**Supplemental Information**

**Beyond Household Socioeconomic Status: Multilevel Modeling of Supply-Side Determinants of LPG Consumption among 5,500 Households in Sub-Saharan Africa**

Matthew Shupler1\*, Judith Mangeni2, Theresa Tawiah3, Edna Sang2, Miranda Baame4, Rachel Anderson de Cuevas1, Emily Nix,1 Emmanuel Betang4, Jason Saah3, Reginald Quansah5, Elisa Puzzolo1,6, Bertrand Mbatchou4, Kwaku Puko Asante3, Diana Menya2, Daniel Pope1, on behalf of the *CLEAN-Air(Africa)* investigators

1. Department of Public Health, Policy and Systems, University of Liverpool, Liverpool, United Kingdom

2. School of Public Health, Moi University, Eldoret, Kenya

3. Kintampo Research Centre, Kintampo, Ghana

4. Douala General Hospital, Douala, Cameroon

5. School of Public Health, University of Ghana, Ghana

6. Global LPG Partnership (GLPGP), 654 Madison Avenue, New York, United States

\*Corresponding author: m.shupler@liverpool.ac.uk

***Variables Available for Modeling***

Information on socioeconomic factors, demographics, cooking behaviors and LPG supply were collected in the rapid census survey conducted during phase one of the CLEAN-Air(Africa) study. Variables from the survey that were considered in multilevel models of LPG consumption and use as a primary or secondary fuel are presented in Table S1. To increase the number of variables able to be tested in modeling, continuous variables were categorized into equal groupings based on their median and inner quartile range. Due to the high number of factors included in all models, between-country effect differences (e.g. random slopes) were not included due to model saturation. Principal component analysis was used to generate a socioeconomic index (additional to household income) based on ownership of assets.1 The first principal component was categorized into quartiles and retained for inclusion as a predictor in the modeling. The highest quartile was most strongly associated with (ranked in order) (1) having an electricity connection and ownership of a (2) refrigerator, (3) television, (4) shower/bath in the house, (5) smartphone, (6) compact disc (CD) player, (7) computer and (8) car. Participants in the lowest quartile were more likely to own (1) off-grid solar electricity, (2) animals (e.g. cows, sheep, goats), (3) a mobile phone (not a smartphone), (4) farmland and (5) a scooter/motorbike.

**Table S1. Variables available for multilevel modeling**

|  |  |  |  |
| --- | --- | --- | --- |
| **Type** | **Variable** | **Variable Description** | **Categories** |
| Socioeconomic | Household income | Total household income | Country-specific quartiles |
| Education | Highest education level in household | (1) No formal education (2) Primary (3) Junior high (4) Secondary/high (5) College/university |
| Seasonal income | Does household income vary with seasons? | Yes/no |
| Household assets | Principal component analysis | First principal component grouped into quartiles |
| Demographic | Age  | Age of head of household | (1) 18-24 (2) 25-39 (3) 30-35 (4) 36+ |
| Sex of fuel decision maker | Sex of individual in household in charge of cooking fuel decision | (1) Male (2) Female (3) Joint decision |
| Household size | Number of individuals living in the home | (1) 1-2 (2) 3-4 (3) 5-6 (4) 7+ |
| Children under 5 | Number of children under 5 living in the home | Continuous variable |
| Marital status | Current living situation | (1) Married (2) Single (3) Cohabitating with partner (4) Widowed (5) Divorced |
| Cooking patterns | Frequency of LPG stove usage | Number of days in previous week using LPG stove | (1) 7 days (everyday) (2) Less than 7 days |
| Stove stacking | Number of cooking fuels used | (1) One cooking fuel (no stacking) (2) More than one cooking fuel used (stacking) |
| Duration of LPG use | Number of years spent cooking with LPG | (1) < 1 year (2) 1-2 years (3) 2-5 years (4) 6-10 years (5) >10 years |
| Number of burners | Number of burners on LPG stove | (1) 1 burner (2) 2 burners (3) 3 burners (4) 4+ burners |
| Supply | LPG fuel availability | How frequently is LPG unavailable when you are at the retail point? | (1) Never (always available) (2) Rarely (<4 times/yr) (3) Occasionally (4-12 times/yr) (4) Often (>12 times/yr)  |
| Cost of refill | Cost of an LPG cylinder refill | Reported in local currency and converted to USD: (1) 0-0.85 USD/kg (2) $0.86-1.00 USD/kg (3) $1.01-$1.10 USD/kg (4) $1.10 USD/kg or greater |
| Duration of transport | Travel time to reach the LPG cylinder retail point? | (1) Home delivery (2) 1-10 minutes (3) 11-20 minutes (4) 21-30 minutes (5) > 30 minutes |
| Cost of transport | Transportation costs for obtaining LPG cylinder | Reported in local currency and converted to USD: (1) $0-0.27 USD (2) $0.28-$0.36 USD (3) $0.37-$0.54 USD (4) $0.55 USD or greater |
| Mode of transport | Method of transportation typically used to obtain refill  | (1) On foot (2) Motorbike/scooter (3) Car (4) Public transport |
| Free biomass | Obtain biomass cooking fuels for free | Yes/no |

***Assessing Model Performance***

Final models were assessed based on the highest explained variability (R2) while minimizing the Akaike information criterion (AIC), with parsimonious models favored. Both the logistic (use of LPG as a primary versus secondary cooking fuel) and log-linear (kg LPG/capita/year) regression models performed modestly well (R2marginal = ~0.40-0.50). A larger percent of variability was explained by household SES in the model of whether LPG was used a primary or secondary fuel (pseudo R2marginal=0.03) (Table S2) than the model of annual LPG per capita consumption (R2marginal=0) (Table S3).

**Table S2. Goodness of fit statistics for LPG primary versus secondary fuel model (logistic regression) (N=2,247)**

|  |  |  |
| --- | --- | --- |
|  | **Main model performance** | **Cross Validation**  |
| **Model** | **Conditional R2 a** | **Marginal R2 b** | **Change in Marginal R2 c** | **ICCc** | **AICd** | **AUCe** |
| Base random effects model (clustering variables)1 | 0.20 | 0.12 | -- | 0.08 | 3503 | 0.70 |
| Base + socioeconomic variables2 | 0.23 | 0.15 | 0.03 | 0.10 | 3357 | 0.73 |
| Base + socioeconomic variables + demographic variables3 | 0.34 | 0.26 | 0.11 | 0.11 | 2892 | 0.78 |
| Base + socioeconomic variables + demographic variables + cooking pattern variables4 | 0.39 | 0.32 | 0.06 | 0.12 | 2696 | 0.80 |
| Base + socioeconomic variables + demographic variables + cooking pattern variables + supply-side variables5 | 0.52 | 0.42 | 0.10 | 0.17 | 2272 | 0.82 |

1. Base model includes a random intercept for community and a fixed effect for country

2. Socioeconomic variables included in final logistic regression model: *household income, highest household education level*

3. Demographic variables included in final logistic regression model: *age of household head, sex of cooking fuel decision maker, number of household members, number of children under 5*

4. Cooking pattern variables included in final logistic regression model: *days cooking with LPG in previous week*

5. Supply-side variables included in final logistic regression model: *transportation costs to LPG retail point, travel time to retail point, LPG availability, cost of cylinder refills*

a. Variance explained by random + fixed effects. Represents the Nagelkerke’s pseudo R2

b. Variance explained by fixed effects only. Represents the Nagelkerke’s pseudo R2

c. Intraclass correlation coefficient, representing the proportion of variability explained by between-community differences

d. Akaike information criterion

e. Area under the Receiver Operator Characteristic (ROC) curve

**Table S3. Goodness of fit statistics for annual per capita LPG consumption model (N=2,330)**

|  |  |  |
| --- | --- | --- |
| **Model** | **Main model performance** | **Cross Validation**  |
| **Conditional R2 a** | **Marginal R2 b** | **Change in Marginal R2 c** | **ICCc** | **RMSEd** | **AICe** | **Marginal R2 b** | **RMSEd** |
| Base random effects model (clustering variables)1 | 0.14 | 0.13 | -- | 0.01 | 0.68 | 5416 | 0.16 | 0.69 |
| Base + socioeconomic variables2 | 0.14 | 0.13 | 0.00 | 0.01 | 0.68 | 5303 | 0.15 | 0.69 |
| Base + socioeconomic variables + demographic variables3 | 0.44 | 0.44 | 0.31 | 0.01 | 0.55 | 4020 | 0.31 | 0.59 |
| Base + socioeconomic variables + demographic variables + cooking pattern variables4 | 0.48 | 0.48 | 0.04 | 0.01 | 0.53 | 3894 | 0.39 | 0.54 |
| Base + socioeconomic variables + demographic variables + cooking pattern variables + supply-side variables5 | 0.49 | 0.49 | 0.01 | 0.01 | 0.52 | 3867 | 0.39 | 0.54 |

1. Base model includes a random intercept for community and a fixed effect for country

2. Socioeconomic variables included in final log-linear regression model: *household income*

3. Demographic variables included in final log-linear regression model: *age of household head, sex of cooking fuel decision maker, number of household members, number of children under 5, marital status*

4. Cooking pattern variables included in final log-linear regression: *days cooking with LPG in previous week*, *fuel stacking, number of stove burners, years cooking with LPG*

5. Supply-side variables included in final log-linear regression model: *transportation costs to LPG retail point, travel time to LPG retail point, cost of LPG cylinder refills*

a. Variance explained by random + fixed effects.

b. Variance explained by fixed effects only.

c. Intraclass correlation coefficient, representing the proportion of variability explained by between-community differences

d. Root mean squared error

e. Akaike information criterion

***Logistic Regression Model Coefficients***

**Table S4. Coefficients from LPG primary versus secondary cooking fuel logistic regression model (with continuous covariates mean-centered and categorical predictors held at the population proportion) (N=2,247)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coefficient** | **Estimate** | **Std. Error** | **p < 0.05** | **Predicted probability** **[95% CI]** |
| Intercept | 1.38 | 0.56 | \* |  |
| **Country (REF = Ghana)** |  |  |  | 0.75[0.60, 0.86] |
|  Cameroon | -0.93 | 0.61 |  | 0.43[0.21, 0.68] |
|  Kenya | -1.09 | 0.71 |  | 0.33[0.13, 0.61] |
| **Household income quartile (REF=1st quartile (lowest))** |  |  |  | 0.56[0.39, 0.72] |
|  2nd quartile | 0.19 | 0.16 |  | 0.63[0.46, 0.78] |
|  3rd quartile | -0.04 | 0.20 |  | 0.65[0.48, 0.79] |
|  4th quartile (highest) | 0.18 | 0.24 |  | 0.73[0.57, 0.85] |
| **Highest level of education in household (REF=none)** |  |  |  | 0.44[0.23,0.68] |
|  Primary | 0.13 | 0.34 |  | 0.51[0.34, 0.68] |
|  Junior high | 0.53 | 0.42 |  | 0.67[0.44, 0.84] |
|  Secondary | 0.68 | 0.32 | \* | 0.63[0.46, 0.77] |
|  College/university | 1.10 | 0.33 | \* | 0.71[0.55, 0.83] |
| **Age of head of household (REF=18-24)** |  |  |  | 0.68[0.51, 0.81] |
|  25-29 | -0.11 | 0.16 |  | 0.65[0.48, 0.79] |
|  30-35 | -0.24 | 0.16 |  | 0.61[0.44, 0.76] |
|  36-75 | -0.51 | 0.15 | \* | 0.56[0.39, 0.72] |
| **Sex of cooking fuel decision maker (REF = female)** |  |  |  | 0.59[0.43, 0.74] |
|  Male | 0.58 | 0.15 | \* | 0.71[0.55, 0.84] |
| **Number of household members (REF=1-2)** |  |  |  | 0.89 [0.79, 0.95] |
|  3-4 | -0.8 | 0.21 | \* | 0.73[0.58, 0.84] |
|  5-6 | -1.55 | 0.21 | \* | 0.57[0.40, 0.72] |
|  7+ | -2.32 | 0.23 | \* | 0.39[0.24, 0.56] |
| **Number of children under 5 living in household (REF=1)** |  |  |  | 0.62[0.45, 0.76] |
|  2 | 0.7 | 0.15 |  | 0.64[0.47, 0.77] |
|  3-6 | -0.17 | 0.18 |  | 0.58[0.40, 0.73] |
| **Days cooking with LPG in previous week (REF=6 or less)** |  |  |  | 0.44[0.30, 0.59] |
|  7 days (everyday) | 1.22 | 0.12 | \* | 0.72[0.59, 0.83] |
| **Transportation cost for obtaining LPG refill (REF=<$0.25 USD)** |  |  |  | 0.71[0.57, 0.82] |
|  $0.26-0.50 USD | -0.41 | 0.18 | \* | 0.60[0.44, 0.74] |
|  $0.51-1.00 USD | -0.50 | 0.15 | \* | 0.57[0.42, 0.72] |
|  >$1.00 USD | -0.37 | 0.15 | \* | 0.61[0.45, 0.75] |
| **Travel time for obtaining LPG refill (REF = <10 minutes)** |  |  |  | 0.63[0.46, 0.77] |
|  11-20 minutes | 0.02 | 0.14 |  | 0.62[0.46, 0.77] |
|  21-30 minutes | -0.36 | 0.15 | \* | 0.56[0.39, 0.72] |
|  >30 minutes | 0.40 | 0.25 |  | 0.74[0.56, 0.86] |
| **LPG unavailable (REF = often (more than once a month))** |  |  |  | 0.49[0.30, 0.69] |
|  Sometimes (4-12 times a year) | -0.02 | 0.28 |  | 0.47[0.29, 0.65] |
|  Rarely (less than 4 times a year) | 0.43 | 0.25 |  | 0.59[0.42, 0.74] |
|  Never (always available) | 0.83 | 0.24 | \* | 0.70[0.55, 0.82] |
|  Do not know/Unsure | 0.47 | 0.33 |  | 0.57 [0.37, 0.75] |
| **Cost of refills (per kilogram) (REF<$0.86 USD/kg)** |  |  |  | 0.76[0.60, 0.87] |
|  $0.86-1.00 USD/kg | -0.32 | 0.25 |  | 0.65[0.51, 0.78] |
|  $1.01-1.10 USD/kg | -0.69 | 0.28 | \* | 0.58[0.42, 0.72] |
|  >$1.10 USD/kg | -1.18 | 0.33 | \* | 0.45[0.28, 0.62] |

***Log-Linear Regression Model Coefficients***

**Table S5. Coefficients (log scale) and exponentiated kilogram/capita/year usage (with covariates mean-centered) from LPG consumption model (N=2,330)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coefficient** | **Estimate** | **Std. Error** | **p < 0.05** | **Mean** **[95% CI] kg/capita/yr** |
| Intercept | 3.81 | 0.09 | \* |  |
| **Country (REF = Ghana)** |  |  |  | 15.66[12.69, 19.32] |
|  Cameroon | 0.05 | 0.05 |  | 14.54[11.71, 18.05] |
|  Kenya | -0.04 | 0.07 |  | 13.08[10.80, 15.83] |
| **Age of head of household (REF=18-24)** |  |  |  | 14.59[11.85, 17.96] |
|  25-29 | -0.04 | 0.03 |  | 14.05[11.47, 17.22] |
|  30-35 | -0.01 | 0.03 |  | 14.43[11.73, 17.74] |
|  36-75 | -0.05 | 0.03 |  | 13.77[11.21, 16.91] |
| **Household income quartile (REF=1st quartile (lowest))** |  |  |  | 14.01 [11.47, 17.10] |
|  2nd quartile | 0.05 | 0.04 |  | 14.52[11.86, 17.77] |
|  3rd quartile | 0.06 | 0.04 |  | 14.73[12.02, 18.06] |
|  4th quartile (highest) | 0.15 | 0.05 | \* | 16.59[13.54, 20.32] |
| **Sex of cooking fuel decision maker (REF = female)** |  |  |  | 14.01[11.47, 17.10] |
|  Male | 0.01 | 0.03 |  | 14.33[11.69, 17.56] |
| **Marital status (REF = married)** |  |  |  | 14.01[11.47, 17.10] |
|  Divorced/separated | -0.11 | 0.07 |  | 11.61[ 9.02, 14.96] |
|  Living together with partner/cohabiting | -0.08 | 0.04 |  | 13.56[11.00, 16.71] |
|  Single | 0.01 | 0.02 |  | 14.81[12.17, 18.02] |
|  Widowed | -0.03 | 0.05 |  | 14.34[11.39, 18.05] |
| **Number of household members (REF=1-2)** |  |  |  | 29.08[24.01, 35.23] |
|  3-4 | -0.60 | 0.05 | \* | 16.36[13.53, 19.79] |
|  5-6 | -0.92 | 0.06 | \* | 11.03[ 9.11, 13.36] |
|  7+ | -0.61 | 0.06 | \* | 8.10[ 6.66, 9.86] |
| **Number of children under 5 living in household (REF=1)** |  |  |  | 15.23[12.48, 18.59] |
|  2 | -0.17 | 0.03 | \* | 12.88[10.52, 15.77] |
|  3-6 | -0.18 | 0.04 | \* | 12.67[10.33, 15.54] |
| **Years cooking with LPG (REF = < 1 year)** |  |  |  | 14.01[11.47, 17.10] |
|  1-2 years | -0.08 | 0.05 |  | 12.82[10.62, 15.46] |
|  2-5 years | -0.07 | 0.05 |  | 13.31[11.11, 15.94] |
|  5- 10 years | -0.01 | 0.05 |  | 14.18[11.85, 16.97] |
|  >10 years | 0.00 | 0.05 |  | 14.90[12.40, 17.91] |
|  **Fuel stacking (REF = no)** |  |  |  | 14.01[11.47, 17.10] |
|  Yes | -0.13 | 0.05 | \* | 11.68[10.28, 13.27] |
| **Days cooking with LPG in previous week (REF=6 or less)** |  |  |  | 13.59[11.10, 16.63] |
|  7 days (everyday) | 0.08 | 0.02 | \* | 14.37[11.73, 17.59] |
| **Number of LPG stove burners (REF=1)** |  |  |  | 14.09[11.53, 17.21] |
|  2 | 0.40 | 0.03 | \* | 21.92[17.84, 26.93] |
|  3+  | 0.31 | 0.03 | \* | 20.60[16.80, 25.26] |
| **Transportation cost for obtaining LPG refill (REF=<$0.25 USD)** |  |  |  | 14.52[12.28, 17.17] |
|  $0.26-0.50 USD | 0.01 | 0.03 |  | 14.39[12.09, 17.14] |
|  $0.51-1.00 USD | -0.02 | 0.03 |  | 14.10[11.87, 16.75] |
|  >$1.00 USD | 0.00 | 0.03 |  | 14.45[12.14, 17.21] |
| **Travel time for obtaining LPG refill (REF = <10 minutes)** |  |  |  | 14.61[11.91, 17.91] |
|  11-20 minutes | -0.06 | 0.03 | \* | 13.93[11.37, 17.07] |
|  21-30 minutes | -0.09 | 0.03 | \* | 13.67[11.14, 16.76] |
|  >30 minutes | -0.10 | 0.05 | \* | 13.88[11.12, 17.33] |
| **Cost of refills (per kilogram) (REF=<$0.85 USD/kg)** |  |  |  | 18.98[15.31, 23.54] |
|  $0.86-1.00 USD/kg | -0.09 | 0.04 | \* | 15.87[12.82, 19.66] |
|  $1.01-1.10 USD/kg | -0.15 | 0.05 | \* | 15.12[12.20, 18.74] |
|  >$1.10 USD/kg | -0.35 | 0.06 | \* | 12.31[10.15, 14.94] |

***Households Cooking Exclusively with Polluting Fuels***

Nearly half (47%, n=1,248) of households cooking with polluting fuels previously used LPG for cooking (Table S6). The proportion of households reporting initial LPG equipment costs as a reason for not using LPG was 22% lower among households previously cooking with LPG (50%) compared with households not previously cooking with LPG (72%). However, the proportion of households previously cooking with LPG that listed the high cost of cylinder refills (34%) was 19% higher than the proportion of households not previously cooking with LPG (19%). This suggests that many households that pay the upfront cost of the LPG equipment likely discontinue their use of LPG due to the inability to pay for cylinder refills.

There was an 8% lower prevalence of safety concerns among households with previous LPG cooking experience cooking in Obuasi (35% of households previously cooking with LPG expressed safety concerns, compared to 27% that had not previously cooked with LPG) and 6% lower prevalence in Cameroon (15% to 9%) (Table S6).

**Table S6. Association between experience cooking with LPG and reason(s) for not currently using LPG for cooking among households exclusively cooking with biomass (N=2,685)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Reason(s) for not currently using LPG (multiple responses allowed)** | **Overall** | **Cameroon** | **Ghana** | **Kenya** |
| **Has cooked with LPG before****(N=1,478)** | **Has not cooked with LPG before****(N=1,587)** | **Has cooked with LPG before****(N=550)** | **Has not cooked with LPG before****(N=459)** | **Has cooked with LPG before****(N=612)** | **Has not cooked with LPG before****(N=354)** | **Has cooked with LPG before****(N=316)** | **Has not cooked with LPG before****(N=774)** |
| Cannot afford initial equipment cost | 742 (50%) | 1143 (72%) | 356 (65%) | 388 (85%) | 216 (35%) | 217 (61%) | 170 (54%) | 538 (70%) |
| Cannot afford gas refills | 549 (37%) | 302 (19%) | 212 (39%) | 131 (29%) | 320 (53%) | 164 (46%) | 17 (5%) | 7 (1%) |
| Safety concerns | 228 (15%) | 241 (15%) | 50 (9%) | 67 (15%) | 168 (27%) | 122 (34%) | 10 (3%) | 52 (7%) |
| Satisfied with current fuel | 109 (7%) | 151 (10%) | 49 (9%) | 34 (7%) | 29 (5%) | 13 (4%) | 31 (10%) | 104 (13%) |
| Too far to travel/fuel not available | 80 (5%) | 69 (4%) | 53 (10%) | 65 (14%) | 17 (3%) | 1 (0%) | 10 (3%) | 3 (0%) |
| Cannot cook all meals with LPG | 47 (3%) | 21 (1%) | 33 (6%) | 10 (2%) | 11 (2%) | 7 (2%) | 3 (1%) | 4 (1%) |
| Gave equipment away to child/relative  | 47 (3%) | 5 (0%) | 2 (0%) | 0 | 8 (1%) | 0 | 37 (12%) | 5 (1%) |
| Damaged equipment | 26 (2%) | 2 (0%) | 11 (2%) | 0 | 11 (2%) | 1 (0%) | 4 (1%) | 1 (0%) |
| Do not know how to use equipment | 3 (0%) | 13 (1%) | 0 | 1 (0%) | 0 | 0 | 3 (1%) | 9 (1%) |
| Do not know/will not answer | 23 (2%) | 46 (3%) | 10 (2%) | 7 (2%) | 8 (1%) | 5 (1%) | 5 (2%) | 34 (4%) |

***Assessing Bias In Self-Reporting of LPG Consumption***

A previous study comparing self-reported LPG fuel consumption (based on number of annual refills) with sales data on number of cylinders purchased among households in Karnataka state in India revealed that the majority (85%) of participants overreported their LPG use.2 The average overestimation was 3.53 cylinders/year (equivalent to 50 kg). In this study, using two different self-reported LPG consumption metrics (number of annual refills and average life of a cylinder), we find that the number of annual refills leads to overestimation of consumption by a median of 87 kg (range: 3-210 kg), compared to average cylinder lifetime (Figure S1).



**Figure S1. Distribution of discrepancies between self-reported values of annual LPG household consumption (N=2,830)**

In all cases, LPG consumption based on number of annual cylinder refills was higher than what was reported by average lifetime of a cylinder (Table S7). This may be due to the average length of a cylinder lasts is an easier quantity to recall because it determines how the cooking fuel decision maker must travel to the retail point. Users may find it difficult to convert average time between trips to the retail point to number of annual refills. In absolute terms, the discrepancy in this study is larger than the difference of 50 kg between self-report (using number of annual cylinder refills) and cylinder sales data reported in India. The median difference in self-reported annual LPG household consumption varied by country: 60 kg in Kenya, 87 kg in Ghana and 112 kg in Cameroon.

**Table S7. Difference between self-reported annual per capita consumption (kg) by average duration of a typical cylinder and average number of cylinder refills obtained per year**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Overall (N=2,830)** | **Cameroon (N=958)** | **Ghana (N=1,020)** | **Kenya (N=852)** |
|  | **Avg time btwn refills** | **# refills per year** | **Diff (%)** | **Avg time btwn refills** | **# refills per year** | **Diff (%)** | **Avg time btwn refills** | **# refills per year** | **Diff (%)** | **Avg time btwn refills** | **# refills per year** | **Diff (%)** |
| **LPG exclusive** | 16.5 | 19.4 | 2.9 (18%) | -- | -- | -- | 21.3 | 42.6 | 21.3(100%) | 5.7 | 7.1 | 1.4 |
| **LPG primary** | 7.4 | 13.5 | 6.1(82%) | 7.7 | 13.8 | 6.1(79%) | 9.7 | 17.4 | 7.7(79%) | 4.8 | 7.0 | 2.2 |
| **LPG secondary** | 5.3 | 9.2 | 3.9(74%) | 6.3 | 11.8 | 5.5(87%) | 12.0 | 17.7 | 5.7(48%) | 3.0 | 6.0 | 3.0 |

In the absence of LPG sales data in CLEAN-Air(Africa), it was not possible to evaluate the accuracy of the self-reported quantities. To be conservative in our model estimation, we used LPG consumption as reported by average lifetime of a cylinder, as this may have minimized the level of overreporting.

To check the robustness of the modeling results, the LPG consumption regression model with inclusion of only 50% of households at or below the median difference of 87 kg in self-reported consumption. The coefficients from the reduced model remained largely unchanged from the model with all households (Table S8). Model performance was slightly improved (R2marginal=0.53) from the model with all households included (R2marginal=0.47).

**Table S8. Comparison of coefficients (log scale) (with covariates mean-centered) from LPG consumption model with all households (Full Model) and households with below median difference in two self-reported estimates of LPG consumption (Reduced Model) (N=2,330)**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficient** | **Full Model Estimate** | **Full Model Std. Error** | **p < 0.05** | **Reduced Model Estimate** | **Reduced Model****Std. Error** | **p < 0.05** |
| Intercept | 3.81 | 0.09 | \* | 3.75 | 0.12 | \* |
| **Country (REF = Ghana)** |  |  |  |  |  |  |
|  Cameroon | 0.05 | 0.05 |  | 0.39 | 0.10 | \* |
|  Kenya | -0.04 | 0.07 |  | 0.16 | 0.11 |  |
| **Age of head of household (REF=18-24)** |  |  |  |  |  |  |
|  25-29 | -0.04 | 0.03 |  | -0.07 | 0.05 |  |
|  30-35 | -0.01 | 0.03 |  | -0.03 | 0.05 |  |
|  36-75 | -0.05 | 0.03 |  | -0.09 | 0.05 | \* |
| **Household income quartile (REF=1st quartile (lowest))** |  |  |  |  |  |  |
|  2nd quartile | 0.05 | 0.04 |  | 0.05 | 0.05 |  |
|  3rd quartile | 0.06 | 0.04 |  | 0.02 | 0.05 |  |
|  4th quartile (highest) | 0.15 | 0.05 | \* | 0.14 | 0.05 | \* |
| **Sex of cooking fuel decision maker (REF = female)** |  |  |  |  |  |  |
|  Male | 0.01 | 0.03 |  | 0.05 | 0.04 |  |
| **Marital status (REF = married)** |  |  |  |  |  |  |
|  Divorced/separated | -0.11 | 0.07 |  | -0.01 | 0.11 |  |
|  Living together with partner/cohabiting | -0.08 | 0.04 |  | -0.12 | 0.07 |  |
|  Single | 0.01 | 0.02 |  | 0.01 | 0.04 |  |
|  Widowed | -0.03 | 0.05 |  | -0.09 | 0.08 |  |
| **Number of household members (REF=1-2)** |  |  |  |  |  |  |
|  3-4 | -0.60 | 0.05 | \* | -0.63 | 0.05 | \* |
|  5-6 | -0.92 | 0.06 | \* | -1.04 | 0.06 | \* |
|  7+ | -0.61 | 0.06 | \* | -1.35 | 0.06 | \* |
| **Number of children under 5 living in household (REF=1)** |  |  |  |  |  |  |
|  2 | -0.17 | 0.03 | \* | -0.02 | 0.04 |  |
|  3-6 | -0.18 | 0.04 | \* | -0.10 | 0.05 | \* |
| **Years cooking with LPG (REF = < 1 year)** |  |  |  |  |  |  |
|  1-2 years | -0.08 | 0.05 |  | -0.10 | 0.06 |  |
|  2-5 years | -0.07 | 0.05 |  | -0.11 | 0.06 |  |
|  5- 10 years | -0.01 | 0.05 |  | -0.07 | 0.06 |  |
|  >10 years | 0.00 | 0.05 |  | -0.04 | 0.06 |  |
|  **Fuel stacking (REF = no)** |  |  |  |  |  |  |
|  Yes | -0.13 | 0.05 | \* | -0.13 | 0.07 |  |
| **Days cooking with LPG in previous week (REF=6 or less)** |  |  |  |  |  |  |
|  7 days (everyday) | 0.08 | 0.02 | \* | 0.07 | 0.03 | \* |
| **Number of LPG stove burners (REF=1)** |  |  |  |  |  |  |
|  2 | 0.40 | 0.03 | \* | 0.39 | 0.06 | \* |
|  3+  | 0.31 | 0.03 | \* | 0.29 | 0.05 | \* |
| **Transportation cost for obtaining LPG refill (REF=<$0.25 USD)** |  |  |  |  |  |  |
|  $0.26-0.50 USD | 0.01 | 0.03 |  | 0.04 | 0.05 |  |
|  $0.51-1.00 USD | -0.02 | 0.03 |  | -0.02 | 0.04 |  |
|  >$1.00 USD | 0.00 | 0.03 |  | -0.02 | 0.04 |  |
| **Travel time for obtaining LPG refill (REF = <10 minutes)** |  |  |  |  |  |  |
|  11-20 minutes | -0.06 | 0.03 | \* | -0.02 | 0.04 |  |
|  21-30 minutes | -0.09 | 0.03 | \* | -0.07 | 0.04 |  |
|  >30 minutes | -0.10 | 0.05 | \* | -0.08 | 0.07 |  |
|  Home delivery | -0.03 | 0.09 |  | 0.12 | 0.09 |  |
| **Cost of refills (per kilogram) (REF=<$0.85 USD/kg)** |  |  |  |  |  |  |
|  $0.86-1.00 USD/kg | -0.09 | 0.04 | \* | 0.03 | 0.06 |  |
|  $1.01-1.10 USD/kg | -0.15 | 0.05 | \* | -0.05 | 0.09 |  |
|  >$1.10 USD/kg | -0.35 | 0.06 | \* | -0.45 | 0.09 | \* |

**References**

1. Vyas S, Kumaranayake L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy Plan. 2006 Nov;21(6):459–68.

2. Kar A, Brauer M, Bailis R, Zerriffi H. The risk of survey bias in self-reports vs. actual consumption of clean cooking fuels. World Development Perspectives. 2020 Jun 1;18:100199.