A Structural Model of Cooking Fuel Choices in Developing Countries

Miguel Poblete-Cazenave, Shonali Pachauri

International Institute for Applied Systems Analysis (IIASA),
Schlossplatz 1, A-2361 Laxenburg, Austria

Post-print Version

Published in Energy Economics 75: 449-463. DOI:10.1016/j.eneco.2018.09.003.

1 Introduction

There are many dimensions of poverty, one of them related to the availability and accessibility of different fuel options for cooking. Approximately forty percent of the world's population uses solid fuels for cooking, such as firewood and charcoal (International Energy Agency (IEA) and the World Bank, 2017). These fuels, along with the use of rudimentary stoves, creates a series of problems because of poor fuel quality and incomplete combustion. In particular, an estimated 2.6 million people worldwide die prematurely (Health Effects Institute, 2018) because of air pollution caused mainly by the use of poor quality fuels in rudimentary stoves within household premises. Several efforts to improve the adoption of modern cooking fuels and stoves have been implemented, especially in developing regions, to reduce the risks associated with the use of low quality fuels and stoves. In order to analyze the potential impact of such policies ex ante, and project possible future scenarios of clean fuel adoption, several models of household cooking fuel choices have been developed in the past decades. However, most existing models are based either on the assumption that there is an "energy ladder" and households ascend this ladder i.e. move to using cleaner, more expensive fuels, as their income rises, (e.g. OTA 1992; Hosier and Dowd 1987; Kowsari and Zerriffi 2011; van der Kroon et al. 2013), or that the adoption of cleaner fuels is gradual as income increases, and households "stack" their fuel options (e.g. Masera et al. 2000; Cheng and Urpelainen 2014; Smith and Sagar 2014).

Here, following on Ekholm et al. (2010) and Cameron et al. (2016), we present the latest version of the MESSAGE-Access model, a behavioral choice model to estimate the demand and choices for household cooking fuels. Unlike other models in the literature, we make no explicit a priori assumptions about preferences between fuels, that is, we do not assume either an "energy ladder" or a "stacking" theory for the transition. We estimate our model using the Method of Simulated Moments (McFadden, 1989) on data for Ghana, Guatemala, India, Nigeria and Uganda [1]. We find that our model estimates of the pattern of fuel adoption by income are a close match to the empirical data derived from the surveys in all the selected countries. We also undertake ex post simulations using the estimated parameters of the model to test the responsiveness of demand to variations in the price of fuels and the level of per capita income.

The rest of the paper is organized as follows. In Section 2, we review literature on models of household fuel choices that have been applied in scenario analysis. In section 3, we present our theoretical model of household cooking fuel choices. In section 4, we discuss the datasets used in the study and present our estimation methods and results. In section 5, we use the model to assess the responsiveness of fuel demands to changes in prices and income. Finally, Section 6 concludes.

2 Literature on household fuel choices

Several empirical analysis of the determinants of household energy choices in developing countries can be found in the literature (e.g. Campbell et al. 2003; Heltberg 2004, 2005; Alem et al. 2016). Many of these rely on statistical analysis using multiple linear regressions or discrete choice models. Recent reviews of the literature point to the fact that much of the evidence on factors influencing household choices remains largely scattered and qualitative and that

[1] India was selected as representative of South and South East Asia and Guatemala as representative of Central America, whereas the remaining countries are representative of Sub-Saharan Africa. These geographic regions represent those with the largest concentration of biomass dependent populations for cooking in the world.

quantitative analysis is constrained by the lack of sufficient data, especially on energy prices and expenditures (Lewis and Pattanayak, 2012). Here we do not attempt to undertake a comprehensive review of literature on household fuel choices, but focus specifically on studies that have taken a forward looking perspective to analyze future scenarios of cooking fuel transitions.

As noted in Section 1, this model is developed as the next step in the evolution of the MESSAGE-Access framework. Ekholm et al. (2010) present the earliest version of this model. In this early version of the model, households face a utility maximization problem that translates into an equation that represents a choice between fuel alternatives based on a trade-off between inconvenience costs of different fuels and differences in actual costs and prices. The main drawback of this approach is that, as a result of this being a linear choice model, households choose only one fuel among the alternatives, something that is in contrast with empirical evidence. To address this issue, Cameron et al. (2016) provides a second version of the MESSAGE-Access model, where households are allowed to stack multiple fuel options. To this end, demand curves for clean fuels are estimated for different population subgroups, as well as their total demand for cooking fuel. Based on the estimated demand curves, households are assumed to first choose cleaner fuel options up to the point that these are affordable at the given prices they are available at. Afterwards, if the total household fuel demand has not been fulfilled, the remaining is met with non-clean fuel options. Although a significant improvement compared to the earlier model, the key problem of this method is that the only estimated demand is for clean fuels, and therefore, there is no demand response to changes in prices of non-clean fuels. In addition to this, and because for estimation purposes the population is divided among subgroups, there is a weak response to changes in income, as the model has no clear substitution effect between clean and non-clean fuel options.

We also discuss some of the alternative models that have been proposed in the literature. In van Ruijven et al. (2011), we find the first use of the IMAGE-REMG model to estimate household fuel choices for multiple end-uses such as cooking, lightning and heating. In particular for the cooking module, the model assumes a constant level of household cooking energy demand and then uses a multinomial logit model to estimate preferences for different available fuel alternatives. In this model too, the population is separated in groups and there is an assumption that cleaner fuels are always chosen first. Therefore, the same issues arise. Additionally, although the use of a multinomial logit model for this purpose seems natural, it severely limits the applicability of this model to other countries as it requires enough data to make the estimated coefficients significant. For example, in Sub-Saharan countries where the adoption of clean fuels is still lagging, it would be hard to correctly estimate preferences for these fuels. Finally, with a multinomial logit model we can only obtain the probabilities of choosing between mutually exclusive alternatives. Therefore, under such assumption, individual households can choose only one of the possible fuel alternatives, something that is at odds with empirical data.

A recent study of Sub-Saharan African countries that uses a multinomial model approach is Rahut et al. (2016). To overcome the problem of data scarcity, it merges data from three different countries in the estimation process. However, the reduced-form nature of this paper limits its applicability for an actual modeling of household choices. For example, some of the control variables would be hard to project for future scenarios (e.g. distance to markets). Moreover, two critical factors are not included: fuel prices and household income. It would be interesting to see whether an estimation including these factors would render significant coefficients that could be used to estimate responses to them. Nevertheless, as with all atheoretical approaches, its appropriateness for scenario analysis will always be limited, as the estimated parameters are only valid as statistical descriptions of the data, while the mechanisms behind the choice decision remain obscure (Koopmans, 1947; Heckman, 2008; Keane, 2010).

Recently, Fuso Nerini et al. (2017) propose a new approach, which they refer to as "levelized cost of cooking a meal", that involves calculating the cost of cooking with different fuel-technology combinations. This only allows for a comparison between the cost of cooking using a calculated predefined level of energy using a variety of available options. Though the method does include responses to changes in price, it does not include responses to changes in income. Therefore, as a consequence, it possesses limited adequacy for scenario analysis.

3 Model

We present a parsimonious model of household choice between cooking fuels and consumption of other goods, subject to a budget constraint. Households choose cooking fuels according to their preferences for consumption of other household goods and each of the available fuel options. Individual households are considered to be price-takers, and therefore, prices are assumed to be exogenous. In particular, we assume a Cobb-Douglas utility function, such that:

$$\max_{C,F} U(C,F) = \left[C^{\alpha} \left(\sum_{f=1}^{N_f} e_f F_f \right)^{1-\alpha} \right]^{\gamma} \left[\chi \left(F_1 ... F_{N_f} \right) \right]^{1-\gamma} \tag{1}$$

s.t.

$$p_c C + \sum_{f=1}^{N_f} (p_f F_f + \mathbb{A}_f) = I \tag{2}$$

$$C, F_f \geq 0 \tag{3}$$

where C is consumption of other items, F_f is cooking fuel consumption of fuel f, $\mathbb{A}_f$ is an annualized value of the cooking stove of fuel f and I is income (or, a better proxy, expenditure). $\chi \left(F_1 ... F_{N_f} \right)$ is a function that represents the household preferences for each of the available fuel options, or, if we think in dual terms, the implicit "inconvenience cost" of the fuels used by the household (e.g. collection costs and health costs)[2]. The unknown parameters that we need to estimate are the preference for fuel consumption vs consumption of other items α, and overall consumption vs implicit cost γ. Additionally, we have to make assumptions about the function representing the preferences for each fuel χ. In particular, we can assume that this function is a second degree polynomial on each of the fuels:

$$\chi = \chi_0 - \sum_{f=1}^{N_f} \left(\delta_{1f} F_i + \delta_{2f} F_i^2 \right) - \mathbb{K} \tag{4}$$

[2] We can think of this function as an analogue of the preference for leisure / disutility of work in a standard labor supply model, that is, a function that represents the need to produce something in order to increase utility (in this case, food), but it has some associated non-monetary costs depending on the particular choice that is made.

This may be the strongest parametric assumption in the model. However, we found empirical support for it mainly from the fact that households, as stated in the previous sections, do not choose a single fuel for their energy needs. Without including a non-linear implicit cost function households would only choose one fuel among the available options. Also, as pointed out in the solution of the model (see Appendix A), the household optimization problem will have a unique solution if the implicit cost function is a strictly concave function of the fuel options. Therefore, we opt to use the simplest possible strictly concave function available, i.e., a second degree polynomial, to avoid imposing a heavier additional structure on the model. Finally, we also include a fixed cost $\mathbb{K}$ that becomes non zero only when firewood is collected for free.

Assuming that we know the parameters, the model is solved as a constrained optimization problem (see Appendix A). In the end, the total quantity of fuel $\tilde{f}$ demanded by the household comes implicitly from:

$$F_{\tilde{f}} = \begin{cases} \frac{1}{2\delta_{2f}} \left[\frac{\gamma(1-\alpha)e_{\tilde{f}}}{\sum_{f=1}^{N_f} e_f F_f} - \frac{\gamma\alpha p_{\tilde{f}}}{I - \sum_{f=1}^{N_f}(p_f F_f + \mathbb{A}_f)} - \frac{(1-\gamma)\delta_{1f}}{\chi} \right] & \text{if } \mu_i = 0 \\ 0 & \text{if } \mu_i > 0 \end{cases} \tag{5}$$

where $\mu_{\tilde{f}}$ is the Lagrange multiplier associated to the $F_{\tilde{f}} \geq 0$ condition.

4 Data and estimation

The basic concept underlying the structural simulated method of moments is to set up a theoretical model to represent an economic decision and use data to find primitive parameters of the model that would explain real life observations. This is done by generating a simulated dataset using the model, and then matching the moments estimated from the simulated data to the moments of the observed data. In this case, the following parameters from the model are unknown and need to be estimated: α, γ, $\mathbb{K}$ and δ_{1f}, δ_{2f} for all the fuel options.

The model was estimated independently for 3 countries that can be used to represent Sub-Saharan Africa, for Guatemala, representative of Central America and for India, representative of South Asia. Table 1 summarizes the data sources used for this study. All datasets

Country	Dataset	Years	Obs*	Perc**
Ghana	Ghana Living Standards Survey (GLSS)	2012-2013	7,039	49.2%
Guatemala	Encuesta Nacional de Condiciones de Vida (ENCOVI)	2014	6,737	61,2%
India	National Sample Survey (NSS)	2011-2012	62,201	49.8%
Nigeria	General Household Survey (GHS)	2012-2013	2,104	51.7%
Uganda	Uganda National Household Survey (UNHS)	2012-2013	2,472	27.0%

* Number of observations after data cleaning.

** Weighted percentage of the total sample.

Table 1: Datasets Used by Country

were subjected to the same data cleaning processes, which consisted of excluding outliers[3] in expenditure per household per capita, cooking fuel consumption and cooking fuel consumption over expenditure per household per capita. Additionally, household fuel choices were cross checked with the possession of an appropriate stove for the fuel (e.g. electricity use for cooking with the possession of an electric stove). Finally, households where the identification of fuel usage for cooking was unclear were also dropped. Sample weights were included and used for the calculation of the observed moments. A summary of the data cleaning process, as well as some descriptive statistics can be found in Table 2.

	Ghana	Guatemala	India	Nigeria	Uganda
Data Cleaning					
Initial Numbers of Households	16,772	11,563	101,662	4,728	6,891
Outliers in Cooking Expenditure (including missing)	7,331	7,017	82,372	2,192	2,585
Outliers in Household Expenditure	7,183	6,785	80,724	2,148	2,533
Outliers in Cooking Consumption over Expenditure	7,039	6,737	62,201	2,104	2,472
Descriptive Statistics*					
Percentage Urban	69.2%	57.4%	52.6%	51.3%	50.6%
Mean Household Size	5.4	5.5	5.6	7.6	6.7
Mean Household Expenditure**	12,293.6	14,158.3	7,435.9	12,386.9	6,755.3
Mean Firewood Expenditure**	16.2	24.6	60.6	120.4	87.4
Mean Charcoal Expenditure**	150.7	0.3	6.9	12.1	175.8
Mean Kerosene Expenditure**	0.7	0.2	26.6	180.9	10.6
Mean LPG Expenditure**	36.1	18.7	145.8	17.1	4.2
Mean Electricity Expenditure**	0.2	0.7	0.8	0.2	1.5

* Additional statistics can be found in the data-related moments.

** All expenditures are in 2010 USD.

Table 2: Data Cleaning Process and Descriptive Statistics

All of these datasets present similar difficulties. First, expenditure information and information on quantities consumed by each household were not available, except in the datasets of

[3] i.e. 1%-tile bottom or top observations, as well as households with no observations.

India and Guatemala. For the remaining countries, the datasets used contained additional market modules that contained price information for the surveyed regions, therefore, average regional prices were used as a proxy to calculate household consumption given the expenditure data available. Nevertheless, in some cases, price information was given in quantities that are ambiguous (for example, firewood quantities were sometimes presented in quantities such as "bundles" or "bunches"). To solve this, we used a variety of external sources to find representative units for the quantities (e.g. how many kilos is a bunch), and then, we tested whether this unit-corrected prices would imply aggregate consumption levels that are consistent with national energy statistics of each country. Additionally, stove prices and efficiencies where not available, therefore, similar assumptions as in Cameron et al. (2016) were followed. In Table 5 in the Appendix we show the stove options used in the model, which are representative of the most used stoves for each fuel option[4]. Finally, since it is not always possible to distinguish whether the fuels are used for cooking or for other purposes, the following assumptions were made. On the one hand, if electricity was not disclosed as the main cooking fuel source, it was not included into the cooking fuel mix of the household. On the other hand, for households that listed electricity as the main cooking fuel, the average consumption of electricity of households with similar expenditure levels was subtracted as a proxy of electricity usage for other purposes.

Estimation was done using the Method of Simulated Moments. We estimate 26 moments corresponding to mean log total fuel consumption and expenditure of other items per household per capita for the aggregate (to identify α and γ), plus mean log consumption per household per capita per fuel, the mean percentage of each fuel in the total household fuel consumption by rural/urban groups (to identify γ and the δs) and the percentage of firewood that is obtained for free (to identify $\mathbb{K}$). For the estimation we assume that the general preference parameters α and γ are the same for all households, but the parameters of the inconvenience cost function are different between rural and urban households. With this we try to explain the heterogeneity in behavior arising from the differences in the budget constraint of different households, without disregarding the inherent differences in the inconvenience of obtaining or using a particular fuel between urban and rural households.

[4]For a detailed description of these, see the Supplementary Information of Cameron et al. 2016

The steps of the Method of Simulated Moments are the following:

- Calculate the selected moments from the sample observations and construct a column vector with the observed moments, M^o.

- Using an initial guess for the primitive parameters of the model that we are trying to estimate, $\hat{\theta}$, create 10,000 simulated households.

- Putting together initial conditions, shocks and decision rules, get the simulated choices of each of the 10,000 households.

- Obtain the corresponding moments from the simulated data and generate the column vector $M^s(\theta)$.

- Calculate the value of the following criterion function

$$G\big(\hat{\theta}\big) = \Big(M^s(\hat{\theta}) - M^o\Big)' W^{-1} \Big(M^s(\hat{\theta}) - M^o\Big)$$

  where W is a diagonal matrix, where each of the elements of the diagonal represent the inverse of the variance of the corresponding moment estimated in the data.

- Iterate on the parameters $\hat{\theta}$ until $G\big(\hat{\theta}\big)$ is minimized.

Moments matched and parameter estimates for each country can be found in the Appendix.

In Figure 1 we compare the actual data to the results of the model in terms of percentage of each fuel in the total cooking fuel consumption of the household. In all cases, we find that the model provides a close approximation to the observed patterns in the data. Additionally, in Figure 2 we show how the preferences for each fuel change by expenditure level, that is, how the part of the function χ corresponding to each fuel behaves as expenditure increases[5]. From both sets of figures we can see how, as expenditure increases, household switch towards cleaner fuels. These illustrate the rate of transition to cleaner fuels with rises in income. In Figure 2, the parameter on preferences reflects some of the non-economic factors

[5]Namely $\delta_{1f}F_i + \delta_{2f}F_i^2$

that contribute to the inertia in fuel switching, such as tastes, preferences, reliability and ease of fuel supply. As seen from the figure, this clearly differs across nations and between urban and rural households. In general, the level of income at which urban households switch to cleaner fuels, like LPG, is lower than that at which rural households do so. This may reflect the easier access to fuels and stoves in urban centers and the higher opportunity cost of labor in towns and cities.

5 Demand responsiveness to price and income changes

In this section, we use the model to estimate responses to changes in some of the key factors that affect household's fuel choices. We undertake three different simulations, two to test responses to variations in price and one to test responses to changes in income. It is important to note that, in this model framework, the elasticities are not constant, as there are various channels of response to changes in the model parameters. For example, an increase in income would not always increase the consumption of one of the fuels already used by a household by a certain particular amount. It could well be that, after a certain threshold, the higher income pushes the household to switch to a different fuel source altogether. Therefore, here we present some scenarios to show some responses of the model to changes in some of the relevant decision factors, with the caveat that these responses would be noticeably different depending on the size of the changes.

In our first simulation, we increase the prices of biomass fuels (i.e. firewood and charcoal/coal) by 20%. In our second, we reduce the price of LPG, the most used clean fuel, also by 20%. Finally, we undertake simulations in which the average household per capita income is set to either increase or decrease by 20%. In all simulations, everything else is held constant, consistent with the partial-equilibrium nature of the model.

We present the results of these simulations for the overall population and for rural or urban households in Tables 3 and 4 and in Figures 3 to 7. In all cases, the responses are in the expected direction, that is, when biomass price increases, households increase their usage of clean fuels; when LPG price decreases, households increase their usage of this fuel. In addition, when income rises, households are better able to afford more clean fuels[6]. Also, the response of urban households is higher than rural households for changes in biomass prices, whereas is lower to changes in income levels, consistent with the greater availability of clean fuel sources in urban areas.

Additionally, we find that, save in Nigeria, the factor that affects the percentage of clean fuels used by households most significantly is income. Noticeably, the effect of a decrease in income is higher than the effect of an increase in income. Also, not surprisingly, the effect is much stronger in countries where clean fuel adoption is lower. Finally, biomass prices seem to have a lower impact on demand for this fuel, as in the model, households always have the option to gather firewood for free, if convenient. Indeed, in Table 4 we see the effects of these simulations on the demand for freely collected biomass. As expected, an increase in biomass prices as well as a decrease in income leads to an increase in the collection of free firewood. On the contrary, higher income leads to a decrease in the amount of biomass collected for free.

6 Conclusions

Earlier efforts at modeling household fuel choices in developing countries have assumed either a fuel ladder or fuel stacking as the underlying theoretical construct for estimating parameters such as income and price elasticities of demand. In this paper we present the latest version of the MESSAGE-Access model. Unlike previous theoretical models, we do not impose a limit on the amount of cooking fuel used by households nor specify a hierarchy between different fuel options. Also, compared to reduced form discrete choice models, the structural nature of this model makes it more appropriate for scenario analysis. Understanding how consumers choices may change given changes in incomes and energy prices is important for both policy makers and researchers alike. By observing consumers choices and using a structural model form, we provide insights that go beyond what is possible using other methods. In particular, by linking theoretical models and empirical estimation methods, structural models are well suited for analysis of counterfactuals, as several channels of responses are modeled and jointly

[6]Here we only consider LPG and electricity as clean fuels.

Country	Baseline Level*	20% Higher Biomass Prices	20% Lower LPG Price	20% Higher Income	20% Lower Income
Ghana	27.94	4.90	11.99	15.50	-17.97
Guatemala	47.60	7.18	7.92	12.06	-16.39
India	64.14	0.67	6.13	12.25	-16.57
Nigeria	4.04	1.49	46.29	32.92	-33.17
Uganda	1.18	34.75	24.58	38.98	-43.22

(a) Overall Population

Country	Baseline Level*	20% Higher Biomass Prices	20% Lower LPG Price	20% Higher Income	20% Lower Income
Ghana	14.13	4.03	14.51	34.61	-16.42
Guatemala	16.47	12.08	7.29	27.32	-25.26
India	48.52	1.92	8.86	20.71	-22.38
Nigeria	2.33	4.72	79.40	44.21	-40.77
Uganda	0.86	4.65	22.09	36.05	-50.00

(b) Rural Population

Country	Baseline Level*	20% Higher Biomass Prices	20% Lower LPG Price	20% Higher Income	20% Lower Income
Ghana	33.88	5.05	11.10	14.26	-18.03
Guatemala	69.40	6.38	7.56	10.01	-14.50
India	77.63	0.00	4.51	7.87	-13.24
Nigeria	5.60	0.36	33.75	29.11	-30.00
Uganda	1.46	49.32	25.34	40.41	-39.04

(c) Urban Population

Table 3: Percentage Change of the Proportion of Clean Fuel Use in Total Cooking Fuel Consumption for Different Scenarios

(*) Represents the Percentage Level of Clean Fuel Adoption at the Baseline Simulation

estimated to fit relevant characteristics of the data.

The results of the estimation for 5 countries presented in this study shows a close fit to the empirical data. This allows us to undertake simulations to test demand responses to changes in prices and income. The model simulations show results in line with expected behavioral responses. The strong response of demand for clean fuels, like LPG, to income that we observe has important policy implications. It suggests that public policies that provide targeted and social transfers should be explored in addition to traditional support via fuel subsidies. In addition, the responsiveness of self-collected free biomass to both changes in fuel prices and income also hints to the fact that in areas where wood is abundant and freely

Country	Baseline Level*	20% Higher Biomass Prices	20% Lower LPG Price	20% Higher Income	20% Lower Income
Ghana	1.56	47.44	-3.21	-30.77	52.56
Guatemala	12.61	66.38	16.34	-32.04	62.17
India	35.07	26.15	19.93	-26.15	23.13
Nigeria	24.91	8.35	14.05	-19.75	29.51
Uganda	15.74	29.10	3.68	-22.94	37.29

(a) Overall Population

Country	Baseline Level*	20% Higher Biomass Prices	20% Lower LPG Price	20% Higher Income	20% Lower Income
Ghana	4.40	36.59	6.59	-28.18	42.50
Guatemala	12.82	48.44	3.67	-35.96	61.47
India	37.85	23.43	25.15	-28.08	28.96
Nigeria	29.44	8.22	16.51	-20.31	27.07
Uganda	18.46	26.76	5.63	-23.24	34.24

(b) Rural Population

Country	Baseline Level*	20% Higher Biomass Prices	20% Lower LPG Price	20% Higher Income	20% Lower Income
Ghana	0.34	108.82	-47.06	-50.00	100.00
Guatemala	12.21	111.88	48.73	-23.10	64.05
India	16.77	74.00	-20.04	-21.59	16.22
Nigeria	20.07	8.62	11.06	-19.88	34.03
Uganda	13.39	31.81	1.49	-22.70	40.70

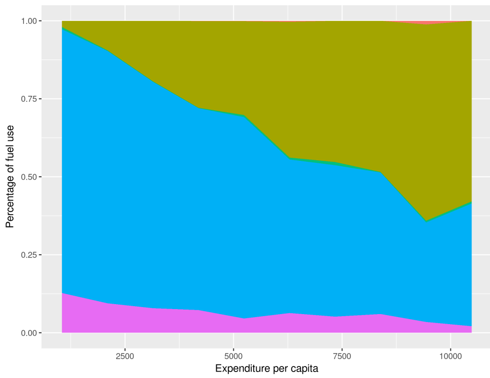
(c) Urban Population

Table 4: Percentage Change of the Proportion of Free Biomass in Total Biomass Consumption for Different Scenarios

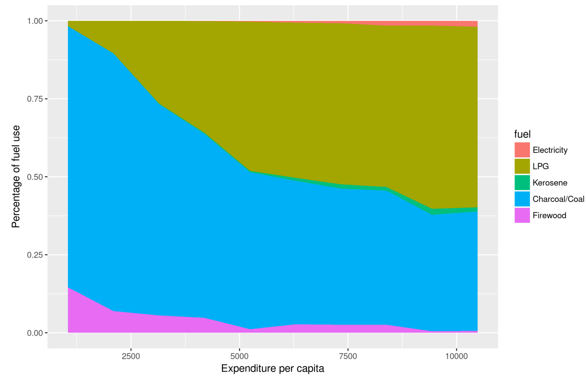
(*) Represents the Percentage of Free Biomass in Total Biomass Consumption at the Baseline Simulation

available, policies that inform and educate people about the adverse impacts of cooking with solid fuels are required.

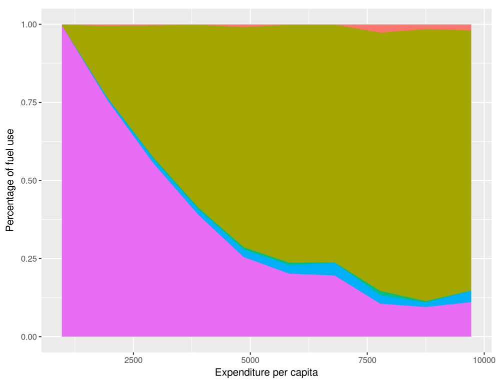
Agreement on the United Nation's Sustainable Development Goals (SDG) is providing greater impetus for achieving universal access to clean cooking by 2030. This requires analysis of household fuel choices and assessments of policy scenarios that can facilitate this. Overall, the results of this study shows that the MESSAGE-Access model can be used as a powerful policy tool for scenario analysis where a multiplicity of conditions change simultaneously. In particular, the model could be used to assess the effects of alternative policy instruments for accelerating a transition to cleaner fuels, the policy costs of such efforts, and the implications of such transitions for other sustainable development goals such as those pertaining to health and the environment.



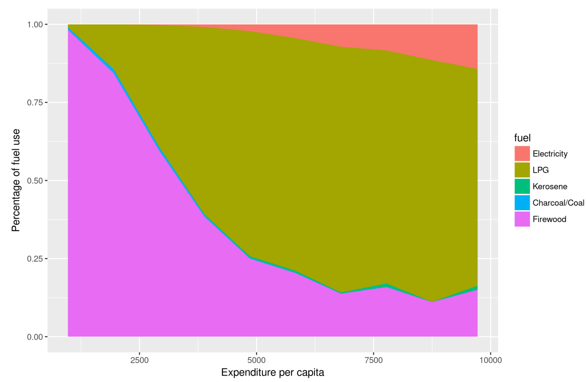
(a) Ghana: GLSS



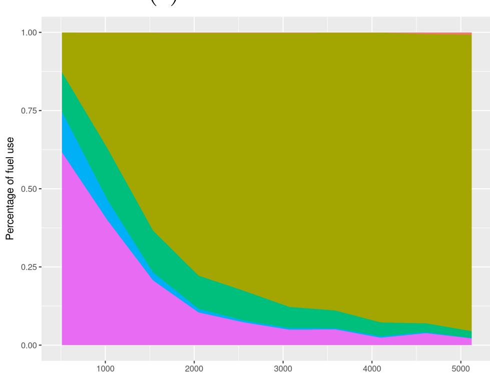
(b) Ghana: Model Simulation



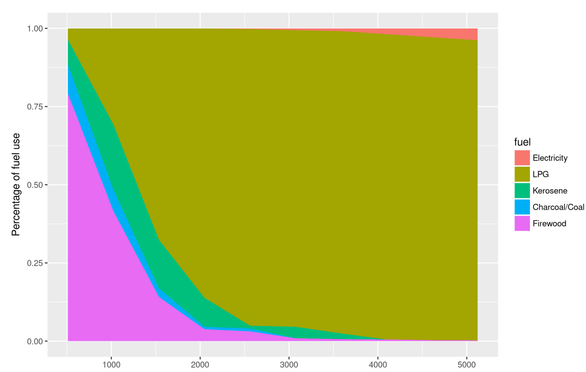
(c) Guatemala: ENCOVI



(d) Guatemala: Model Simulation

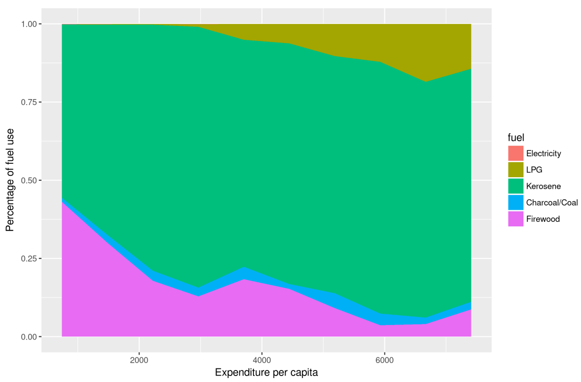


(e) India: NSS

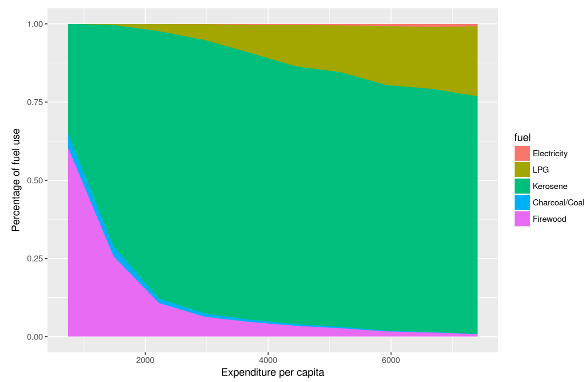


(f) India: Model Simulation

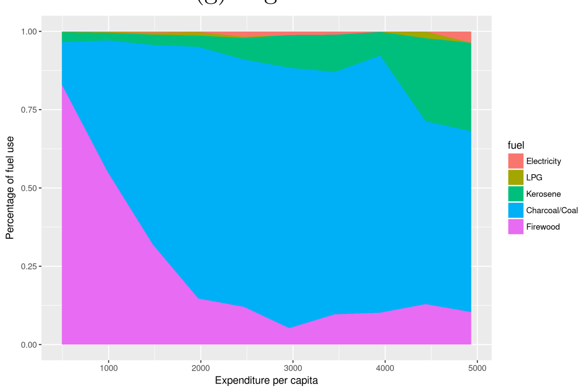
Figure 1: Percentage of fuel use in total fuel consumption by expenditure per capita per household for different countries, Data vs Model Simulation.



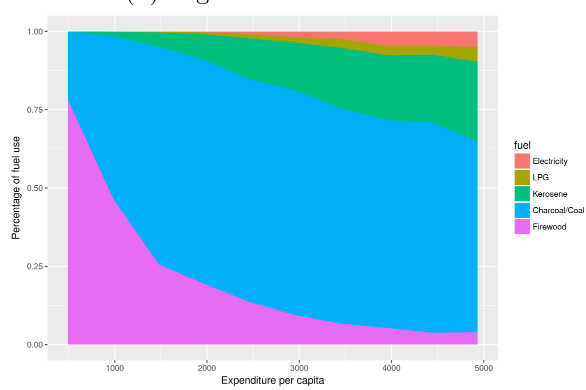
(g) Nigeria: GHS



(h) Nigeria: Model Simulation

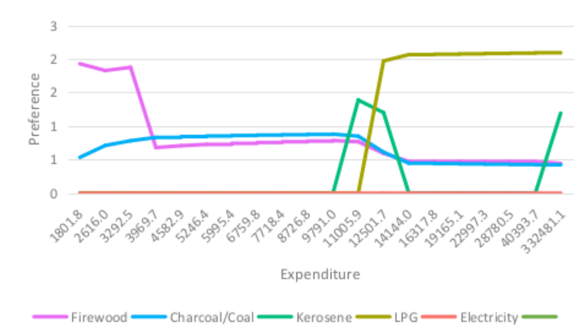


(i) Uganda: UNHS

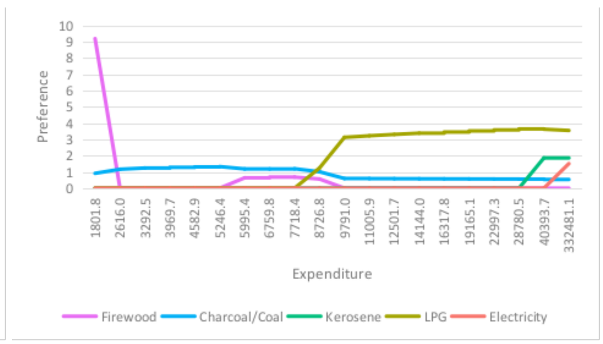


(j) Uganda: Model Simulation

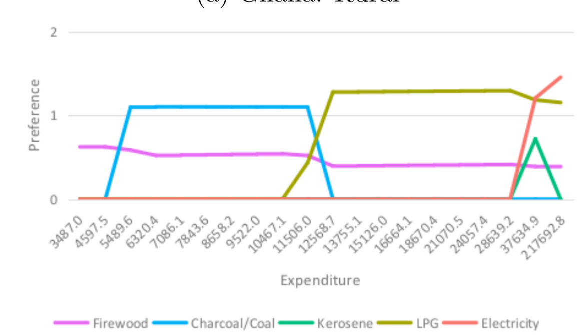
Figure 1: Percentage of fuel use in total fuel consumption by expenditure per capita per household for different countries, Data vs Model Simulation.



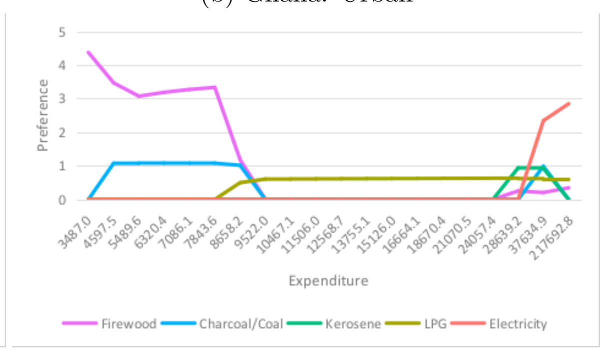
(a) Ghana: Rural



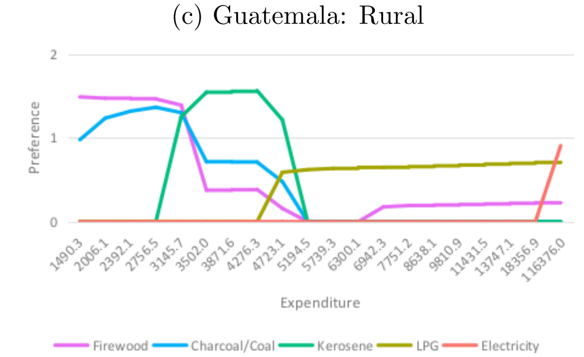
(b) Ghana: Urban



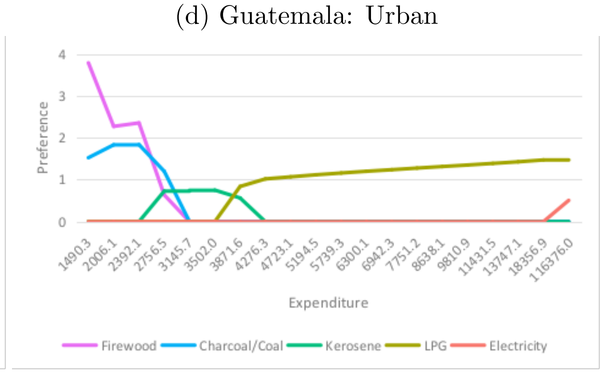
(c) Guatemala: Rural



(d) Guatemala: Urban

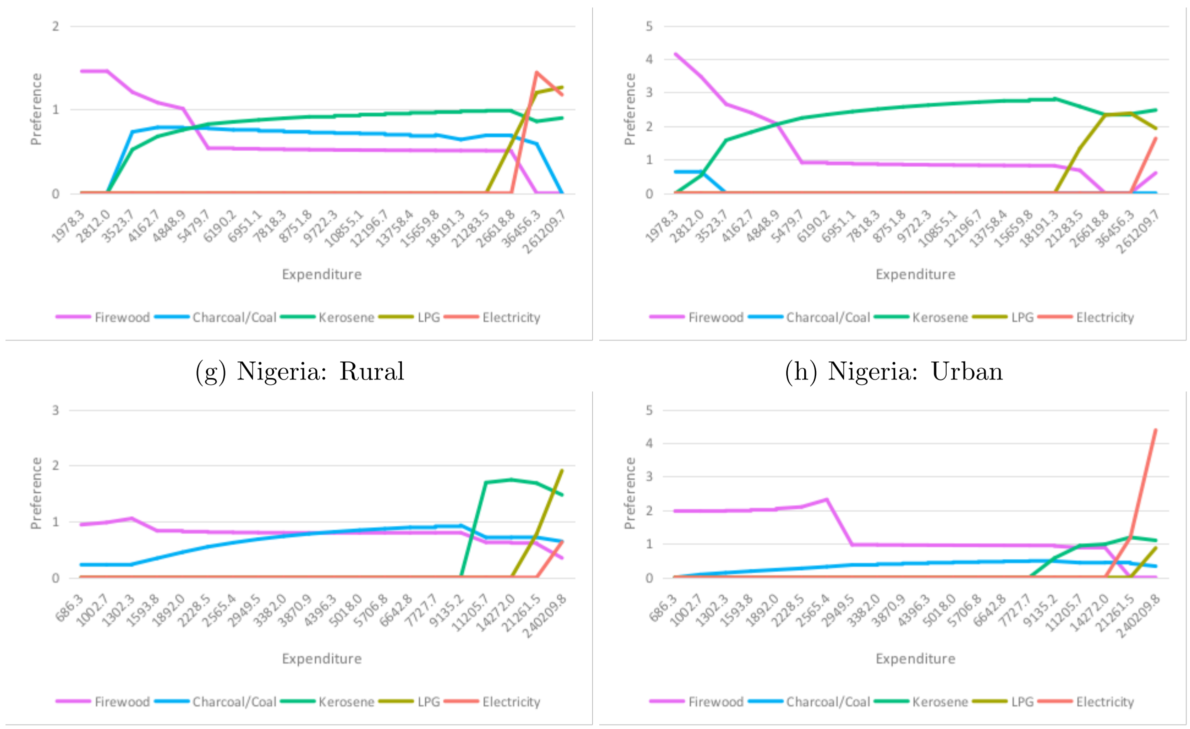


(e) India: Rural



(f) India: Urban

Figure 2: Preferences for Each Fuel by Expenditure Level.



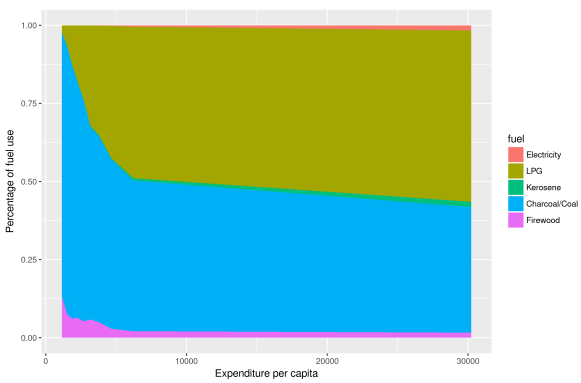
(g) Nigeria: Rural

(h) Nigeria: Urban

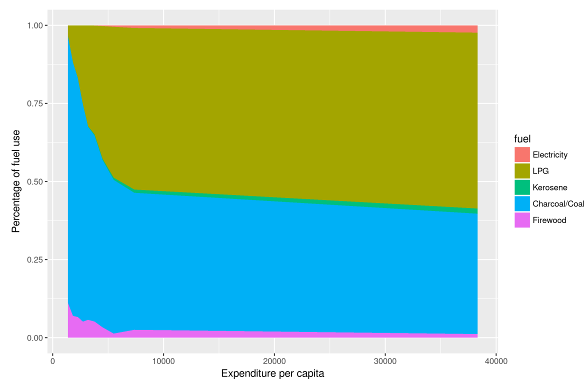
(i) Uganda: Rural

(j) Uganda: Urban

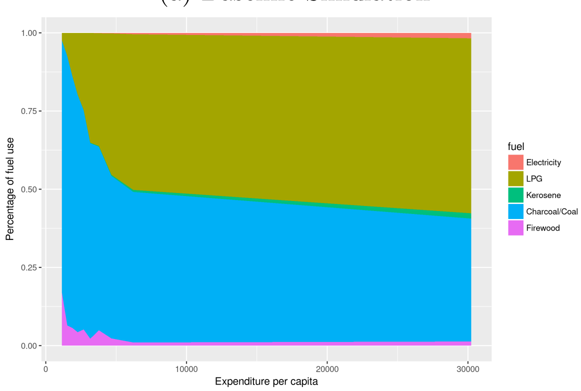
Figure 2: Preferences for Each Fuel by Expenditure Level.



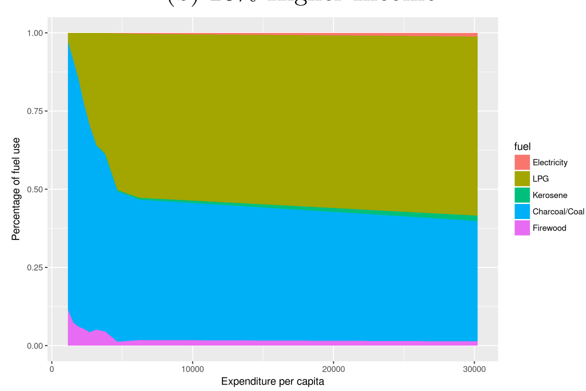
(a) Baseline Simulation



(b) 20% Higher Income

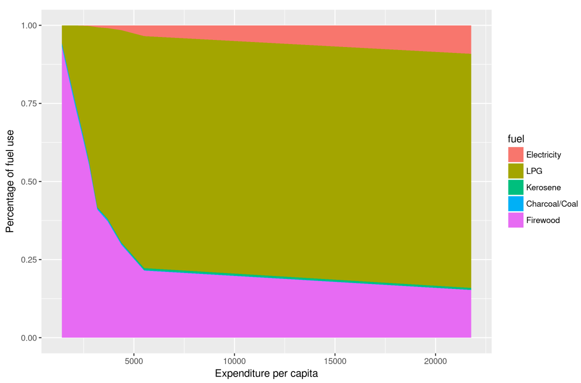


(c) 20% Higher Biomass Prices

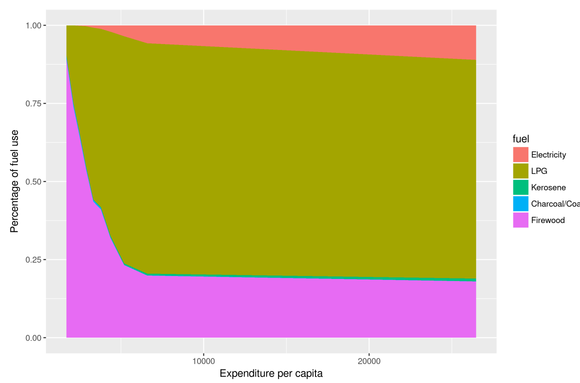


(d) 20% Lower LPG Prices

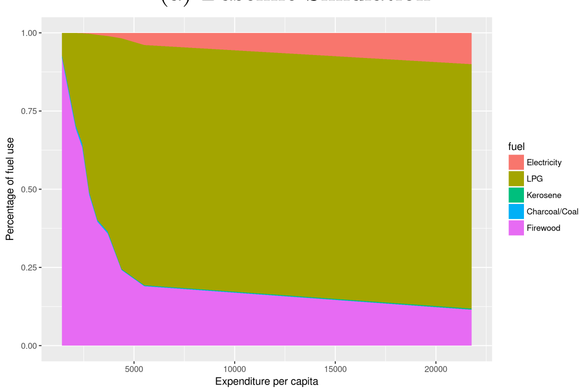
Figure 3: Ghana: Base Simulation and 20% Increase Scenarios



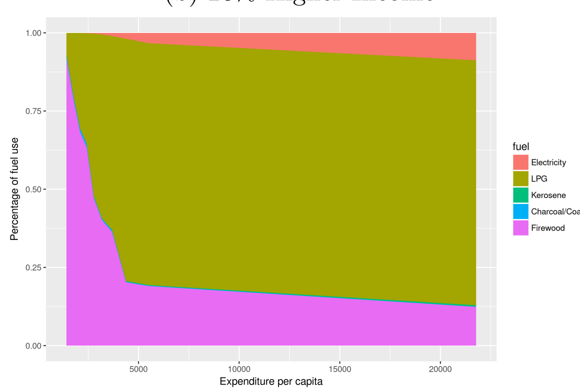
(a) Baseline Simulation



(b) 20% Higher Income

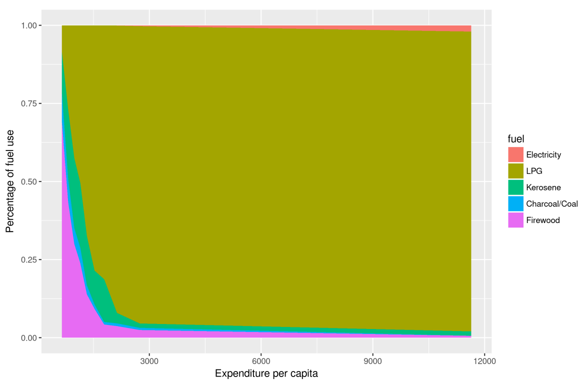


(c) 20% Higher Biomass Prices

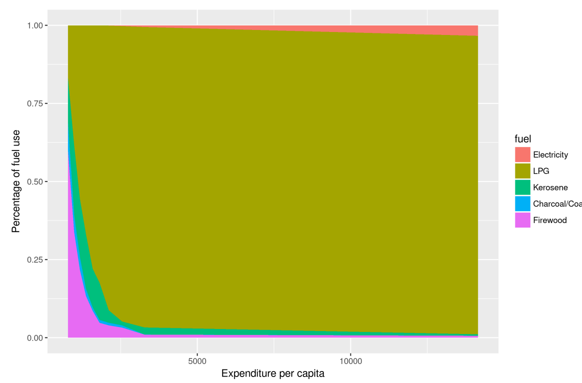


(d) 20% Lower LPG Prices

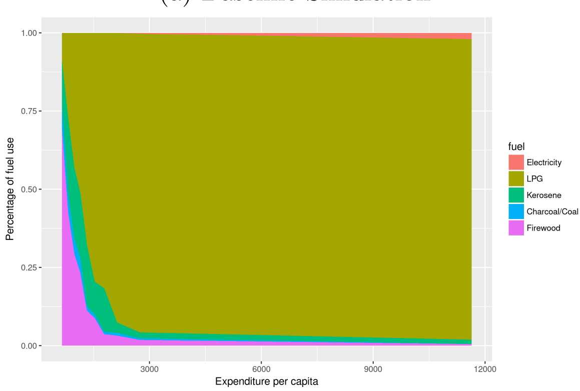
Figure 4: Guatemala: Base Simulation and 20% Increase Scenarios



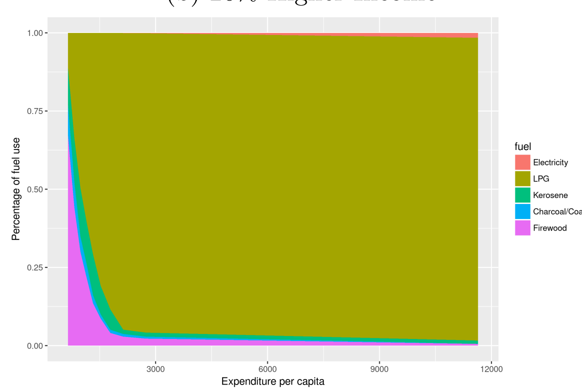
(a) Baseline Simulation



(b) 20% Higher Income

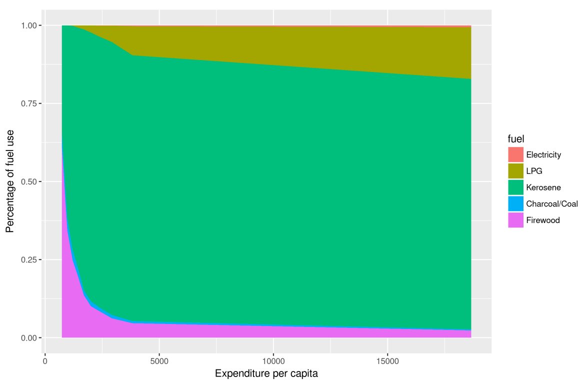


(c) 20% Higher Biomass Prices

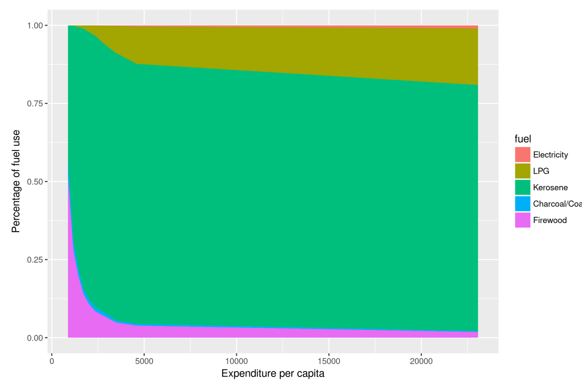


(d) 20% Lower LPG Prices

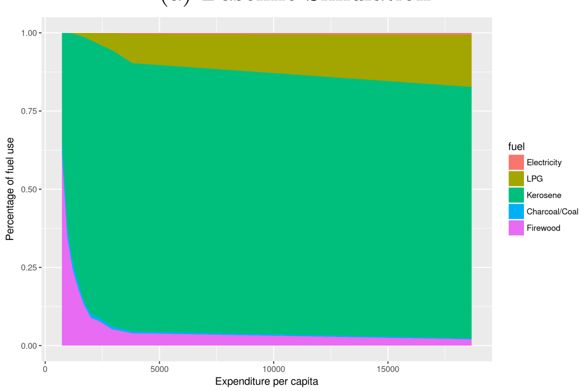
Figure 5: India: Base Simulation and 20% Increase Scenarios



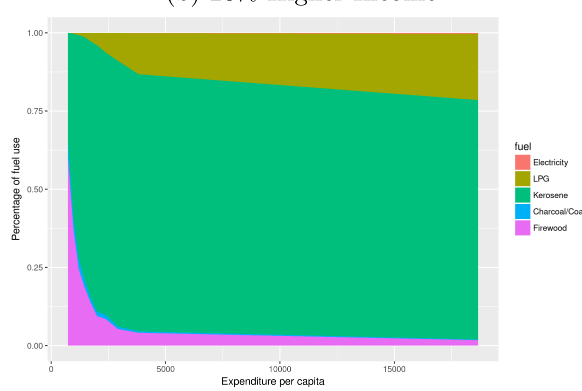
(a) Baseline Simulation



(b) 20% Higher Income

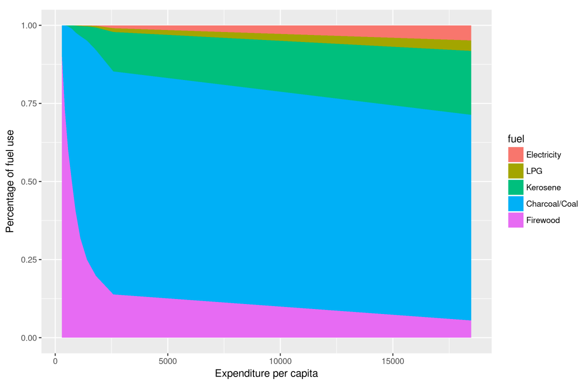


(c) 20% Higher Biomass Prices

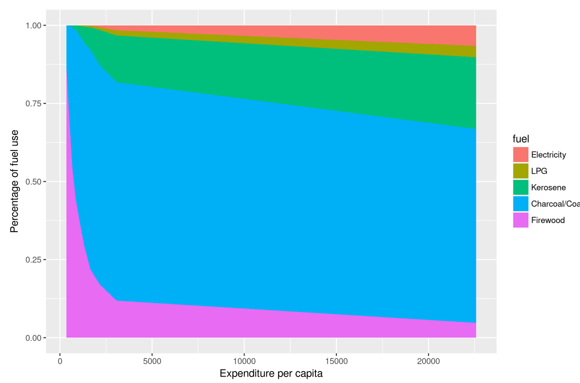


(d) 20% Lower LPG Prices

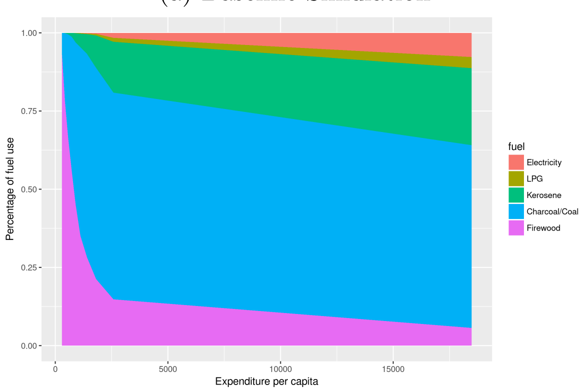
Figure 6: Nigeria: Base Simulation and 20% Increase Scenarios



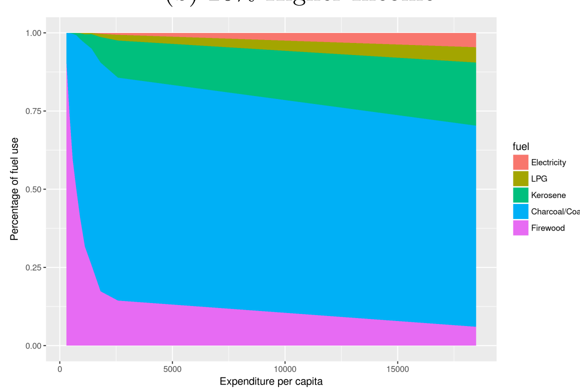
(a) Baseline Simulation



(b) 20% Higher Income



(c) 20% Higher Biomass Prices



(d) 20% Lower LPG Prices

Figure 7: Uganda: Base Simulation and 20% Increase Scenarios

References

Alem, Y., A. D. Beyene, G. Khlin, and A. Mekonnen (2016). Modeling household cooking fuel choice: A panel multinomial logit approach. *Energy Economics 59*, 129 – 137.

Cameron, C., S. Pachauri, N. Rao, D. McCollum, J. Rogelj, and K. Riahi (2016). Policy trade-offs between climate mitigation and clean cook-stove access in south asia. *Nature Energy 1*, 15010.

Campbell, B., S. Vermeulen, J. Mangono, and R. Mabugu (2003). The energy transition in action: urban domestic fuel choices in a changing zimbabwe. *Energy Policy 31*(6), 553 – 562.

Cheng, C.-y. and J. Urpelainen (2014). Fuel stacking in india: Changes in the cooking and lighting mix, 19872010. *Energy 76*, 306 – 317.

Ekholm, T., V. Krey, S. Pachauri, and K. Riahi (2010). Determinants of household energy consumption in india. *Energy Policy 38*, 5696–5707.

Fuso Nerini, F., C. Ray, and Y. Boulkaid (2017, June). The cost of cooking a meal. The case of Nyeri County, Kenya. *Environmental Research Letters 12*(6), 065007.

Health Effects Institute (2018). State of global air 2018. Technical report, Health Effects Institute.

Heckman, J. J. (2008). Econometric causality. *International Statistical Review 76*(1), 1–27.

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*(3), 337361.

Hosier, R. H. and J. Dowd (1987). Household fuel choice in zimbabwe: An empirical test of the energy ladder hypothesis. *Resources and Energy 9*(4), 347 – 361.

International Energy Agency (IEA) and the World Bank (2017). Global tracking framework 2017 - progress toward sustainable energy. Technical report, World Bank.

Keane, M. P. (2010). Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics 156*(1), 3 – 20. Structural Models of Optimization Behavior in Labor, Aging, and Health.

Koopmans, T. C. (1947). Measurement without theory. *The Review of Economics and Statistics 29*(3), 161–172.

Kowsari, R. and H. Zerriffi (2011). Three dimensional energy profile:: A conceptual framework for assessing household energy use. *Energy Policy 39*(12), 7505 – 7517. Clean Cooking Fuels and Technologies in Developing Economies.

Lewis, J. J. and S. K. Pattanayak (2012, February). Who adopts improved fuels and cookstoves? a systematic review. *Environmental Health Perspectives 120*(5), 637–645.

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

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica 57*(5), 995–1026.

OTA (1992). *Fueling development: energy technologies for developing countries.* Washington, DC: Congress of the U.S., Office of Technology Assessment : For sale by the U.S. G.P.O., Supt. of Docs.

Rahut, D. B., B. Behera, and A. Ali (2016). Patterns and determinants of household use of fuels for cooking: Empirical evidence from sub-saharan africa. *Energy 117*(Part 1), 93 – 104.

Smith, K. R. and A. Sagar (2014). Making the clean available: Escaping indias chulha trap. *Energy Policy 75*(Supplement C), 410 – 414.

van der Kroon, B., R. Brouwer, and P. J. van Beukering (2013). The energy ladder: Theoretical myth or empirical truth? results from a meta-analysis. *Renewable and Sustainable Energy Reviews 20*, 504 – 513.

van Ruijven, B. J., D. P. van Vuuren, B. J. de Vries, M. Isaac, J. P. van der Sluijs, P. L. Lucas, and P. Balachandra (2011). Model projections for household energy use in india. *Energy Policy 39*(12), 7747 – 7761. Clean Cooking Fuels and Technologies in Developing Economies.

Appendices

A Model solution

To simplify the calculations, we can rewrite the utility function by taking logarithm as:

$$
\begin{aligned}
U(C, F_f) &= \gamma \log \left(C^{\alpha} \left(\sum_{f=1}^{N_f} e_f F_f \right)^{1-\alpha} \right) + (1-\gamma) \log \chi \\
&= \gamma\alpha \log C + \gamma(1-\alpha) \log \left(\sum_{f=1}^{N_f} e_f F_f \right) + (1-\gamma) \log \chi
\end{aligned}
$$

Then the Lagrangian and FOCs:

$$
\begin{aligned}
\mathcal{L} : \gamma\alpha \log C &+ \gamma(1-\alpha) \log \left(\sum_{f=1}^{N_f} e_f F_f \right) + (1-\gamma) \log \chi \\
&+ \lambda \left[I - p_c C - \sum_{f=1}^{N_f} (p_f F_f + \mathbb{A}_f) \right] - \sum_{f=1}^{N_f} \mu_f F_f
\end{aligned}
$$

$$
\frac{\partial \mathcal{L}}{\partial C} : \frac{\gamma\alpha}{C} - \lambda p_c = 0 \tag{6}
$$

$$
\frac{\partial \mathcal{L}}{\partial F_i} : \frac{\gamma(1-\alpha) e_i}{\sum_{f=1}^{N_f} e_f F_f} + \frac{1-\gamma}{\chi} \cdot \frac{\partial \chi}{\partial F_i} - \lambda p_i - \mu_i = 0 \tag{7}
$$

$$
\frac{\partial \mathcal{L}}{\partial \lambda} : I - p_c C - \sum_{f=1}^{N_f} (p_f F_f + \mathbb{A}_f) = 0 \tag{8}
$$

From (1):

$$
\lambda = \frac{\gamma\alpha}{p_c C}
$$

In (2):

$$
\frac{\gamma(1-\alpha) e_i}{\sum_{f=1}^{N_f} e_f F_f} + \frac{1-\gamma}{\chi} \cdot \frac{\partial \chi}{\partial F_i} = \frac{\gamma\alpha}{C} \frac{p_i}{p_c} + \mu_i \tag{9}
$$

And from (3):

$$C = \frac{1}{p_c}\left(I - \sum_{f=1}^{N_f}(p_f F_f + \mathbb{A}_f)\right)$$

in (4):

$$\frac{\gamma(1-\alpha)e_i}{\sum_{f=1}^{N_f} e_f F_f} + \frac{1-\gamma}{\chi}\cdot\frac{\partial\chi}{\partial F_i} = \frac{\gamma\alpha p_i}{I - \sum_{f=1}^{N_f}(p_f F_f + \mathbb{A}_f)} + \mu_i \tag{10}$$

As long as χ is a strictly concave function of F_f, the system of equation (5)s for all fuels will give a unique solution for all F_i. In particular, as stated in section 3 we can assume that the function χ is a second degree polynomial on each of the fuels:

$$\chi = \chi_0 - \sum_{f=1}^{N_f}\left(\delta_{1f}F_i + \delta_{2f}F_i^2\right)$$

where δ_{1f} and δ_{2f} are constants to be estimated and χ_0 is a constant, in which case, equation (5) ends up as:

$$\begin{aligned}
\frac{\gamma(1-\alpha)e_i}{\sum_{f=1}^{N_f} e_f F_f} + \frac{1-\gamma}{\chi}(-\delta_{1f} - 2\delta_{2f}F_i) &= \frac{\gamma\alpha p_i}{I - \sum_{f=1}^{N_f}(p_f F_f + \mathbb{A}_f)} + \mu_i \\
\frac{1-\gamma}{\chi}(\delta_{1f} + 2\delta_{2f}F_i) &= \frac{\gamma(1-\alpha)e_i}{\sum_{f=1}^{N_f} e_f F_f} - \frac{\gamma\alpha p_i}{I - \sum_{f=1}^{N_f}(p_f F_f + \mathbb{A}_f)} + \mu_i \\
\Rightarrow F_i &= \begin{cases} \frac{1}{2\delta_{2f}}\left[\frac{\gamma(1-\alpha)e_i}{\sum_{f=1}^{N_f} e_f F_f} - \frac{\gamma\alpha p_i}{I-\sum_{f=1}^{N_f}(p_f F_f + \mathbb{A}_f)} - \frac{(1-\gamma)\delta_{1f}}{\chi}\right] & \text{if } \mu_i = 0 \\ 0 & \text{if } \mu_i > 0 \end{cases}
\end{aligned} \tag{11}$$

B Stove characteristics

Stove	Fuel	Price (2015 USD)	Efficiency (%)	Lifetime (yrs)
Traditional	Biomass	0	15	3
Traditional	Charcoal/Coal	0	20	3
Kerosene Stove	Kerosene	20	45	5
Gas Stove	LPG	78	60	10
Electric Induction	Electricity	95	80	15

Table 5: Stove Costs and Attributes

C Moments and estimated parameters

	Data	Simulation
Mean log total fuel consumption per capita	0.12040	0.28410
Mean share of fuel expenditure on total expenditure	0.01943	0.00598
Log mean firewood consumption per capita - Rural	-1.88547	-1.70434
Percentage of households using firewood - Rural	0.12903	0.09684
Log mean charcoal consumption per capita - Rural	0.24140	0.17288
Percentage of households using charcoal - Rural	0.76306	0.75928
Log mean kerosene consumption per capita - Rural	-5.74961	-5.90846
Percentage of households using kerosene - Rural	0.00321	0.00259
Log mean lpg consumption per capita - Rural	-2.27896	-2.25020
Percentage of households using lpg - Rural	0.10470	0.14129
Log mean electricity consumption per capita - Rural	-1e6	-1e6
Percentage of households using electricity - Rural	0.00000	0.00000
Percentage of firewood users who do not pay for it - Rural	0.26211	0.19223
Mean log total fuel consumption per capita - Rural	-0.03000	0.17978
Log mean firewood consumption per capita - Urban	-2.54439	-2.49997
Percentage of households using firewood - Urban	0.06181	0.03794
Log mean charcoal consumption per capita - Urban	0.29128	0.20841
Percentage of households using charcoal - Urban	0.66648	0.61889
Log mean kerosene consumption per capita - Urban	-5.50949	-5.64623
Percentage of households using kerosene - Urban	0.00363	0.00442
Log mean lpg consumption per capita - Urban	-1.22928	-0.98741
Percentage of households using lpg - Urban	0.26689	0.33526
Log mean electricity consumption per capita - Urban	-5.63113	-6.14242
Percentage of households using electricity - Urban	0.00119	0.00350
Percentage of firewood users who do not pay for it - Urban	0.09550	0.01929
Mean log total fuel consumption per capita - Urban	0.16806	0.32897

Table 6: Ghana: Matched Moments, Data vs Simulation

	Data	Simulation
Mean log total fuel consumption per capita	0.12040	0.28410
Mean share of fuel expenditure on total expenditure	0.01943	0.00598
Log mean firewood consumption per capita - Rural	-1.88547	-1.70434
Percentage of households using firewood - Rural	0.12903	0.09684
Log mean charcoal consumption per capita - Rural	0.24140	0.17288
Percentage of households using charcoal - Rural	0.76306	0.75928
Log mean kerosene consumption per capita - Rural	-0.03005	0.03602
Percentage of households using kerosene - Rural	0.00382	0.00374
Log mean lpg consumption per capita - Rural	0.73653	0.25961
Percentage of households using lpg - Rural	0.78923	0.83200
Log mean electricity consumption per capita - Rural	-5.20111	-5.32492
Percentage of households using electricity - Rural	0.00498	0.00303
Percentage of firewood users who do not pay for it - Rural	-6.06798	-8.73238
Mean log total fuel consumption per capita - Rural	0.00176	0.00023
Log mean firewood consumption per capita - Urban	-1.69364	-2.05257
Percentage of households using firewood - Urban	0.20134	0.15706
Log mean charcoal consumption per capita - Urban	-5.94661	-5.51044
Percentage of households using charcoal - Urban	0.00271	0.00769
Log mean kerosene consumption per capita - Urban	0.16441	0.12819
Percentage of households using kerosene - Urban	0.46691	0.18034
Log mean lpg consumption per capita - Urban	-0.25009	-0.31335
Percentage of households using lpg - Urban	0.33399	0.29163
Log mean electricity consumption per capita - Urban	-3.61101	-3.70983
Percentage of households using electricity - Urban	0.02294	0.01078
Percentage of firewood users who do not pay for it - Urban	-5.57049	-6.00389
Mean log total fuel consumption per capita - Urban	0.00464	0.00354

Table 7: Guatemala: Matched Moments, Data vs Simulation

	Data	Simulation
Mean log total fuel consumption per capita	-0.278313	-0.20175
Mean share of fuel expenditure on total expenditure	0.039202	0.008304
Log mean firewood consumption per capita - Rural	-1.033894	-1.06689
Percentage of households using firewood - Rural	0.375219	0.349079
Log mean charcoal consumption per capita - Rural	-3.32603	-3.390158
Percentage of households using charcoal - Rural	0.043549	0.036517
Log mean kerosene consumption per capita - Rural	-2.626113	-2.528752
Percentage of households using kerosene - Rural	0.14605	0.129237
Log mean lpg consumption per capita - Rural	-0.9597	-1.12946
Percentage of households using lpg - Rural	0.434521	0.484343
Log mean electricity consumption per capita - Rural	-6.700294	-8.105037
Percentage of households using electricity - Rural	0.000661	0.000824
Percentage of firewood users who do not pay for it - Rural	0.425781	0.417334
Mean log total fuel consumption per capita - Rural	-0.424986	-0.349781
Log mean firewood consumption per capita - Urban	-2.595581	-2.544708
Percentage of households using firewood - Urban	0.088251	0.071967
Log mean charcoal consumption per capita - Urban	-3.47636	-3.517885
Percentage of households using charcoal - Urban	0.02685	0.030492
Log mean kerosene consumption per capita - Urban	-2.568889	-2.684091
Percentage of households using kerosene - Urban	0.103166	0.121274
Log mean lpg consumption per capita - Urban	-0.225661	-0.168077
Percentage of households using lpg - Urban	0.779375	0.772567
Log mean electricity consumption per capita - Urban	-5.609683	-5.899155
Percentage of households using electricity - Urban	0.002358	0.0037
Percentage of firewood users who do not pay for it - Urban	0.186116	0.183636
Mean log total fuel consumption per capita - Urban	-0.12623	-0.074016

Table 8: India: Matched Moments, Data vs Simulation

	Data	Simulation
Mean log total fuel consumption per capita	-0.05867	-0.05179
Mean share of fuel expenditure on total expenditure	0.03186	0.01019
Log mean firewood consumption per capita - Rural	-0.71815	-0.85936
Percentage of households using firewood - Rural	0.33206	0.27744
Log mean charcoal consumption per capita - Rural	-2.55763	-3.67318
Percentage of households using charcoal - Rural	0.02955	0.02437
Log mean kerosene consumption per capita - Rural	-0.52158	-0.47574
Percentage of households using kerosene - Rural	0.62990	0.67492
Log mean lpg consumption per capita - Rural	-3.71881	-4.44532
Percentage of households using lpg - Rural	0.00797	0.02200
Log mean electricity consumption per capita - Rural	-6.34521	-7.24633
Percentage of households using electricity - Rural	0.00051	0.00127
Percentage of firewood users who do not pay for it - Rural	0.17155	0.31544
Mean log total fuel consumption per capita - Rural	-0.24931	-0.08619
Log mean firewood consumption per capita - Urban	-1.46544	-1.88598
Percentage of households using firewood - Urban	0.15271	0.10042
Log mean charcoal consumption per capita - Urban	-3.03532	-3.54353
Percentage of households using charcoal - Urban	0.02593	0.01740
Log mean kerosene consumption per capita - Urban	0.05885	-0.12293
Percentage of households using kerosene - Urban	0.79394	0.82621
Log mean lpg consumption per capita - Urban	-2.35893	-3.15303
Percentage of households using lpg - Urban	0.02678	0.05485
Log mean electricity consumption per capita - Urban	-6.46672	-7.27716
Percentage of households using electricity - Urban	0.00065	0.00113
Percentage of firewood users who do not pay for it - Urban	0.20985	0.22225
Mean log total fuel consumption per capita - Urban	0.05806	-0.02035

Table 9: Nigeria: Matched Moments, Data vs Simulation

	Data	Simulation
Mean log total fuel consumption per capita	0.05456	0.12100
Mean share of fuel expenditure on total expenditure	0.06156	0.00873
Log mean firewood consumption per capita - Rural	-0.12467	-0.15423
Percentage of households using firewood - Rural	0.70648	0.68628
Log mean charcoal consumption per capita - Rural	-0.92501	-1.26057
Percentage of households using charcoal - Rural	0.25466	0.27491
Log mean kerosene consumption per capita - Rural	-3.59895	-3.92478
Percentage of households using kerosene - Rural	0.03238	0.03018
Log mean lpg consumption per capita - Rural	-4.48360	-5.64586
Percentage of households using lpg - Rural	0.00568	0.00853
Log mean electricity consumption per capita - Rural	-6.74292	-11.00039
Percentage of households using electricity - Rural	0.00080	0.00009
Percentage of firewood users who do not pay for it - Rural	0.25465	0.18757
Mean log total fuel consumption per capita - Rural	-0.00464	-0.01711
Log mean firewood consumption per capita - Urban	-1.32632	-1.28024
Percentage of households using firewood - Urban	0.21118	0.17192
Log mean charcoal consumption per capita - Urban	0.09284	0.08312
Percentage of households using charcoal - Urban	0.71926	0.74338
Log mean kerosene consumption per capita - Urban	-2.74405	-2.56396
Percentage of households using kerosene - Urban	0.06058	0.07014
Log mean lpg consumption per capita - Urban	-5.84160	-5.93159
Percentage of households using lpg - Urban	0.00221	0.00319
Log mean electricity consumption per capita - Urban	-4.58681	-4.71885
Percentage of households using electricity - Urban	0.00677	0.01137
Percentage of firewood users who do not pay for it - Urban	0.24978	0.30678
Mean log total fuel consumption per capita - Urban	0.11987	0.23999

Table 10: Uganda: Matched Moments, Data vs Simulation

	Ghana	Guatemala	India	Nigeria	Uganda
α	0.977592	0.982371	0.975803	0.971912	0.979333
γ	0.984122	0.981700	0.971282	0.988477	0.991146
δ_{11r}	0.344353	0.274818	0.150705	0.404463	0.243272
δ_{21r}	1.117526	0.043807	0.435913	0.221758	0.181049
δ_{12r}	0.023655	1.054438	0.469848	0.549076	0.227166
δ_{22r}	0.221786	1.091348	0.441419	1.167535	0.254044
δ_{13r}	1.117898	0.694047	3.251980	0.268854	1.062099
δ_{23r}	1.906717	0.916931	1.202909	0.152332	0.454790
δ_{14r}	1.831923	0.282050	0.088901	0.526092	0.574437
δ_{24r}	0.136632	0.618817	0.175565	0.477849	0.965616
δ_{15r}	2.838725	0.452384	0.837482	1.145916	0.582249
δ_{25r}	0.442451	1.643963	0.466482	1.853091	3.999875
$\mathbb{K}_r$	0.389794	2.152453	0.717035	0.750419	3.979615
δ_{11u}	0.490966	0.360819	0.535303	0.407159	0.331692
δ_{21u}	0.321955	0.032231	1.481955	0.545691	0.866485
δ_{12u}	0.357782	0.573664	1.162668	1.433420	0.230334
δ_{22u}	0.110764	0.983048	1.073604	0.505507	0.077436
δ_{13u}	0.265354	0.384860	0.531671	0.165049	0.073047
δ_{23u}	1.848725	0.911925	0.150491	0.148398	0.457733
δ_{14u}	0.031888	0.144388	0.068005	0.522367	0.369651
δ_{24u}	0.340486	0.303863	0.268988	0.385637	0.767474
δ_{15u}	1.138297	0.111723	0.364815	1.170401	0.421342
δ_{25u}	1.396755	1.273751	0.487000	1.545850	0.856675
$\mathbb{K}_u$	2.484320	2.494098	1.493208	0.385466	1.332864

Table 11: Estimated Parameters for Different Countries

Notes:
Fuels: 1 - Firewood, 2 - Charcoal/Coal, 3 - Kerosene, 4 - LPG, 5 - Electricity
Groups: r - Rural, u - Urban