CERE

CERE Working Paper, 2016:9

Who adopts LPG as the main cooking fuel and why? Empirical evidence on Ghana based on national survey.

Amin Karimu[a]
Justice Tei Mensah[b]
George Adu[c,d]

[a]Centre for Environmental and Resource Economics
Umeå School of Business and Economics
Umeå University, Sweden

[b]Department of Economics, Swedish University of Agricultural Sciences
Uppsala, Sweden

[c]The Nordic Africa Institute, Box 1703, SE-75147, Uppsala, Sweden
[d]Department of Economics, Kwame Nkrumah University of Science and Technology
Kumasi, Ghana

Department of Economics, Umeå University, S-901 87, Umeå, Sweden

www.cere.se

Who adopts LPG as the main cooking fuel and why? Empirical evidence on Ghana based on national survey.

Amin Karimu[a]
Justice Tei Mensah[b]
George Adu[c,d]

[a]Centre for Environmental and Resource Economics
Umeå School of Business and Economics
Umeå University, Sweden
[b]Department of Economics, Swedish University of Agricultural Sciences
Uppsala, Sweden
[c] The Nordic Africa Institute, Box 1703, SE-75147, Uppsala, Sweden
[d]Department of Economics, Kwame Nkrumah University of Science and Technology
Kumasi, Ghana

April 2016

Abstract

The aim of this paper is to identify the factors that influence the probability of adopting LPG as the main cooking fuel in Ghana using household level data gleaned from last two nationwide household surveys (GLSS 5 & GLSS 6). Using a flexible semi-parametric specification, the following were uncovered. First, we find socioeconomic and demographic factors such as income, education, access to urban infrastructure, location of household, as key drivers of households' choice of LPG as main cooking energy source. Again the influences of these factors are stable across time, and with a strong price effect. The evidence shows that urban households with better socioeconomic and demographic factors are likely to adopt LPG as the main cooking fuel relative to households in rural areas and also urban households with poor socioeconomic and demographic factors. Finally, we observe that the imposition of fully parametric structure (functional form) prior to estimation on factors such as age of household head, income and household size as done in the literature is inappropriate, at least in the case of Ghana and tend to bias the marginal effects. There is strong evidence of variations in the response rate of LPG adoption over the domains of income, household size and the age of the household head. The results suggest a policy dichotomy between rural and urban dwellers for it to be effective.

JEL Classification: C14, O13, Q41, Q42

Key Words: Fuels, cooking, households, development, energy poverty, Ghana

1. Introduction

Household energy use has received considerable attention among policy makers and researchers. The keen interest in fuel use patterns particularly in developing economies stems from the heavy dependence on solid biomass or coal fuels, and their associated negative impacts on livelihoods, environment and health. Coupled with the use of inefficient cooking stoves, emissions from biomass combustion is associated with high incidence of acute respiratory infections (ARI), particularly, among women and children (Ezzati and Kammen, 2001; Mishra, 2003). Also, fuelwood collection has been shown to affect educational and labour outcomes of women and children, due to the long hours spent on collection. On the environmental impacts, emissions of black carbon among other greenhouse gases have negative impacts on the climate. Again, unsustainable harvesting of biomass for energy use is a leading driver of deforestation in tropical regions, especially, sub-Saharan Africa (SSA). Thus, continuous dependence on biomass energy, especially the use inefficient cookstoves and methods by households poses a major threat to the survival of livelihoods and entails significant externalities on the global environment and climate.

The recognition of these impacts has led many governments and development agencies to propose various measures in the quest of addressing the issue of energy poverty by encouraging the up-take of improve cookstoves, sustainable production of biomass fuels and the up-take of "modern" fuels such as LPG. In Ghana, for example, government has over the past three decades implemented policies such as price subsidies on LPG. Even though the policy has recorded some level of success, there is still room for improvement to ensure that LPG up-take serves as one (others include adoption of improve cookstoves for some socio-economic groups, where modern fuel up-take is almost impossible in the near future due to cost, culture and reliability of fuel supply) of the key solutions to the energy poverty issues in Ghana and the associated environmental, health and development impacts. Currently, only 23.1% of Ghanaian households use LPG as the main cooking fuel (Mensah and Adu, 2015) up from under 5% in the early 1990s.

Contrary to predictions from energy theories such as the "energy ladder hypothesis", which suggest a simplistic switch from traditional fuels to modern fuels driven largely by economic forces such as income and price; empirical evidence from many developing economies suggest that the dynamics of fuel switching is far more complex and less understood. There is a high degree of inertia in fuel switching patterns, often attributed to non-economic factors such as taste, preference, reliability in supply, inter alia. This is evident in the Ghanaian context, since despite a rise in income and price subsidies on LPG, less than a quarter of households use LPG as a main cooking fuel.

Given these complexities, it is important to fully understand and establish the motivating factors driving households' choice of a relatively 'clean' cooking fuel like LPG over alternate fuels. More importantly, we address the question of "how does the response of LPG adoption rate vary over time to changes in factors driving it?" Further, even though majority of the extant studies have examined the determinants of cooking energy choices, which can be grouped broadly into two strands, a group that is more descriptive in highlighting cooking fuel transition (e.g., Barnes and Qian,1992; Dowd,1989; Hofstad et al.,2009; Masera et al.,2000) and a group that utilises quantitative(econometric approaches) to examined the transition probabilities and/or demand for cooking fuels (see e.g., Barnes, et al., 2005; Chambwera, 2004; Chambwera and Folmer, 2007;Gebreegziabher et al. 2012; Heltberg, 2005; Kebede et al., 2002;Mekonnen and Köhlin, 2008; Mensah and Adu, 2015). Both strands have their respective merits in providing valuable information in understanding energy poverty dynamics, especially in developing countries. Often, the qualitative strand provides some intuition and hypothesis on cooking fuel transition dynamics, and the quantitative approaches evaluate the hypothesis to ascertain their validity and provide causal dynamics. One key question that remains is whether the functional form specification used by these studies has any impact on the results derived. Theoretically, demand for energy is assumed to be a non-linear function[1] of household size, the age of household head and income. However, the exact form of the non-linearity is unknown, a priori. To account for this non-linearity in the energy demand relationship, previous researchers have applied variety of techniques such as log transformation[2], introduction of quadratic terms, and the use of estimators such as logit/probit models in binary outcome cases. What still remains as an empirical challenge is the extent by which these appropriately capture the inherent non-linearity in the fuel choice functions, since the exact form of the relationship is unknown to the researcher, a priori. We contribute to this discussion by adopting semi-parametric probit model which offers more flexibility in capturing the non-linearity in the relationship without imposition of any functional from prior to estimation. To illustrate the relevance of allowing the data generating process (DGP) to select the precise form of the non-linearity rather than imposing some functional form prior to estimation, we compare our preferred semi-parametric estimates

[1] Relations that does not follow a linear path, but changes overtime or over different levels of a variable, e.g., we derive more satisfaction from the first glass of water when tasty than from the second, the third and so on, implying that our total satisfaction from a class of water will increase but at a decreasing rate, which reflects non-linear relationship between our satisfaction and the number of glass of water we had rather than a linear relationship.

[2] This is generally a data transformation technique, where natural logarithms are applied to the raw data to reduce the skewness of the distribution or apply to non-linear functions to induce linear relationships.

with fully parametric estimates. This comparison allows us to establish possible biases in the evidence reported in the extant literature, especially for Ghana. Further it also highlights the possibility of such biases for other developing countries studies.

The aim of this study therefore is in two folds: first, to analyze the main driving forces behind a household's choice of LPG as its' main cooking fuel over alternate fuels such as biomass, and how the influences of these factors vary over time. Second, we seek to test the effect(s) of functional form bias on the results derived. In line with the aim, the following four questions are addressed: (1) does the key determinants usually cited in the literature on cooking fuels in developing countries also important for LPG uptake in Ghana?; (2) does the effect of the key variables vary over time?; (3) whether the usual fully parametric specification usually imposed on the empirical models is free from functional specification bias; and (4) what is the role of prices on LPG uptake and the policy ramification that emerges from it? We utilize data from two most recent nationwide household surveys to estimate a probit model using both parametric and semi-parametric econometric techniques. Results are then compared across the two datasets and models.

This study offers immense contribution to energy policy design in Ghana and many developing economies, in search of effective strategies to boost household's use of cleaner and efficient fuels such as LPG. The study also offers relevant contribution to the extant literature, as it sheds more light on the role of functional form specification in household energy demand models.

Findings from our study reveal socioeconomic and demographic factors such as income, education, access to urban infrastructure, location of household, as key drivers of households' choice of LPG as main cooking energy source. Again the influences of these factors are stable across time. There is also evidence of a strong price effect. Finally, we observe that the imposition of parametric structure (functional form) prior to estimation on factors such as age of household head, income and household size as done in the literature is inappropriate for the case of Ghana, and tend to bias the marginal effects. There is strong evidence of variations in the response rate of LPG adoption over the domains of income, household size and the age of the household head.

The structure of the remaining sections of the paper is as follows: A review of relevant literature is presented in section 2. Section 3 describes the methodology and empirical strategy for the paper, while Section 4 focuses on data description. The discussion and analysis of the results from the paper are presented in section 5. Section 6 concludes with a summary and policy implications from the study.

2. **Literature Review**

Studies on the determinants of household energy choice in the literature are replete with evidence from both developed and developing countries, albeit studies from the latter are often constrained by the lack of sufficient data especially on energy prices and expenditures. Notable among such studies include Couture et al (2012) using French data, Gebreegziabher et al (2012), Kebede et al. (2002), Mekonnen and Köhlin (2008) for Ethiopian households, Barnes, et al., 2005; Heltberg (2004, 2005) for developing countries, Chambwera (2004), Chambwera and Folmer (2007) for Zimbabwe, Akpalu et al. (2011), Mensah and Adu (2015) for Ghanaian households and Ouedraogo (2006) for households in Burkina Faso among others.

Evidence from Kebede et al. (2002) using Ethiopian data, reveal that non-income factors such as fuel availability and demography are key drivers of demand for modern energy other than income. This evidence offers key insight into the reasons behind the slow adoption rate of modern energy even among high income households in developing economies, as the market for these fuels especially LPG and electricity, is often marred with frequent shortages and inefficiencies in service delivery. This inherent uncertainty in the supply of LPG and electricity creates disincentives for households' energy transition. Mekonnen and Köhlin (2008) found Similar conclusions.

Nevertheless, income remains a key constraint to the transition of households from the use of traditional fuels to modern and clean fuel types like LPG and electricity. Edwards and Langpap (2005) provides an empirical test of this claim by analysing the effects of start-up cost and restricted access to credit on the households' decision to switch from firewood to gas using Guatemalan data. Using the parametric probit model, their results reveal that access to credit via its effect on the purchasing power of households to purchase a gas stove, exerts a significant influence on the household's consumption levels of fuelwood. Their findings also reveal that high start-up cost is indeed a limiting factor for Guatemalan households in the transition from fuelwood to LPG. Hence targeted subsidies towards reducing the cost of gas stove will yield desired results. In contrast, Heltberg (2005) argues that the idea of a complete transition from fuelwood to modern fuels such as LPG does not seem plausible in the Guatemalan case, as many of the households that adopt these clean fuels still retain some preference for fuelwood. Thus any policy such as LPG subsidies that seeks to engender a complete switch to LPG will not yield the intended results. Heltberg (2005) further asserts that aside income, household fuel choice is strongly influenced by traditional cooking techniques, taste and preferences to the extent that demand for fuelwood is price inelastic in some cases. Also a review of the drivers of households' substitution between primary and secondary forest products as cooking fuel, particularly in rural communities where commercial fuels such as LPG are seldom, is presented by Hyde and Köhlin (2000).

Evidence from other studies confirms the findings of Heltberg (2005) that whereas inter-fuel substitution is ubiquitous among households in developing countries, some households remain stacked to biomass fuels (Akpalu et al, 2011). Davis (1998) argues that even with high availability of modern fuels such as LPG and kerosene, it is unlikely that biomass fuels can be completely substituted for the modern fuels, largely because of taste and preference factors. A typical example of this is provided by WEC (1999) which revealed that in one community in Sierra Leone, nearly 70% of the families avoid switching to modern fuels and prefers to retain the use of firewood as wood stove is the most suitable option for the preparation of a traditional food staple within the region (Akpalu et al, 2011). For a comprehensive review on the adoption of improved cooking fuel and stoves by households see Lewis and Pattanayak (2012).

The evidence on Ghana is relatively scanty except the works of Mensah and Adu (2015), Karimu (2015) and Akpalu et al (2011). Findings from Mensah and Adu (2015), and Karimu (2015) lend support to the energy ladder hypothesis that household income is a major determinant of household energy choice. Further, they find social and demographic factors as well as access to energy supplies as key factors influencing the choice of cooking fuel among Ghanaian households. Akpalu et.al. (2011) also find evidence of strong preference for LPG among Ghanaian households, followed by charcoal and kerosene respectively. In addition, their study suggests that with the exception of kerosene that has price-elastic demand, demand for charcoal, LPG and fuelwood are price inelastic.

In terms of the methodological approach, the available studies are largely based on parametric techniques including the discrete choice models (logit and probit), multinomial probit and logit, full information maximum likelihood (FIML) techniques and the ordinary least square approach. See Lewis and Pattanayak (2012) for extensive review of household energy choice studies and methodological approaches. The use of non-parametric and semi-parametric techniques in analyzing household fuel choice is however very scanty in the extant literature and completely non-existent in the Ghanaian context. This paper makes a contribution on this by applying variants of the popular specifications in the literature on household energy studies and estimating the preference for LPG as the main cooking fuel using semi-parametric method. The results are then compared with the fully parametric methods as common in the previous studies.

3. Econometric model

Based on previous literature on household cooking fuels, we estimate a choice model for LPG for cooking by assuming that a household's choice for cooking fuel is influence by income (Y), a vector of household characteristics (Z) such as age, household size, gender of the head of the household and education. We also included access to modern infrastructure (M), location (L), the availability of the various cooking fuels (A)and a vector of prices for the various fuels (P_i'). The household choice for cooking fuel in general is expressed as:

$$Q = Q(Y, Z, A, L, M, P^{'}) \tag{1}$$

Suppose a representative household i faces j alternatives of cooking fuels (where $j = 1, 2$; the first alternative is LPG and the second is the other fuels), and that the indirect utility derived from each of the j alternatives is defined by V_{ij}. The indirect utility function is composed of two parts, an observed part, $\mathbf{x}_i'\beta_j$ and an unobserved part ε_{ij}, where $\mathbf{x}_i$ is a vector of all the variables in equation (1), β_j a vector of parameters to be estimated. The indirect utility for alternative j for household i can then be expressed as:

$$V_{ij} = \mathbf{x}_i'\beta_j + \varepsilon_{ij} \tag{2}$$

The unobserved part, ε_{ij} , is assumed to be jointly normally distributed with $\varepsilon \sim N[0, \Sigma]$. The probability that household i chooses the first alternative is now:

$$\begin{aligned} P_{i1} &= pr(\varepsilon_{i2} - \varepsilon_{i1} < \mathbf{x}_i'\beta_1 - x_i'\beta_2) \\ &= pr[\dot{\varepsilon}_{i,21} < \mathbf{x}_i'(\beta_1 - \beta_2)] \end{aligned} \tag{3}$$

where $\dot{\varepsilon}_{i,21} = \varepsilon_{i2} - \varepsilon_{i1}$. Similar expressions can be obtained for the other alternatives P_{i2}. It is assumed that ε_{ij} has a joint normal density function defined as $f(\varepsilon_{ij}) = f(\varepsilon_{i1,}\varepsilon_{i2})$, and let y_{ij} denote a discrete choice outcome variable that takes a value 1 if household i chooses fuel j and 0 otherwise. The cumulative probability for the choice of the first alternative fuel (LPG) by household i can now be expressed as:

$$P_{i1} = pr[y_i = 1] = \int_{-\infty}^{\dot{V}_{i,12}} f(\dot{\varepsilon}_{i,21}) d\dot{\varepsilon}_{i,21} \tag{4}$$

The expression in equation (4) is specific to the first fuel. In a more general case, the choice probability for household i choosing alternative j is given by $P_{ij} = pr[y_i = j] = m_j(\mathbf{x}_i'\beta_j)$ where $m_j(\mathbf{x}_i'\beta_j)$ takes a similar expression as in equation (4). The log likelihood function for a sample of N independent households with J alternatives can then be expressed as:

$$\ell = Ln\, L = \sum_{i=1}^{N}\sum_{j=1}^{J} y_{ij} Ln(\hat{P}_{ij}) \tag{5}$$

where $\hat{P}_{ij}$ is estimated via a similar expression as in equation (4) using simulation method, which is substituted into the log likelihood function and maximised to obtain the parametric estimates for the $\beta's$.

We could relax the functional relation for some of the variables in the $\mathbf{x}_i$ vector using semi-parametric approach since we often do not know the "correct" functional form a prior. Let define a sub-set of the $\mathbf{x}_i$ vector that we want to relax the functional from as $\ddot{\mathbf{x}}_i$,the indirect utility function as presented in Eq. (2) become:

$$V_{ij} = \mathbf{z}_i'\beta_j + m(\ddot{\mathbf{x}}_i) + \varepsilon_{ij} \tag{6}$$

Where $m(.)$ is a smooth function that is twice differentiable and additive, implying each variable in $\ddot{\mathbf{x}}_i$ has its own smooth function; and $\mathbf{z}$ is a subset of vector $\mathbf{x}$. The smooth functions are estimated using penalized cubic regression splines via the generalized additive model (GAM) framework. Further we assumed the link function to be probit. Details on penalized regression splines is in Hastie and Tibshirani, 1990; Wahba, 1990; Green and Silverman, 1994; and Wood, 2000. Our choice of the $\ddot{\mathbf{x}}_i$ vector in this paper is influence entirely by whether the variable is a continuous variable or not. Since the smooth function must be twice differentiable, it is only the continuous variables that meet this requirement. Our $\ddot{\mathbf{x}}_i$ vector therefore comprises of income, age of household head, household size and prices of the various cooking fuels. Note however that in the empirical application not all the variables are available for some of the samples and in such cases, we only utilised the available variables in the sub-set for the nonparametric part of the model.

4. Data

The data for the paper is taken from the last two Ghana living standards surveys (GLSS 5, 2005/06 & GLSS 6, 2012/2013), which is a nationally representative household survey. The sampling frame for the survey was the population living in private households in Ghana. The above sample frame was divided into two sampling units, a primary and secondary sampling unit. The primary sampling unit was defined as the census enumerated areas (EAs) that are stratified into the ten administrative regions of Ghana based on proportional allocation using the population in each of the ten regions. The second sampling unit on the other hand was defined as the households living in each of the EA's. The sampling design for the survey was that of two-stage stratified random sampling approach, where in the first stage 550 EA's was considered, while in the second stage, 15 households per EA was considered. The combined outcome of the two-stage sampling resulted in a total of 8700 households' nationwide. In the end, however, 8687 households were successfully interviewed representing a 99.85 % response rate for the GLSS5 sample. The sample size used for this study further reduced to 8262 households because of lack of information on energy use for cooking for some of the sampled households. Similar sampling approach was applied to obtain the GLSS6 data. However, in the GLSS6, 1200 EA's were considered in the first stage leading to a total of 18000 households. Out of this, 16,772 were successfully interviewed.

Using the above survey design, data was extracted on the following key variables: household income, consumption, expenditure, education, energy use for cooking, demographic characteristics such as education, gender of the head of the household, age of the household head, size of the household, main fuel for lightening, household location, income, and the availability of the various cooking fuels. The education variable has four levels; basic, secondary and tertiary (university education) with no formal education as the reference category. The main source of lightening for the household is used as a proxy for the level of access to modern infrastructure. Also, to control for supply side factors that may influence households' choice decisions, we include as covariate(s), the reliability in access to LPG and its substitute fuels. This variable(s) measures whether households have reliable supply of the fuel in their community or otherwise. Thus, for example, reliable supply of LPG (=1) if LPG is readily available in the local market and reliable supply of LPG (=0) if LPG is not readily available or characterized by frequent shortages.

The information on the main fuel sources used by each of the households cover fuels such as crop residue, fuelwood, charcoal, kerosene, liquefied petroleum gas (LPG), and electricity. The information on each of the fuels refers to choice and not the quantity; hence our analysis is limited to choices and probabilities associated with these choices. Besides the lack of information on the quantities on each of the fuels, the GLSS 5 data also lack information on price of the various fuels, which further limits our analysis. The GLSS 6 however, has information on relative prices of alternate fuels. Furthermore, the data set does not contain information on combination of cooking fuels by households refer in the literature as "fuel stacking", a common feature especially among urban households in developing countries. A list of the variables in the data set and their definitions is presented in Table A, while the summary statistics in Table A1, both are in the appendix.

5. Empirical Results

Our empirical analysis follows the following steps. First, we run a parametric probit model for each of the two recent waves of surveys, the GLSS 5 and GLSS 6. The purpose of this first step is to assess the development of the determinants for LPG uptake. In the second step, we relax the functional form, especially for the continuous variables (income, age of the household head and household size for both sampled periods) and test if the flexible semi-parametric specification is more appropriate for our LPG model relative to the parametric specification. The final step, we assess the relative price effects of the various cooking fuels on the choice of LPG, based on the chosen model from the second step and utilising the data set with prices.

5.1 Parametric probit model Estimates for each of the two surveys

The results from the first step are reported in Table 1. The table contains the coefficient estimates and the corresponding marginal effects. The estimated marginal effects can be interpreted as follows. A positive estimated marginal effect for a regressor implies that the regressor increases the probability of choosing LPG as a cooking fuel relative to the other fuels (fuelwood, charcoal, kerosene and crop residue/sawdust). Similarly, a negative marginal value implies a decrease in the probability of choosing LPG as a cooking fuel. The estimates as reported in Table 1 reveal the following; (a) that in the GLSS5 sample, reliable supply of firewood and age of the household head are the only variables with estimates that are insignificant at the 5% level, whereas they are significant in the GLSS6 sample. Also, reliable supply of kerosene is significant in the GLSS5 but insignificant in GLSS6, (b) the key drivers of cooking fuel choice suggested by previous studies (e.g., Heltberg, 2005; Gupta and Köhlin, 2006; Ouedraogo, 2006; Karimu, 2015; Mensah and Adu, 2015) such as education, income, urban location, household size and access to modern infrastructure proxy by the use of electricity as the main source of lightening are statistically significant in both samples, (c) all the significant marginal effects across the two samples have the same sign but differ in magnitude. The size of each of the significant marginal effects in GLSS6 is larger than the corresponding counterpart in GLSS5 sample. It is important to stress that, given that the comparison is based on repeated cross sectional national representative household data, the changes are at aggregate national level not individual household level.

Table 1: Parametric probit estimates for Households LPG choice using GLSS 5 and GLSS6 data

Dept.Variable: LPG (1, 0)	**GLSS 5**	**Marginal E.**	**GLSS 6**	**Marginal E.**
(Intercept)	-6.43*		-2.45***	
	(0.64)		(0.17)	
Male head	-0.18**	-0.80**	-0.11**	-2.4**
	(0.07)	(0.004)	(0.05)	(0.01)
Basic Education	0.56***	3.0***	0.35***	6.8***
	(0.08)	(0.01)	(0.05)	(0.01)
Secondary Education	1.24***	14.4***	0.65***	16.4***
	(0.09)	(0.02)	(0.05)	(0.02)
Tertiary Education	1.60***	26.7***	1.42***	44.5***
	(0.11)	(0.04)	(0.07)	(0.03)
Reliable supply of LPG	1.36***	19.6***	0.85***	20.3***
	(0.16)	(0.05)	(0.06)	(0.02)
Reliable supply of Kerosene	-0.49**	-1.52***	0.05	0.9
	(0.22)	(0.01)	(0.06)	(0.01)
Reliable supply of Charcoal	-0.38*	-1.23**	-0.23***	-4.8***
	(0.21)	(0.01)	(0.08)	(0.02)
Reliable supply of Firewood	-0.002	-0.01	-0.47***	-9.3***
	(0.18)	(0.01)	(0.06)	(0.01)
Urban	0.85***	4.6***	0.57***	10.7***
	(0.11)	(0.01)	(0.06)	(0.01)
Electricity	0.78***	3.7***	0.86***	13.8***
	(0.105)	(0.01)	(0.08)	(0.01)
HHsize	-0.07***	-0.31***	-0.05***	-0.8***
	(0.02)	(0.01)	(0.01)	(0.002)
Age_hhd	-0.002	-0.01	-0.01***	-0.1***
	(0.002)	(0.001)	(0.001)	(0.0002)
Log income	0.24***	1.04***	0.09***	1.6***
	(0.04)	(0.002)	(0.015)	(0.002)
AIC	2885.43		8325.45	
Log Likelihood	-1428.71		-4148.73	
Num. obs.	8262		15212	

Note:*** , **and*denote the significance level at the 0.1%,1% and 5%,respectively, the values under the heading "Marginal denote marginal effects" are multiply by 100 to convert into percentage points, while standard errors are in parenthesis.AIC is Akaike information criteria.

The implication of the larger marginal effects, especially for education, urban location, electricity, male headed household and income in the 2012/2013 sample relative to the 2005/2006 sample is that, theses variable tended to increase (positive signed marginal effects)/decrease (negative signed marginal effects) the choice of LPG as a main cooking fuel. For instance, a household's head with tertiary level of education have a 26.7% probability of choosing LPG as a cooking fuel in the 2005/2006 period, while in 2012/2013, the choice probability increases to 44.5%. The increase in the effect of education on the choice probability across the two time periods also holds for lower levels of education –thus basic education choice probability increased from 3% to 6.8% and secondary from 14.5% to16.4%.

The general conclusion from these results is that; income, urban location, access to electricity, reliable supply of LPG and other household's characteristics such as age, gender and size of the household are significant factors for the choice process for LPG as a cooking fuel. Moreover, these variables tended to persist over the two sample periods and more importantly, each of them tended to have a lager (absolute terms) effect in the 2012/13 period relative to the 2005/06 period. However, these results are obtained with the assumption that all the explanatory variables enter the model (log) linearly for LPG, but this might not be the case for some variables, especially the income variable. One cannot rule out the possibility that beyond a certain level of income the increasing marginal effect of income on the choice of LPG will start to diminished as income level further increases. A similar outcome for age and household size cannot be ruled out either. The other explanatory variables are however discrete variables, which restrict them to enter the model linearly. We examine the possible non-linearity of income, age and household size for LPG choice in the next section via a flexible semi-parametric probit model.

5.2 Semi-parametric probit model estimates

The semi-parametric probit model assumed all the explanatory variables, except age, income and household size; enter the model linearly for choice of LPG, while age, income and household size enter the model non-parametrically in a sense that the functional form for the three variables is not specified pre-estimation. Rather, the data determine the functional relation between each of them and the dependent variable (choice of LPG as cooking fuel). The estimated marginal effects for the semi-parametric model can be interpreted in the same way as done in the full parametric section above. The estimated results for semi-parametric model are reported in Table 2 and revealed that age, income and household size are non-linear (in logarithm) as the test for linearity is rejected at 5% level of significance based on the chi-square statistic, implying that a (log) linear restriction for these variables might not be appropriate. Other diagnostics such as AIC and log-likelihood statistics are smaller in the semi-parametric model relative to the full parametric probit model, providing support for the semi-parametric probit model as a better model for estimating the LPG choice model for the two sampled periods. A key revelation from the semi-parametric results is that the estimated marginal effects for education is smaller in each of the sampled periods relative to the parametric model's estimates. This also holds for the other covariates for the 2012/13 sample but not in the 2005/06 sample, where the estimated marginal effects for the other variables are rather lager in the parametric model relative to the semi-parametric model. The conclusion from this exercise is that, given that the (log) linearity test, the AIC and log-likelihood statistics all favour the semi-parametric model, the parametric estimates tended to be biased for both samples. The bias is on both directions depending on the covariate in the 2005/06 sample, whereas it is only upward bias in the 2012/13 sample. Further, the bias is more dramatic for education covariate than the other covariates.

Table 2: Semi-parametric (model1 for GLSS 5) and semi-parametric (model 2 for GLSS6)

Dept.Variable: LPG (1, 0)	**Model 1**	**Marginal E.**	**Model 2**	**Marginal E.**
(Intercept)	-3.07***		-2.43***	
	(0.11)		(0.07)	
Male head	-0.22***	-2.0**	-0.18***	-3.0***
	(0.06)	(0.01)	(0.04)	(0.01)
Basic Education	0.53***	5.1***	0.40***	6.0***
	(0.08)	(0.01)	(0.04)	(0.01)
Secondary Education	1.23***	11.7***	0.67***	12.5***
	(0.08)	(0.01)	(0.04)	(0.01)
Tertiary Education	1.53***	14.5***	1.38***	20.3***
	(0.11)	(0.01)	(0.05)	(0.01)
Reliable supply of LPG	1.36***	12.9***	0.82***	12.1***
	(0.13)	(0.01)	(0.03)	(0.01)
Reliable supply of Kerosene	-0.57**	-5.4**	0.05	0.8
	(0.21)	(0.02)	(0.04)	(0.01)
Reliable supply of Charcoal	-0.38	-3.6*	-0.15**	-2.2***
	(0.21)	(0.02)	(0.05)	(0.01)
Reliable supply of Firewood	-0.06	-0.6	-0.47***	-6.9***
	(0.14)	(0.01)	(0.04)	(0.01)
Urban	0.90***	8.5***	0.53***	7.8***
	(0.08)	(0.01)	(0.04)	(0.01)
Electricity	0.80***	7.6***	0.86***	12.7***
	(0.09)	(0.01)	(0.06)	(0.01)
HHsize	Smooth		Smooth	
Age_hhd	Smooth		Smooth	
Log income	Smooth		Smooth	
AIC	2851.78		8121.07	
Log Likelihood	-1405.35		-4033.79	
GCV score	-0.65		-0.47	
Test of linear restriction for HHsize	[35.91]***		[3.44]**	
Test of linear restriction for Age_head	[19.14]**		[7.66]***	
Test of linear restriction for log income	[119.51]***		[7.74]***	
Num. obs.	8262		15212	
Num. smooth terms	3		3	

Note: ***, **and*denote the significance level at the 0.1%, 1% and 5%, respectively, the values under the heading "Marginal denote marginal effects" are multiply by 100 to convert into percentage points, smooth indicate that the variable enter the model non-parametrically. Numbers in square brackets are chi-square statistics for the test of linear restriction on the smooth variables, while standard errors are in parenthesis. AIC is Akaike information criteria and GCV score stand for generalized cross validation score

The estimated marginal effects for the parametric part of the semi-parametric model as shown in Table 2 indicate that having a male as the head of the household has a negative effect on the choice probability for LPG. Specifically, having a male as the head of the household decreases the adoption probability for LPG by 2% and 3% in the 2005/6 and 2012/13 periods, respectively. The negative relationship of males as head of the household on the choice of LPG may be explain by the differences in decision by female and male heads of households, which more closely reflect their preferences, welfare, and opportunity cost of time, raising the possibility that female-headed households may be more responsive to interventions to improve access or reduce barriers to new cooking technology and fuels relative to male-headed households. This is consistent with Ditlhale and Wright (2003), who finds some evidence on the hypothesis that female decision makers are more likely to opt for modern energy relative to males.

The estimated results also indicate that education at each of the three levels have a significant positive effect on the choice of LPG. Basic, secondary and tertiary education increases the probability of choosing LPG by 5.1%, 11.7% and 14.5%, respectively for the 2005/6 period, relative to the reference category of no formal education. These values however increased to 6%, 12.5% and 20.3% for the 2012/13 period. The implication is that, if one control for the income effect associated with education, education appears to inform individuals of the possible benefits in the use of the LPG. For instance, the time savings associated with the use of LPG compared to say fuelwood, the health implications in terms of the smoke from fuelwood relative to LPG among other things make LPG the preferred option among educated households. The results also reveal that the marginal effect of education increases as level of education rises from basic level to tertiary level.

Furthermore, to examine the effects of supply side constraints on fuel choice adoption, we control for the reliable supply of LPG and alternate cooking fuels in the model. In line with expectations, the results show that whereas the reliable supply of LPG increases the choice of LPG as a cooking fuel, the reliable supply of kerosene and charcoal on the other hand reduces the choice probability for LPG. The estimated marginal effects indicate that the reliable supply of LPG increases the adoption probability for LPG by 12.9 and 12.1 percentage points in the 2005/6 and 2012/13 periods, respectively. Reliable supply of charcoal reduces the adoption probability for LPG in both sampled periods.

The estimated marginal effects on urban location (settlement) indicate that households located in urban centres on the average tends to have a higher probability of adopting LPG as the main cooking fuel relative to their rural counterparts for both sampled periods. However, there is a slight reduction in the size of the marginal effect in the 2012/13 period relative to the 2005/6 period, while households with access to modern infrastructure also increases the adoption probability for LPG for both sampled periods. The estimated marginal effect is however lager (12.7%) in the 2012/13 period relative to the 2005/06 (7.6%). The reason for using electricity, as the main source of light for the household to proxy for access to modern infrastructure is that, when an area have access to electricity, it is very likely that it will also have access to the basic infrastructural requirement for access to LPG and education, especially at the basic level. In line with the parametric results, the semi-parametric estimates indicate that income, urban location, access to electricity, reliable supply of LPG and other households characteristics such as age, gender and size of the household are significant factors for the choice process for LPG as a cooking fuel. In addition, these variables tend to persist over the two sample periods. However the semi-parametric results also reveal that the estimated marginal effects for the reliable supply of LPG, reliable supply of charcoal and urban location are smaller in the 2012/13 period relative to the 2005/06 period. This is in sharp contrast to the finding in the full parametric model, where the estimated marginal effects for each of the explanatory variables in the 2012/13 period is larger than that in the 2005/06 sampled period. A possible reason for this is likely due to the functional restriction on the parametric probit model for some of the covariates resulting in bias estimates.

Non-parametric estimates

The estimated nonparametric variables (household size, income, age) are all significant based on the estimated chi-square statistic for each of the nonparametric variables. The plot for the smooth functions for each of the three variables for 2005/06 and 2012/13 periods is presented in Figure 1 and 2, respectively. The estimated smooth function for household size (hhsize) indicates a negative trend with a possible non-linear relationship for each of the two sampled periods. The estimated function for age however, indicate a positive slope between the ages 15-40 years and thereafter a negative slope for the 2005/06 sampled period, while it is only positive between the ages 15-30 years for the 2012/13 period. In the case of income, the slope is positive for both samples. The functional form for age, income and household size is roughly the same across the two sampled periods, only slight differences in terms of the turning points, especially for the income and age variables. A possible explanation for the differences in the turning point for income between the two periods could be due to the growth and development in the economy, and the consequence of that on incomes and living standards of households. Particularly, the 2012/13 period falls within the era Ghana attained its lower middle income status. This means among other things an appreciable level of income for households, which might lead to a reduction in energy poverty relative to the 2005/06 period, when Ghana was still a low income country.

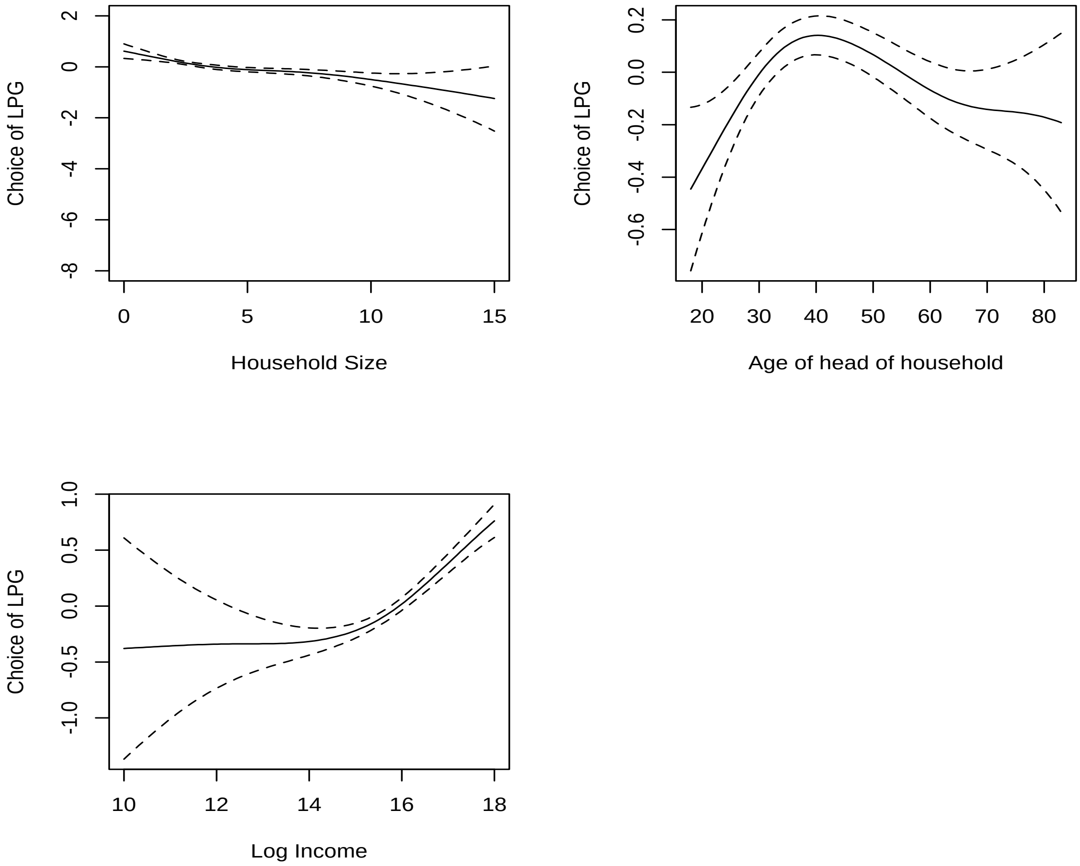

Figure 1: Plot of smooth estimates (model1-GLSS5) with 95% Bayesian confident interval represented by the dotted lines, while the solid lines are the estimated smooth function for each variable.

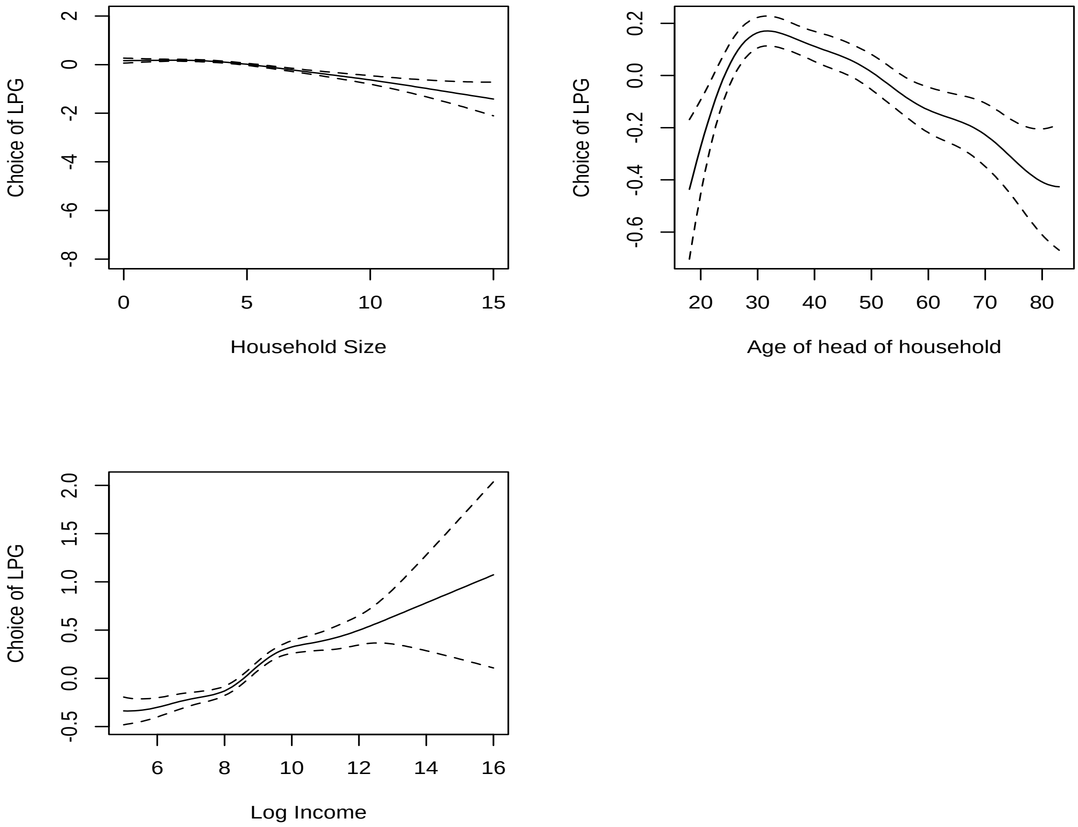

Figure 2: Plot of smooth estimates (model2-GLSS6) with 95% Bayesian confident interval represented by the dotted lines, while the solid lines are the estimated smooth function for each variable.

In order to make it possible to compare the estimates for age, income and household size between the parametric and semi-parametric model, we took the derivatives of the smooth functions to convert them into marginal effects. The calculated derivatives are reported in Table 3, where we considered the distribution of these derivatives for each of the covariates and for each of the two samples. The calculated mean for age, income and household size for the 2005/06 sample are -0.4%, 22% and -17%, respectively. Whiles for the 2012/13 the mean marginal effect for age, income and household size are -0.3%, 22% and -17%, respectively. These values are bigger in absolute terms than those from the parametric model, implying significant under estimation from the parametric model, especially for the estimates for income and household size. Note however that the mean value for age is not significant at the 5% level for each of two samples for the semiparametric model.

Table 3: Estimated marginal effects for non-parametric covariates for Semi-parametric model

	GLSS5 (model 1)			GLSS6 (model 2)		
	Age	Income	hhsize	Age	Income	hhsize
Minimum	-2.5**	4.1	-22.1	-3.6***	2.4	-18.8
	(0.01)	(0.06)	(0.02)	(0.01)	(0.04)	(0.01)
1st Quartile	-2.1**	4.9	-20.5***	-1.8	13.3**	-18.7***
	(0.01)	(0.07)	(0.02)	(0.01)	(0.05)	(0.01)
Median	-1.5	9.1	-18.0***	-1.4	21.9**	-18.5***
	(0.01)	(0.09)	(0.03)	(0.01)	(0.06)	(0.02)
Mean	-0.4	22.0*	-17.1***	-0.3	22.0**	-16.9***
	(0.01)	(0.12)	(0.05)	(0.01)	(0.07)	(0.02)
3rd Quartile	1.2	38.6**	-13.5***	-1.0	23.3**	-14.7***
	(0.01)	(0.19)	(0.07)	(0.01)	(0.09)	(0.03)
Maximum	3.1	62.1**	-10.6**	12.3***	51.0***	-13.0**
	(0.02)	(0.27)	(0.11)	(0.03)	(0.11)	(0.05)

Note:*** , **and*denote the significance level at the 0.1%,1% and 5%,respectively, the "marginal effects" are multiply by 100 to convert into percentage points, while standard errors are in parenthesis.

4.3 Price Effect

The previous sections assumed away the possible price effect on the choice for LPG due to non-availability of prices for the 2005/06 survey, and since one of the aim of our paper is to assess the development of the various key factors across time, we had to drop the prices in the 2012/13 sample in order to be consistent with the variables in the model for each of the samples. A possible consequence for dropping the price variable is omitted variables bias, since price, from previous studies (e.g., Heltberg, 2004, 2005) is considered as an important factor, at least for LPG choice process. In order to assess the lack of price effect on the previous analysis, particularly for the sample we have price data (GLSS6), we re-estimate a full parametric probit and semi-parametric probit models for the 2012/13 sample with prices as additional covariates. The inclusion of prices however reduced our sampled observation to 4918 due to missing data on the prices. The results are reported in Table 4, which reveal that the semi-parametric model fit the data relatively better than the parametric model based on AIC, log likelihood statistic and the chi-square test for linear restriction for age, income, household size and price for each of the fuels (LPG, kerosene, charcoal and firewood). This is in line with the results without the price effect from section 5.2, and as a consequence, we will only focus on the semi-parametric model in terms of comparison to the results without prices and interpretation of the model estimates.

In order to make a reasonable comparison to assess the potential bias due to price omission, we also estimated a semi-parametric model with the same number of observation as the model reported in Table 4, but without prices to assess the impact of omitted prices on the marginal effects of the included covariates. The result from this exercise is reported in Table B in the appendix. We found a slight increase in the estimated marginal effects in the model without prices relative to the model with prices for all the covariates except reliable LPG supply, which the reverse is true. However, the size of the differences is small, approximately 1% for most of the covariates. The indication is that omitted price bias is insignificant and therefore the model without prices does not suffer severely from such a bias. A further comparison to the semi-parametric model result reported in Table 2 for the GLSS6 sample, utilising the full sample observation, indicate similar conclusion regarding omitted price bias being insignificant, irrespective of the significant differences in the number of observations between the two (4918 for model with prices and 15212 for the model without price and utilizing the full sample observation).

The results for the smooth functions for the non-parametric part of the model with prices are reported in figure 3. All the price variables also depict patterns of non-linearity and therefore consistent with the chi-square test. Furthermore, LPG price had (generally) a negative relation on the choice of LPG. The log LPG price ranges from 0.87(0.30 US$)[3] Ghana cedis per kg to 2.37(0.72$) Ghana cedis per kg, where prices below 2.12 (0.64$) Ghana cedis per kg contain some form of government price subsidy via the LPG promotion programme. The other prices, especially firewood and charcoal are generally positively related to the choice of LPG as a cooking fuel. The implication of the negative own price effect and the positive firewood and charcoal prices is that, a policy that increases the price of firewood and charcoal via taxes and reduces the price of LPG via a subsidy is likely to encourage households into switching to LPG as a cooking fuel.

The positive effect of firewood price on the choice process of LPG is not due to a complete switch away from firewood to LPG fuel by households but rather a partial switch in the sense that most households do not rely on one fuel for their cooking needs. Households usually have more than one type of cooking fuel and utilized them depending on the type of food being cooked as well as the prevailing prices and fuel availability. Therefore higher prices for firewood is likely to reduce the frequency in the use of firewood to cook the daily meals, implying among other things an increase in the frequency of use for LPG for cooking purposes.

Next, we transformed the smooth functions of each of the non-parametric covariates into marginal effects across the entire domain of the function by taking the first derivative of each function. The estimated marginal effects are reported in figure 4, which indicated a varying marginal effect for household size, age of the head of the household, income and prices of the cooking fuels. The reason for this transformation is to make it easy for the reader to trace out the effects of each of the nonparametric variables on LPG. The estimated marginal effect of

[3] We apply the bank of Ghana exchange rate (1Ghana cedis to 3.3 US$) as at 16th July, 2015 between Ghana cedis and the US$ to convert the prices to dollar equivalent.

household size on the choice of LPG varies across household size, from -6% to -17%(scaled by multiplying by 100) as household size increases from 2 to 15.

Table 4: Semi-paramtric (model1) and Parametric (model2) results from GLSS 6 data with price effects

Dept.Variable: LPG (1, 0)	**Model 1**	**Marginal E.**	**Model 2**	**Marginal E.**
(Intercept)	-2.14***		-0.49	
	(0.13)		(0.60)	
Male head	-0.21***	-4.3**	-0.07	-2.2
	(0.06)	(0.01)	(0.07)	(0.02)
Basic Education	0.35***	7.1***	0.35***	10.2***
	(0.06)	(0.01	(0.07)	(0.02)
Secondary Education	0.72***	14.7***	0.67***	22.6***
	(0.06)	(0.01)	(0.08)	(0.03)
Tertiary Education	1.29***	26.6***	1.32***	47.6***
	(0.08)	(0.02)	(0.10)	(0.04)
Reliable supply of LPG	0.84***	17.3***	0.78***	23.9***
	(0.05)	(0.01)	(0.08)	(0.02)
Reliable supply of Kerosene	-0.10	-2.1	-0.09	-2.7
	(0.06)	(0.01)	(0.07)	(0.02)
Reliable supply of Charcoal	-0.07	-1.4	-0.15	-4.7
	(0.08)	(0.02)	(0.12)	(0.04)
Reliable supply of Firewood	-0.44***	-9.1***	-0.52***	-15.5***
	(0.06)	(0.01)	(0.08)	(0.03)
Urban	0.32***	6.5***	0.52***	14.2***
	(0.07)	(0.02)	(0.09)	(0.02)
Electricity	0.91***	18.7***	0.88***	20.4***
	(0.06)	(0.02)	(0.12)	(0.02)
HHsize	Smooth		-0.01	-0.4
			(0.02)	(0.004)
Age_hhd	Smooth		-0.01***	-0.2***
			(0.002)	(0.001)
Log income	Smooth		0.06***	1.8***
			(0.02)	(0.01)
Log price of LPG	Smooth		-0.54***	-16.0***
			(0.16)	(0.05)
Log price of firewood	Smooth		-0.07	-2.0
			(0.07)	(0.02)
Log price of Charcoal	Smooth		0.15	4.3
			(0.15)	(0.04)
Log price of Kerosene	Smooth		-0.12***	-3.5*
			(0.07)	(0.02)
AIC	3698.17		3898.78	
Log Likelihood	-1797.23		-1931.39	
GCV score	-0.25		…….	
Test of linear restrict. for HHsize	[30.35]***			
Test of linear restrict. for Age_head	[38.17]***			
Test of linear restrict. for log income	[97.82]***			
Test of linear restrict. for log P. (LPG)	[32.50]***			
Test of linear restrict. for log P.(firewood)	[11.33]*			
Test of linear restrict. for log P.(charcoal)	[17.96]**			
Test of linear restrict. for log P.(kerosene)	[39.36]***			
Num. obs.	4918		4918	
Num. smooth terms	7		0	

Note:*** , **and*denote the significance level at the 0.1%,1% and 5%,respectively, the values under the heading "marginal effects" are multiply by 100 to convert into percentage points, smooth indicate that the variable enter the model non-parametrically. Numbers in square brackets are chi-square statistics for the test of linear restriction on the smooth variables, while standard errors are in parenthesis. AIC is the Akaike information criteria and GCV score stand for generalized cross validation score. P stands for price and restrict for restriction.

The marginal effect of age of the household head on the other hand is positive but decreasing between the ages 18 to 30 years and negative for all ages older than 30 years. The estimated marginal effect for income, LPG price, firewood price and charcoal price also indicated a non-linear and varying effect on the choice of LPG among Ghanaian households. The sizes of the estimated marginal effect from the semi-parametric model are very different from those from the parametric model. For instance, the estimated marginal effect on log income is 1.8% (parametric model), which is much smaller than the estimated mean marginal effect from the semi-parametric model (19.5%), which we argue is due to the functional restriction imposed on the parametric probit model. This arguably provide some support to the importance of flexible modelling approaches as they are able to accommodate all forms of non-linearity that the parametric model cannot reveal and it is this aspect of flexibility that the semi-parametric have an edge over the parametric model.

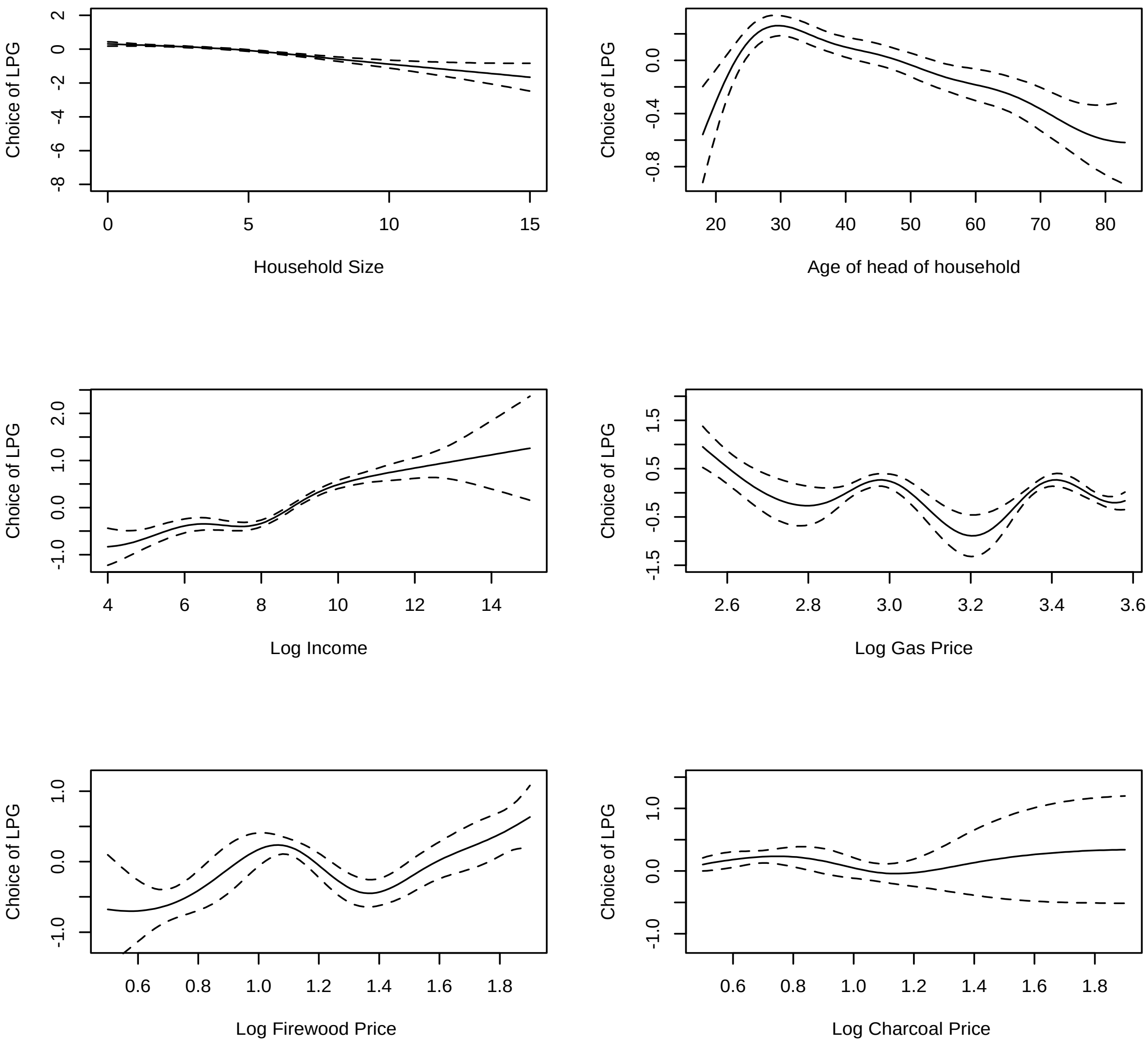

Figure 3: Plot of smooth estimates (GLSS 6 with price effects) with 95% Bayesian confident interval represented by the dotted lines, while the solid lines are the estimated smooth function for each of the variables.

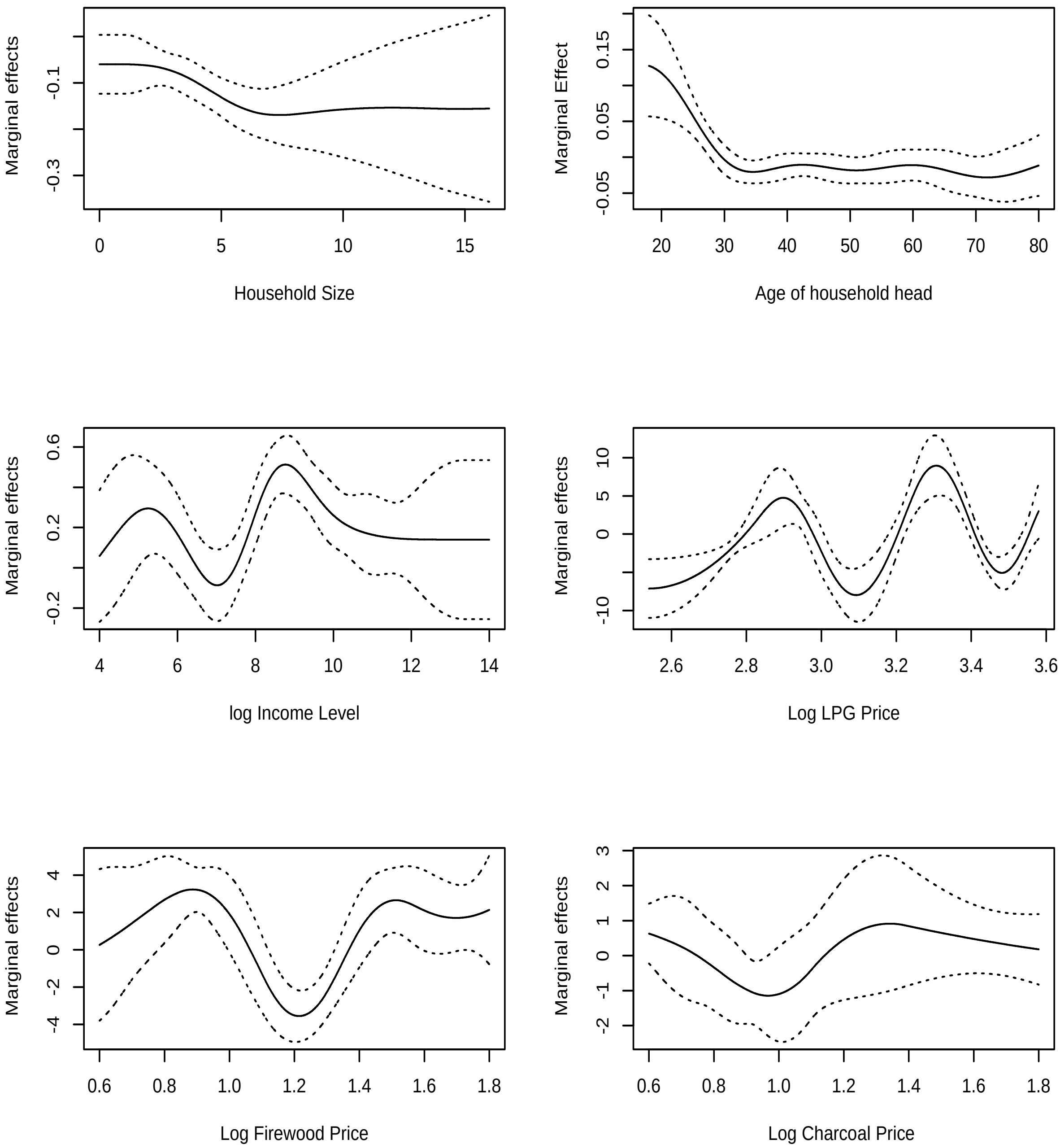

Figure 4: Plot of gradient estimates (marginal effects) for the non-parametric variables. The 95% Bayesian confident intervals are represented by the dotted lines, while the solid lines are the estimated marginal effects.

6. Summary and policy implications

In this paper, we analysed the factors that influence the choice for LPG as a cooking fuel in Ghanaian household. The estimated model is based on previous studies such as (e.g., Mekonnen and Köhlin, 2008; Barnes, et al., 2005; Heltberg, 2005, Karimu, 2015; Mensah and Adu, 2015) derived from the household production theory. We examined the four key objectives of our study - (1) if the key determinants for cooking fuels for developing countries such as income, education, access to modern infrastructure, location and household characteristics as commonly cited in previous literature are consistently important for LPG fuel choice in Ghana for two separate periods, (2) whether the effects of the key variables vary across time, utilising repeated cross-sectional data, (3) whether the usual parametric probit model is free from functional specification bias and (4) the role of price effect on the choice process for LPG in Ghana, via both parametric and semi-parametric discrete choice econometric techniques. The result revealed the following; (a) that education, income, household characteristics, access to urban infrastructure and urban location are consistently important factors in the choice process for LPG for cooking in Ghana, irrespective of the sampled period. Reliable supply of some of the fuels is also an important factor, especially the reliable supply of LPG. Additionally, in general, there are differences in the size of the estimated marginal effects between the 2005/06 and 2012/13 periods, especially in the parametric probit model. (b) We found, consistently, that imposing parametric structure on age, income and household size prior to estimation is not appropriate in the choice model for LPG, irrespective of the sample period. More so, parametric restriction tends to bias the estimated marginal effects, especially for the educational covariates. (c) Last but not the least, we found a significant price effect on the adoption probability for LPG consistent with previous studies (e.g., Heltberg, 2004, 2005). Our estimated price effects are however different from previous literature in the sense that we found also that the price effects varies across the level of prices. Also, comparing the marginal effects from the model with price effect with that without price effect based on the 2012/13 sample indicated that, omitted price bias is not severe.

The results from the study indicated that LPG use is positively related to income and it consistently hold irrespective of sample period, functional form and whether prices of the various fuels are included as covariates or not. Also, education is positively related to the adoption of LPG as a cooking fuel relative to the alternative fuels, and more importantly, increasing levels of education is associated with switching into LPG. Urban location and access to modern infrastructure are positively associated to LPG use, while household size is generally negatively associated to the choice of LPG as a cooking fuel. Other significant variables in the LPG switching process are; reliable supply of LPG, reliable supply of firewood, LPG price, firewood price and charcoal price.

The general policy implication from the price effect of the various fuels on LPG uptake is that, a subsidy on LPG in addition to a tax on firewood and charcoal is likely to promote a switch into LPG. However, such a policy will only yield the desire outcome if the initial cost constraints in the switching process are reduced significantly, especially the cost of LPG cylinders and cost of stoves. Also, cultural factors that serve as berries to LPG uptake will have to be taken into consideration in the design of the subsidy policy. It is also important to recognize the fact that, fuel choice is a complex process and depends on many factors apart from fuel prices and therefore the energy policy should go beyond subsidies. For instance in poor rural areas with cheaply available firewood, improved cookstoves should be the policy goal rather than fuel substitution. In areas where firewood is expensive, especially in urban centers, a fuel substitution policy is likely to be successful if efforts are made to provide cooking technologies that are cost effective and that the technologies are designed to meet the taste and cultural preferences of the people.

Reference

Akpalu W., Dasmani I., Aglobitse P.B (2011). Demand for cooking fuels in a developing country: To what extent do taste and preferences matter? *Energy Policy*; 39:6525–6531.

Adkins E, Eapen S, Kaluwile F, Nair G, Modi V. 2010. Off-grid energy services for the poor: introducing LED lighting in the Millennium Villages Project in Malawi. *Energy Policy* 38:1087–1097.

Amacher GS, Hyde WF, Joshee BR. 1992. The adoption of consumption technologies under uncertainty: a case of improved stoves in Nepal. *J Econ Dev* 17(2):93–105.

Amacher GS, Hyde WF, Kanel KR. 1996. Household fuelwood demand and supply in Nepal's Tarai and mid-hills: choice between cash outlays and labor opportunity. *World Dev* 24(11):1725–1736.

Arthur MFSR, Zahran S, Bucini G. 2010. On the adoption of electricity as a domestic source by Mozambican households. *Energy Policy* 38(11):7235–7249.

Barnes, D., & Qian, L. (1992). *Urban Interfuel substitution, energy use, and equity in developing countries: some preliminary results*. Industry and energy department Working Paper, Energy series paper #53. The World Bank, Washington, DC.

Barnes, D., Krutilla, K., Hyde, W., 2005. The Urban Household Energy Transition: Social and Environmental Impacts in the Developing World. Resources for the Future, Washington, DC.

Becker, G.S., 1965. A Theory of the Allocation of Time, in G.S. Becker, The Economic Approach to Human Behavior, pp. 89-114.

Chambwera, M. 2004. Economic analysis of urban fuelwood demand: the case of Harare in Zimbabwe. PhD diss., Department of Environmental Sciences, Wageningen University.

Chambwera, Folmer, 2007. Fuel switching in Harare: an almost ideal demand system approach. *Energy Policy* 35, 2538–2548.

Chaudhuri S, Pfaff ASP. 2003. Fuel-Choice and Indoor Air Quality: A Household-Level Perspective on Economic Growth and the Environment. New York: Department of Economics and School of International and Public Affairs, Columbia University.

Couture S., Garcia, S., Reynaud A., (2012)Household energy choices and fuelwood consumption: An econometric approach using French data. Energy Economics 34 1972–1981

Davis, M., 1998. Rural household energy consumption: the effects of access to electricity—evidence from South Africa. *Energy Policy* 26(3), 207–217.

Damte A, Koch SF. 2011. Clean Fuel Saving Technology Adoption in Urban Ethiopia. Department of Economics Working Paper Series. Pretoria, South Africa: University of Pretoria.

Ditlhale, N. and M. Wright. 2003. The Importance of Gender in Energy Decision making: the case of Rural Botswana. Journal of Energy in Southern Africa, Vol. 14 No. 2.

Dowd, J. (1989). *The urban energy transition and interfuel substitution in developing countries: a review of the literature.* Internal Report, ESMAP, World Bank, Washington, DC.

Edwards J.H.Y, Langpap, C. 2005. Startup costs and the decision to switch from firewood to gas fuel. *Land Econ* 81(4):570–586.

Efron, B.,1978. Regression and anova with zero-one data: Measures of residual variation. *Journal of the American Statistical Association* 73, 113-121.

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

Green, P.J. and Silverman, B.W., 1994. *Nonparametric Regression and Generalized Linear Models*. London: Chapman and Hall.

Gundimeda H, Köhlin G. 2008. Fuel demand elasticities for energy and environmental policies: Indian sample survey evidence. *Energy Economics* 30(2):517–546.

Gupta G, Köhlin G. 2006. Preferences for domestic fuel: analysis with socio-economic factors and rankings in Kolkata, India. *Ecological Economics* 57(1):107–121.

Hastie, T.J. and Tibshirani, R.J.,1990. *Generalized Additive Models.* London: Chapman and Hall.

Heltberg, R., 2004. Fuel switching: evidence from eight developing countries. *Energy Economics* 26 (5), 869–887.

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

Hofstad O, Köhlin G, Namaalway F. 2009. How can emissions from woodfuel be reduced? In: Realising REDD+: National Strategy and Policy Options (Angelsen A, Brockhaus M, Kanninen M, Sills E, Sunderlin WD, Wertz-Kanounnikoff S, eds). Bogor, Indonesia:Center for International Forestry Research, 237–248.

Hyde W. and G. Kohlin, 2000. Social Forestry Reconsidered. *Silva Fennica* 34(3):285-315

Hosier RH, Dowd J. 1987. Household fuel choice in Zimbabwe: an empirical test of the energy ladder hypothesis. *Resource and Energy* 9(4):347–361.

Jack DW. 2006. Household Behavior and Energy Demand: Evidence from Peru [PhD Dissertation]. Cambridge, MA: Harvard University.

Karimu, A., 2015. Cooking fuel preferences among Ghanaian Households: An empirical analysis. *Energy for Sustainable Development* 27, 10-17.

Kavi Kumar KS, Viswanathan B. 2007. Changing structure of income indoor air pollution relationship in India. *Energy Policy* 35(11):5496–5504.

Kebede, B., Bekele, A., Kedir, E., 2002. Can the urban poor afford modern energy? The case of Ethiopia. *Energy Policy* 30, 1029–1045.

Khandker SR, Barnes DF, Samad HA. 2010. Energy Poverty in Rural and Urban India: Are the Energy Poor Also Income Poor? Policy Research Working Paper No. 5463.Washington, DC:World Bank.

Lamarre-Vincent J. 2011. Household Determinants and Respiratory Health Impacts of Fuel Switching in Indonesia [Master's Thesis]. Durham, NC:Duke University.

Lewis J.J., and Pattanayak S.K (2012). Who Adopts Improved Fuels and Cookstoves? A Systematic Review. *Environmental Health Perspectives* 120:637–645

Louw K, Conradie B, Howells M, Dekenah M. 2008. Determinants of electricity demand for newly electrified low-income African households. Energy Policy 36:2812–2818.

McEachern M, Hanson S. 2008. Socio-geographic perception in the diffusion of innovation: solar energy technology in Sri Lanka. *Energy Policy* 36(7): 2578–2590.

Mekonnen, A., Köhlin, G., 2008. Determinants of household fuel choice in major cities in Ethiopia. RFF Discussion Paper EfD 08–18, Environment for Development (EfD) Initiative and Resources For the Future (RFF), Washington DC, August 2008. http://www.rff.org/RFF/Documents/EfD-DP-08-18.pdf

Mensah, J.T. and Adu, G., 2015. An Empirical Analysis of Household Energy Choice in Ghana. *Renewable and Sustainable Energy Review*, 51: 1401-1411.

Masera, O.R., Saatkamp, B.D., and Kammen, D.M., 2000. From linear fuel switching to multiple cooking strategies: a critique and alternative to the energy ladder model. *World Development*, 28 (12): 2083-2103.

Muneer E.T, Mohamed S.M.W. 2003. Adoption of biomass improved cookstoves in a patriarchal society: an example from Sudan. *Sci Total Environ* 307(1–3):259–266.

Ouedraogo B. 2006. Household energy preferences for cooking in urban Ouagadougou, Burkina Faso. *Energy Policy* 34(18):3787–3795.

Peng W, Hisham Z, Pan J. 2010. Household level fuel switching in rural Hubei. *Energy Sustain Dev* 14(3):238–244.

Pine K, Edwards R, Masera O, Schilmann A, Marrón-Mares A, Riojas-Rodríguez H. 2011. Adoption and use of improved biomass stoves in Rural Mexico. *Energy Sustain Dev* 15(2):176–183.

Ramanathan V, Carmichael G. 2008. Global and regional climate changes due to black carbon. *Nat Geosci* 1(4):221–227.

Rao MN, Reddy BS. 2007. Variations in energy use by Indian households: an analysis of micro level data. *Energy* 32(2):143–153.

Rebane KL, Barham BL. 2011. Knowledge and adoption of solar home systems in rural Nicaragua. *Energy Policy* 39(6):3064–3075.

Reddy BS. 1995. A multilogit model for fuel shifts in the domestic sector. *Energy* 20(9):929–936.

Wahba, G., 1990. Spline models for observational data. *Regl Conf.Ser.APPL.Math.*, 59.

Walekhwa PN, Mugisha J, Drake L. 2009. Biogas energy from family-sized digesters in Uganda: critical factors and policy implications. *Energy Policy* 37(7):2754–2762.

Wendland K.J, Pattanayak S.K, Sills E. 2011. Democracy and Dictatorship: Comparing Household Innovation across the Border of Benin and Togo. Raleigh, NC: North Carolina State University, Department of Forestry and Natural Resources.

Wood, S.N., 2000. Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society*. Series B 62(2), 413–428.

Wooldridge, J.M., 2002. *Econometric Analysis of cross section and panel data*. Cambridge, MA, MIT Press.

World Energy Council (WEC),1999.The Challenge for Rural Energy Poverty in Developing Countries. WEC/FAO Report.

Yan HJ. 2010. The Theoretic and Empirical Analysis on the Compatibility of Sustainable Development Strategies and Poverty Reduction Policies at Micro Level. Aix-en-Provence, France:Université de la Méditerranée Aix-Marseille II.

Appendix:

Table A: List of variables and definitions

Variables	Type of Variable
Male head of the household (Male head)	Dummy (Male=1, Female=0)
Age of household head (age_hdd)	Continuous
Log annual household income (Lincome)	Continuous
Household size(hhsize)	Continuous
Basic education (Basic Edu.)	Dummy (Basic=1,No formal education=0)
Secondary education (Secondary Edu.)	Dummy (Secondary=1,No formal education=0)
Tertiary education (Tertiary Edu.)	Dummy (Tertiary=1,No formal education=0)
Urban location (Urban)	Dummy (Urban=1, Rural =0)
Electricity as the main lightening source	Dummy (electricity=1, Others=0)
LPG	Dummy (LPG=1, Others=0)
Reliable supply of LPG (Rel. LPG)	Dummy (Yes=1, No=0)
Reliable supply of kerosene (Rel. Kero)	Dummy (Yes=1, No=0)
Reliable supply of charcoal (Rel. Char)	Dummy (Yes=1, No=0)
Reliable supply of fuelwood (Rel. Firewd)	Dummy (Yes=1, No=0)
Price of LPG (lnprice_LPG)	Continuous
Price of firewood (lnprice_fire)	Continuous
Price of Charcoal (lnprice_charc)	Continuous
Price of Kerosene (lnprice_kero)	Continuous

Table A1: Summary statistics for GLSS5 (2005/06) and GLSS6 (2012/13) samples

Variable	Obs	Mean	SD	Min	Max
		2005/06 Sample			
LPG	8262	0.09	0.29	0.00	1.00
hhsize	8262	4.33	2.82	1.00	29.00
Age_hhd	8262	45.65	15.46	15.00	99.00
Lincome	8262	15.60	1.26	7.63	20.08
Male head	8262	0.71	0.45	0.00	1.00
Basic Edu.	8262	0.30	0.46	0.00	1.00
Secondary Edu.	8262	0.10	0.30	0.00	1.00
Tertiary Edu.	8262	0.04	0.19	0.00	1.00
Rel. LPG	8262	0.03	0.18	0.00	1.00
Rel. Kero	8262	0.15	0.36	0.00	1.00
Rel. Char	8262	0.14	0.35	0.00	1.00
Rel. Firewd	8262	0.10	0.29	0.00	1.00
Urban	8262	0.40	0.49	0.00	1.00
Electricity	8262	0.44	0.50	0.00	1.00
		2012/13 Sample			
	Obs	Mean	SD	Min	Max
LPG	16772	0.17	0.37	0.00	1.00
hhsize	16772	4.26	2.78	1.00	29.00
Age_hhd	16772	45.84	15.90	15.00	99.00
Lnincome	15212	8.14	1.80	-0.96	14.57
Male head	16772	0.72	0.45	0.00	1.00
Basic Edu.	16772	0.45	0.50	0.00	1.00
Secondary Edu.	16772	0.11	0.31	0.00	1.00
Tertiary Edu.	16772	0.08	0.27	0.00	1.00
Rel. LPG	16759	0.22	0.42	0.00	1.00
Rel. Kero	16772	0.45	0.50	0.00	1.00
Rel. Char	16772	0.78	0.41	0.00	1.00
Rel. Firewd	16772	0.57	0.50	0.00	1.00
Urban	16772	0.44	0.50	0.00	1.00
Electricity	16772	0.61	0.49	0.00	1.00
lnprice_LPG	16772	1.03	1.53	0.00	4.45
lnprice_fire	16772	0.27	0.53	0.00	2.40
lnprice_charc	16772	0.36	0.37	0.00	2.56
lnprice_kero	16772	0.33	0.55	0.00	3.47

Note: electricity denotes electricity as the main lightening source, SD stand for standard deviation, Obs for number of observation, Min for minimum and Max for maximum.

Table B: Semi-parametric estimates without price effects for GLSS6 data with the same number of observation as the data with prices.

Dept.Variable: LPG (1, 0)	**Betas**	**Marginal Effects**
(Intercept)	-2.13***	
	(0.13)	
Male head	-0.21***	-4.5**
	(0.05)	(0.01)
Basic Education	0.39***	8.3***
	(0.06)	(0.01
Secondary Education	0.76***	16.2***
	(0.06)	(0.01)
Tertiary Education	1.31***	27.9***
	(0.08)	(0.02)
Reliable supply of LPG	0.75***	15.9***
	(0.05)	(0.01)
Reliable supply of Kerosene	-0.04	-0.9
	(0.05)	(0.01)
Reliable supply of Charcoal	-0.07	-1.5
	(0.08)	(0.02)
Reliable supply of Firewood	-0.52***	-11.1***
	(0.05)	(0.01)
Urban	0.36***	7.7***
	(0.06)	(0.02)
Electricity	0.93***	19.8***
	(0.10)	(0.02)
HHsize	Smooth	
Age_hhd	Smooth	
Log income	Smooth	
AIC	3800.16	
Log Likelihood	-1874.26	
GCV score	-0.23	
Test of linear restrict. for HHsize	[29.98]***	
Test of linear restrict. for Age_head	[39.08]***	
Test of linear restrict. for log income	[93.65]***	
Num. obs.	4918	
Num. smooth terms	3	