuel mix of, each cluster. For the sake of simplicity, in Figure 3 we only show the fuels that make up more than one percent of the fuel use of households in each fuel mix cluster, and the name of each cluster does not reflect fuels that make up less than five percent of fuel use within that cluster. Cluster 2 (awadfw), Cluster 3 (fwadaw), and Cluster 7 (adfwaw) represent fuel mixes that are predominantly based on solid fuels (agricultural waste; firewood; and animal dung, respectively). We take these fuel mix clusters as the reference categories in the MLM models that follow. Cluster 4 (ngfw) is the fuel mix that has the greatest proportion of clean fuels, with 82 percent natural gas consumption. We considered Cluster 4 as a base category for our fourth MLM model. These different reference categories were taken to explore in detail the factors associated with choosing a predominantly solid fuel mix (Clusters 2, 3, and 7), as opposed to a fuel mix based on cleaner fuels (Cluster 4).

Figure 3: Fuel Cluster

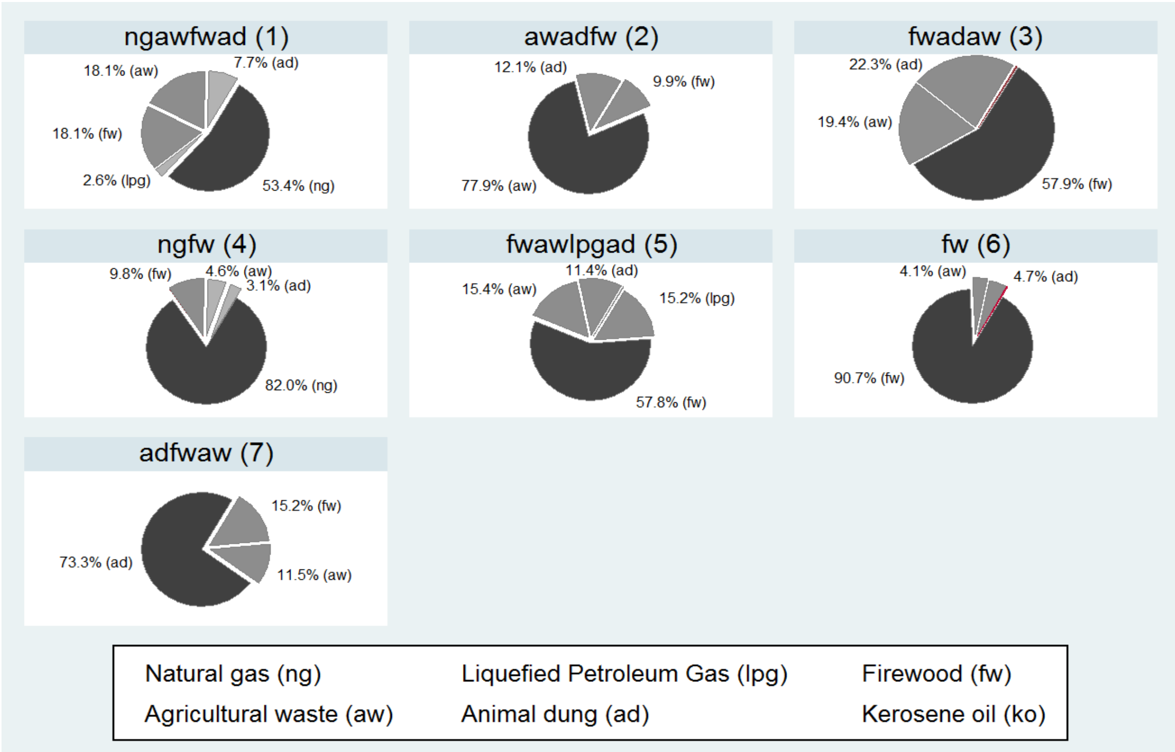


Table 3 Description of the Clusters

Cluster Number	Cluster Name	Number of Households	Description of the Cluster
1	ngawfwad	5881	Natural gas, agri. Waste, firewood, animal dung
2	awadfw	1626	Agri. Waste, animal dung, firewood
3	fwadaw	6037	Firewood, animal dung, agricultural waste
4	ngfw	1052	Natural gas, firewood, agri. waste, animal dung
5	fwawlpgad	494	Firewood, agri. waste, LPG, animal dung
6	fw	1872	Firewood, animal dung, agri. waste, kerosene oil
7	adfwaw	1027	Animal dung, firewood, agri. waste
Total		17989	

5.2 Multinomial Logit Model Results

We ran four multinomial logit models, using fuel mix clusters 2, 3, 4, and 7 as reference categories. The resulting model results (expressed as relative-risk ratios, or exponentiated coefficients) are shown in Table 4, with each column representing the results using a different fuel mix cluster as the reference category. The reference categories in first three models are fuel mixes that are predominantly solid fuels, while in the fourth model the reference category is a fuel mix that is predominantly natural gas (a cleaner fuel). Given the extensive nature of the results in Table 4, and the policy imperative to reduce use of solid fuels in favour of cleaner fuels, we concentrate our discussion here on variables that show a consistent association with a greater preference for cleaner fuels (as represented by fuel mix cluster 4), and a significantly lower preference for solid fuels (as represented by fuel mix clusters 2, 3, and 7).

The results show that households from urban areas are more likely to adopt the cleaner fuel cluster. In other words, urban households are more likely to adopt the cleaner fuel cluster (ngfw) than any of the solid fuel clusters (awadfw, fwadaw, or adfwaw). As noted above, this is likely due to the availability or accessibility of natural gas connections in the urban areas, and the lack of such connections in rural areas. Interestingly, based on the size of coefficients this variable appears to be the most influential factor associated with the use of clean fuels. Similarly, electricity connections (also more likely to be associated with urban households) are associated with with a greater likelihood of choosing cleaner fuels than two of the solid fuel clusters (the exception being Cluster 7, adfwaw) where the relative risk ratio is of the same sign but is statistically insignificant.

In contrast, agricultural households are much more likely to adopt solid fuel mixes, based on the results for the agricultural occupation and cattle variables. Both variables are negatively associated with the choice of Cluster 4 (ngfw), and positively associated with the solid fuel clusters (awadfw, fwadaw, or adfwaw). The free availability of agricultural waste and animal dung likely lead agricultural households to be more likely to make use of these fuel sources rather than cleaner fuels. Larger households are less likely to adopt the cleaner fuel mix. This may also relate to agricultural households and relative labour scarcity. The free availability of fuel collecting labor may make households more likely to adopt solid fuels rather than financially-costly cleaner fuels.

In contrast, demographic factors associated with the household head had a weaker relationship with fuel mix choice. Age of the household head has a small effect on fuel mix choice, with older household heads more likely to choose Cluster 2 (awadfw), and marginally more likely to choose Cluster 3 (fwadaw) than the clean fuel cluster (ngfw), but Cluster 7 (adfwaw) was not statistically significantly more likely to be selected than the clean fuel cluster. In contrast to much of the previous literature, female-headed households were mostly not significantly more likely to choose cleaner fuels than solid fuels.

Unlike other demographic factors, the education of the household head was a strong factor associated with the selection of cleaner fuel mix rather than the solid fuel mixes. This could be for a combination of two main reasons. First, education increases the opportunity cost of time spend collecting solid fuels such as firewood or agricultural waste. Second, greater education bring awareness about risks associated with indoor air pollution, so more educated households may be considering health benefits and thus avoid the use of solid fuels.

Households with a greater number of rooms in the dwelling were more likely to adopt the cleaner fuel mix and less likely to use solid fuels. This likely demonstrates a wealth effect since a greater number of rooms is associated with a larger house and more wealth. Moreover, larger houses require more heating than smaller households, and so modern fuels will be more efficient in heating these larger spaces.

Finally, total expenditures were used as a proxy of income, and we found that this is significant and positively associated with choosing the cleaner fuel mix. Higher income households are more likely to be able to afford cleaner fuels such as natural gas and LPG, which are comparatively expensive, especially when compared with firewood and agricultural waste, which can often be collected at no financial cost to the household. These results support the phenomenon of fuel stacking, i.e. that as income increases, households tend to move towards the use of modern cleaner fuels.

Table 4: Multinomial Logit Model Results

	Model 1 Base 2 (awadfw)	Model 2 Base 3 (fwadaw)	Model 3 Base 7 (adfwaw)	Model 4 Base 4 (ngfw)
ngawfwad (Cluster 1)				
Urban	19.07*** (3.935)	13.29*** (1.909)	20.39*** (5.263)	1.285 (0.217)
Age	1.008** (0.002)	1.013*** (0.002)	1.006* (0.002)	1.001 (0.003)
Male head	0.655 (0.160)	0.888 (0.146)	0.884 (0.231)	1.321 (0.375)
Family size	0.806*** (0.015)	0.888*** (0.013)	0.834*** (0.015)	0.957* (0.017)
Education	1.081*** (0.010)	1.068*** (0.008)	1.085*** (0.010)	0.999 (0.009)
Agri. Occupation	0.160*** (0.020)	0.293*** (0.0319)	0.137*** (0.020)	0.932 (0.163)
Cattle	0.157*** (0.025)	0.379*** (0.053)	0.113*** (0.019)	1.452 (0.477)
Rooms	1.077 (0.043)	1.076* (0.034)	1.138** (0.054)	0.905** (0.032)
Elt. Connection	1.596* (0.348)	1.811** (0.329)	0.613 (0.157)	0.565* (0.140)
Ln of expenditures	4.413*** (0.574)	3.073*** (0.322)	2.091*** (0.280)	0.174*** (0.026)

awadfw (Cluster 2)

Urban	Base category	0.697 (0.132)	1.069 (0.313)	0.0674*** (0.016)
Age	Base category	1.004 (0.002)	0.998 (0.002)	0.992* (0.003)
Male head	Base category	1.355 (0.276)	1.349 (0.372)	2.016* (0.691)
Family size	Base category	1.101*** (0.016)	1.034 (0.018)	1.188*** (0.027)
Education	Base category	0.988 (0.007)	1.004 (0.010)	0.924*** (0.010)
Agri. Occupation	Base category	1.830*** (0.175)	0.852 (0.110)	5.819*** (1.137)
Cattle	Base category	2.406*** (0.291)	0.720* (0.111)	9.219*** (3.222)
Rooms	Base category	0.999 (0.035)	1.056 (0.050)	0.840*** (0.040)
Elt. Connection	Base category	1.135 (0.174)	0.384*** (0.086)	0.354*** (0.108)
Ln of expenditures	Base category	0.696*** (0.059)	0.474*** (0.058)	0.0395*** (0.007)

fwadaw (Cluster 3)

Urban	1.435 (0.271)	Base category	1.534 (0.373)	0.0967*** (0.019)
Age	0.996 (0.002)	Base category	0.993* (0.002)	0.988*** (0.003)
Male head	0.738 (0.150)	Base category	0.995 (0.216)	1.488 (0.443)
Family size	0.908*** (0.013)	Base category	0.939*** (0.015)	1.079*** (0.022)
Education	1.012 (0.007)	Base category	1.016 (0.008)	0.936*** (0.009)
Agri. Occupation	0.546*** (0.0523)	Base category	0.466*** (0.050)	3.180*** (0.571)
Cattle	0.416*** (0.050)	Base category	0.299*** (0.040)	3.832*** (1.322)
Rooms	1.001 (0.035)	Base category	1.057 (0.044)	0.841*** (0.035)
Elt. Connection	0.881 (0.134)	Base category	0.338*** (0.066)	0.312*** (0.088)
Ln of expenditures	1.436*** (0.122)	Base category	0.680*** (0.075)	0.0567*** (0.009)

ngfw (Cluster 4)

Urban	14.85*** (3.623)	10.35*** (2.054)	15.87*** (4.630)	Base category
Age	1.008* (0.003)	1.012*** (0.003)	1.005 (0.003)	Base category
Male head	0.496* (0.170)	0.672 (0.200)	0.669 (0.240)	Base category
Family size	0.842*** (0.019)	0.927*** (0.018)	0.871*** (0.020)	Base category
Education	1.082*** (0.012)	1.069*** (0.011)	1.086*** (0.012)	Base category
Agri. Occupation	0.172*** (0.033)	0.314*** (0.056)	0.146*** (0.028)	Base category
Cattle	0.108*** (0.037)	0.261*** (0.090)	0.0781*** (0.027)	Base category
Rooms	1.191*** (0.056)	1.189*** (0.050)	1.257*** (0.065)	Base category
Elt. Connection	2.824*** (0.865)	3.205*** (0.913)	1.085 (0.357)	Base category
Ln of expenditures	25.30*** (4.657)	17.62*** (2.968)	11.99*** (2.223)	Base category

fwawlpgad (Cluster 5)

Urban	3.480*** (0.835)	2.425*** (0.442)	3.720*** (1.009)	0.234*** (0.059)
Age	1.001 (0.005)	1.006 (0.005)	0.999 (0.005)	0.994 (-0.005)
Male head	0.354** (0.119)	0.479* (0.144)	0.477* (0.173)	0.713 (0.250)
Family size	0.759*** (0.021)	0.836*** (0.021)	0.785*** (0.022)	0.902*** (0.027)
Education	1.090*** (0.013)	1.077*** (0.012)	1.094*** (0.013)	1.008 (0.012)
Agri. Occupation	0.388*** (0.0709)	0.710* (0.118)	0.331*** (0.063)	2.257*** (0.491)
Cattle	0.235*** (0.075)	0.565 (0.177)	0.169*** (0.055)	2.164 (0.918)
Rooms	1.295*** (0.059)	1.293*** (0.052)	1.367*** (0.074)	1.088 (0.049)
Elt. Connection	1.622 (0.722)	1.841 (0.814)	0.623 (0.294)	0.574 (0.302)
Ln of expenditures	15.10*** (2.466)	10.52*** (1.505)	7.155*** (1.180)	0.597* (0.120)

fw (Cluster 6)

Urban	1.441 (0.375)	1.004 (0.182)	1.54 (0.440)	0.0970*** (0.024)
Age	0.997 (0.002)	1.001 (0.002)	0.994 (0.003)	0.989** (0.003)
Male head	0.422*** (0.107)	0.572** (0.117)	0.569* (0.154)	0.851 (0.285)
Family size	0.972 (0.016)	1.071*** (0.015)	1.006 (0.018)	1.155*** (0.024)
Education	1.003 (0.011)	0.99 (0.009)	1.006 (0.011)	0.927*** (0.011)
Agri. Occupation	0.895 (0.110)	1.638*** (0.159)	0.763* (0.103)	5.209*** (1.009)
Cattle	0.463*** (0.074)	1.115 (0.159)	0.333*** (0.055)	4.272*** (1.501)
Rooms	1.076 (0.044)	1.074* (0.038)	1.136* (0.057)	0.903* (0.042)
Elt. Connection	0.337*** (0.066)	0.382*** (0.058)	0.129*** (0.029)	0.119*** (0.036)
Ln of expenditures	3.483*** (0.435)	2.426*** (0.261)	1.650*** (0.234)	0.138*** (0.025)

adfwaw (Cluster 7)

Urban	0.935 (0.274)	0.652 (0.158)	Base category	0.0630*** (0.0184)
Age	1.003 (0.002)	1.007* (0.002)	Base category	0.995 (0.003)
Male head	0.741 (0.205)	1.005 (0.218)	Base category	1.495 (0.535)
Family size	0.967 (0.017)	1.065*** (0.017)	Base category	1.149*** (0.026)
Education	0.997 (0.010)	0.984 (0.008)	Base category	0.921*** (0.010)
Agri. Occupation	1.173 (0.152)	2.147*** (0.234)	Base category	6.827*** (1.337)
Cattle	1.390* (0.215)	3.343*** (0.454)	Base category	12.81*** (4.522)
Rooms	0.947 (0.045)	0.946 (0.040)	Base category	0.795*** (0.041)
Elt. Connection	2.603*** (0.587)	2.955*** (0.576)	Base category	0.922 (0.303)
Ln of expenditures	2.111*** (0.263)	1.470*** (0.163)	Base category	0.0834*** (0.015)
N	17989	17989	17989	17989

Note: Standard errors in parentheses, * p<0.05, ** p<0.01, *** p<0.001

The importance of income raises the question of whether countries such as Pakistan can simply grow out of using solid fuels. That is, as incomes rise through economic growth, how rapidly will households shift up the energy ladder and adopt cleaner fuels? To investigate this important question, we used the results from the MLM model in columns 1-3 of Table 4 to calculate the level of income where the probability of selecting the clean fuel mix (Cluster 4, ngfw) in preference to the other clusters was exactly equal to 50 percent, holding other variables constant at their original values (and weighting to account for the stratified nature of the sample). That is, we calculated the income level where households would be more likely than not to switch to the clean fuel mix, using the following formula:

$$0.5 = \frac{1}{1 + \sum_{j=0}^{m} \exp^{\beta_j X_i'}} \qquad (3)$$

Table 5 shows this calculated level of income where the probability of selecting the clean fuel mix is exactly 50 percent. At incomes higher than this, the probability of choosing the clean fuel mix are greater than 50 percent and at incomes lower than this, the probability of selecting the clean fuel mix are less than 50 percent. It is evident from the table that the income level for choosing the clean fuel mix (ngfw) in preference to Cluster 3 (fwadaw) is higher as compared with the other clusters (awadfw and adfwaw). In cluster 3 (fwadaw), households are mostly consuming firewood, while in other two clusters they are mainly consuming agricultural waste and animal dung respectively. Firewood is a comparatively more expensive fuel than the other solid fuels, so as households' incomes increase they may shift from Clusters 2 and 7 first to Cluster 3, and then to the cleaner fuel mix at even higher incomes. This interpretation supports the energy ladder and energy stacking hypotheses. The table also shows that rural households would require much higher incomes than urban households to shift to the cleaner fuel mix. This reflects that only very high income rural households have access to piped natural gas, which forms the main energy source in the cleaner fuel mix cluster.

Given that the average income (monthly expenditures as proxy) in the weighted sample was 21,444 PKR (27,546 PKR in urban areas and 17,859 PKR in rural areas), this also demonstrates that, in the absence of a significant increase in the availability of piped gas connections, it is unlikely that Pakistan will grow out of solid fuel use, particularly in rural areas. Similarly, the availability of electricity connection reduces the income level at which households are likely to switch to cleaner fuel mix use. In part, this is related to the rural/urban findings, since electricity connections are much less common in rural areas.

Table 5: Monthly Income (PKR) Threshold for Choosing Clean Fuel Mix

Base clusters	National level	Urban level	Rural level	Electricity connection Yes	Electricity connection No
Cluster 2	33297.8	12946.2	45324.2	30954.0	61249.9
Cluster 3	52778.8	22516.8	70652.3	48594.6	101788.8
Cluster 7	35431.6	9490.5	50625.1	33720.4	56291.9

Note: 1 USD = 100 PKR, 2014.

6. CONCLUSION AND POLICY IMPLICATION

This study examined the fuel combinations and determinants of fuel mix choice in Pakistan using a nationally representative data set. We found that the accessibility to piped natural gas is the most influential factor associated with the use of a clean fuel mix. Income and education were significant demographic factors associated with the use of cleaner fuels, supporting the hypothesis of energy stacking. We also demonstrate that income growth in Pakistan is unlikely to be conducive to households growing out of the use of solid fuels, particularly in rural areas. If government is concerned about indoor air pollution and wants to incentivise the use of cleaner fuel mixes by households, our results have clear implications. Expanding the availability of piped natural gas connections from main urban areas to include smaller urban areas and nearby villages is likely to encourage many, particularly non-agricultural households, to switch to cleaner fuels.

Similarly, extending the electricity grid throughout the country, with particular focus on rural villages, would allow these households to reduce their reliance on solid fuels (especially for lighting). While economic growth will raise incomes, it is unlikely to have a substantial impact on the use of solid fuels in rural areas without this increased accessibility of natural gas.

References

Arnold, J. E. M., Köhlin, G., and Persson, R. (2006). Woodfuels, livelihoods, and policy interventions: Changing Perspectives. *World Development*, *34*(3), 596–611. https://doi.org/10.1016/j.worlddev.2005.08.008

Bhatt, B. P., and Sachan, M. S. (2004). Firewood consumption pattern of different tribal communities in Northeast India. *Energy Policy*, *32*(1), 1–6. https://doi.org/10.1016/S0301-4215(02)00237-9

Bonan, G. B. (2008). Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, *320*(5882), 1444–1449. https://doi.org/10.1126/science.1155121

Brouwer, R., and Falcão, M. P. (2004). Wood fuel consumption in Maputo, Mozambique. *Biomass and Bioenergy*, *27*(3), 233–245. https://doi.org/10.1016/j.biombioe.2004.01.005

Caliński, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Campbell, B. M., Vermeulen, S. J., Mangono, J. J., and Mabugu, R. (2003). The energy transition in action: urban domestic fuel choices in a changing Zimbabwe. *Energy Policy*, *31*(6), 553–562. https://doi.org/10.1016/S0301-4215(02)00098-8

Colbeck, I., Nasir, Z. A., and Ali, Z. (2010). The state of indoor air quality in Pakistan—a review. *Environmental Science and Pollution Research*, *17*(6), 1187–1196. https://doi.org/10.1007/s11356-010-0293-3

Cooper, J. A. (1980). Environmental Impact of Residential Wood Combustion Emissions and its Implications. *Journal of the Air Pollution Control Association*, *30*(8), 855–861. https://doi.org/10.1080/00022470.1980.10465119

Edwards, J. H. Y., and Langpap, C. (2012). Fuel choice, indoor air pollution and children's health. *Environment and Development Economics*, *17*(4), 379–406. https://doi.org/http://dx.doi.org.ezproxy.waikato.ac.nz/10.1017/S1355770X12000010

Farsi, M., Filippini, M., and Pachauri, S. (2007). Fuel choices in urban Indian households. *Environment and Development Economics*, *12*(06), 757–774. https://doi.org/10.1017/S1355770X07003932

Heltberg, R. (2004). Fuel switching: evidence from eight developing countries. *Energy Economics*, *26*(5), 869–887. https://doi.org/10.1016/j.eneco.2004.04.018

Heltberg, R. (2005). Factors determining household fuel choice in Guatemala. *Environment and Development Economics*, *10*(3), 337–361.

Hosier, R. H., and Dowd, J. (1987). Household fuel choice in Zimbabwe. *Resources and Energy*, *9*(4), 347–361. https://doi.org/10.1016/0165-0572(87)90003-X

Irfan, M., Cameron, M. P., and Hassan, G. (2018). Household energy elasticities and policy implications for Pakistan. *Energy Policy*, *113*, 633–642. https://doi.org/10.1016/j.enpol.2017.11.041

Jan, I., Khan, H., and Hayat, S. (2012). Determinants of Rural Household Energy Choices: An Example from Pakistan. *Polish Journal of Environmental Studies*, *21*(3), 635–641.

Joon, V., Chandra, A., and Bhattacharya, M. (2009). Household energy consumption pattern and socio-cultural dimensions associated with it: A case study of rural Haryana, India. *Biomass and Bioenergy*, *33*(11), 1509–1512. https://doi.org/10.1016/j.biombioe.2009.07.016

Jumbe, C. B. L., and Angelsen, A. (2011). Modeling choice of fuelwood source among rural households in Malawi: A multinomial probit analysis. *Energy Economics*, *33*(5), 732–738. https://doi.org/10.1016/j.eneco.2010.12.011

Kim, K.-H., Jahan, S. A., and Kabir, E. (2011). A review of diseases associated with household air pollution due to the use of biomass fuels. *Journal of Hazardous Materials*, *192*(2), 425–431. https://doi.org/10.1016/j.jhazmat.2011.05.087

Lakshmi, P. V. M., Virdi, N. K., Thakur, J. S., Smith, K. R., Bates, M. N., and Kumar, R. (2012). Biomass fuel and risk of tuberculosis: a case—control study from Northern India. *Journal of Epidemiology and Community Health (1979-)*, *66*(5), 457–461.

Landrigan, P. J., Fuller, R., Acosta, N. J. R., Adeyi, O., Arnold, R., Basu, N. (Nil), … Zhong, M. (n.d.). The Lancet Commission on pollution and health. *The Lancet*. https://doi.org/10.1016/S0140-6736(17)32345-0

Leach, G. (1992). The energy transition. *Energy Policy*, *20*(2), 116–123. https://doi.org/10.1016/0301-4215(92)90105-B

Lee, L. Y.-T. (2013). Household energy mix in Uganda. *Energy Economics*, *39*, 252–261. https://doi.org/10.1016/j.eneco.2013.05.010

Miah, M. D., Foysal, M. A., Koike, M., and Kobayashi, H. (2011). Domestic energy-use pattern by the households: A comparison between rural and semi-urban areas of Noakhali in Bangladesh. *Energy Policy*, *39*(6), 3757–3765. https://doi.org/10.1016/j.enpol.2011.04.004

Mishra, V. (2003). Indoor air pollution from biomass combustion and acute respiratory illness in preschool age children in Zimbabwe. *International Journal of Epidemiology*, *32*(5), 847–853. https://doi.org/10.1093/ije/dyg240

Narasimha Rao, M., and Reddy, B. S. (2007a). Variations in energy use by Indian households: An analysis of micro level data. *Energy*, *32*(2), 143–153. https://doi.org/10.1016/j.energy.2006.03.012

Narasimha Rao, M., and Reddy, B. S. (2007b). Variations in energy use by Indian households: An analysis of micro level data. *Energy*, *32*(2), 143–153. https://doi.org/10.1016/j.energy.2006.03.012

Nasir, Z. A., Murtaza, F., and Colbeck, I. (2015). Role of poverty in fuel choice and exposure to indoor air pollution in Pakistan. *Journal of Integrative Environmental Sciences*, *12*(2), 107–117. https://doi.org/10.1080/1943815X.2015.1005105

Osiolo, H. (2009). *Enhancing household fuel choice and substitution in Kenya /*. Nairobi, Kenya : Kenya Institute for Public Policy Research and Analysis,.

Ouedraogo, B. (2006). Household energy preferences for cooking in urban Ouagadougou, Burkina Faso. *Energy Policy*, *34*(18), 3787–3795. https://doi.org/10.1016/j.enpol.2005.09.006

Pachauri, S., and Jiang, L. (2008). The household energy transition in India and China. *Energy Policy*, *36*(11), 4022–4035. https://doi.org/10.1016/j.enpol.2008.06.016

Peng, W., Hisham, Z., and Pan, J. (2010). Household level fuel switching in rural Hubei. *Energy for Sustainable Development*, *14*(3), 238–244. https://doi.org/10.1016/j.esd.2010.07.001

Pundo, M. O., and Fraser, G. C. (2006). Multinomial logit analysis of household cooking fuel choice in rural Kenya: The case of Kisumu district. *Agrekon*, *45*(1), 24–37. https://doi.org/10.1080/03031853.2006.9523731

Rahut, D. B., Das, S., De Groote, H., and Behera, B. (2014). Determinants of household energy use in Bhutan. *Energy*, *69*, 661–672. https://doi.org/10.1016/j.energy.2014.03.062

Romesburg, C. (2004). *Cluster Analysis for Researchers*. Lulu.com.

Scott, A. J., and Knott, M. (1974). A Cluster Analysis Method for Grouping Means in the Analysis of Variance. *Biometrics*, *30*(3), 507–512. https://doi.org/10.2307/2529204

Sehgal, M., Rizwan, S. A., and Krishnan, A. (2014). Disease burden due to biomass cooking-fuel-related household air pollution among women in India. *Global Health Action*, *7*. https://doi.org/http://dx.doi.org.ezproxy.waikato.ac.nz/10.3402/gha.v7.25326

Sehjpal, R., Ramji, A., Soni, A., and Kumar, A. (2014). Going beyond incomes: Dimensions of cooking energy transitions in rural India. *Energy*, *68*, 470–477. https://doi.org/10.1016/j.energy.2014.01.071

Torres-Duque, C., Maldonado, D., Pérez-Padilla, R., Ezzati, M., and Viegi, G. (2008). Biomass Fuels and Respiratory Diseases. *Proceedings of the American Thoracic Society*, *5*(5), 577–590. https://doi.org/10.1513/pats.200707-100RP

WHO (n.d.) Indoor air pollution: national burden of disease estimates. Retrieved October 8, 2016, from http://www.who.int/indoorair/publications/nationalburden/en/