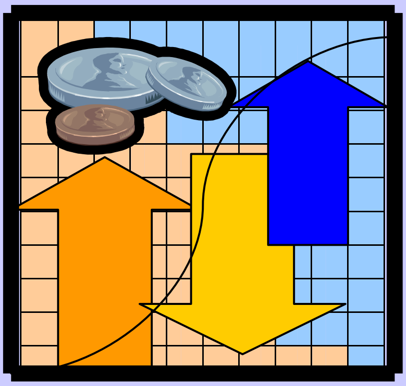
World Health Organization

HSS/HSF/DP.07.7



Coping with out-of-pocket Health Payments:

Applications of Engel Curves and Two-Part Models in Six African Countries

DISCUSSION PAPER

NUMBER 7 - 2007

Department "Health System Financing" (HSF)
Cluster "Health Systems and Services" (HSS)

World Health Organization 2007 ©

This document was prepared by Adam Lieve of The World Bank, Washington, DC and Xu Ke. We thank Andrew Jones for his advice on the econometric analysis and David Evans, Guy Carrin, and Eleonora Cavagnero for their valuable comments and suggestions. The authors also benefited from the discussion in the Health Systems Financing Department seminar held in WHO in September 2006.

Coping with the Out-of-Pocket Health payments:

Applications of Engel Curves and Two-Part Models in Six African Countries

by

Adam Leive and Ke Xu

World Health Organization
GENEVA
2007

SUMMARY

The costs of health services are often catastrophic in countries where people are not financially protected. Without sufficient safety nets, out-of-pocket payments for health care (OOP) make households vulnerable to future income shocks.

The objective of this paper is to examine how households modify their consumption of food, housing, education, and other goods in order to cope with OOP. The paper uses data from the 2003 World Health Survey in Burkina Faso, Chad, Kenya, Senegal, Zambia, and Zimbabwe. A system of Working-Leser Engel curves is estimated in the form of budget shares corresponding to proportions of total non-OOP expenditure. Two-part models (2PMs) are used to estimate absolute expenditure changes.

Households begin to cope at different levels of OOP spending across countries, but there are strong signs of coping when OOP spending is greater than 40 percent of non-subsistence expenditure. The expenditure share allocated to food increases with higher levels of OOP and falls for other goods. However, households at the highest OOP level are predicted to reduce expenditures on food, housing, and education. There is evidence of a negative gradient with rising OOP levels across goods and countries in the 2PMs.

Since prepayment for health care is minimal in these countries, prepayment mechanisms or other forms of social protection could protect household consumption. The analysis also points to the link between financial protection from OOP with a multi-sectoral approach to poverty alleviation.

INTRODUCTION

The costs of illness can be substantial in countries where people are not financially protected. Such costs include both direct expenses, such as out-of-pocket payments (OOP) for medical treatment, and indirect costs, such as the loss of income from an inability to work or travel to a hospital. These expenses often constitute a large share of a household's disposable income and drive many into poverty. People pursue a variety of coping strategies depending on the type and extent of costs incurred. For example, families may alter their labour allocation decisions; if a household head falls ill, some family members previously unemployed may begin working to substitute for lost income. If they are unable to finance the cost of medical care from current income, people may use savings or sell assets. Some may also borrow money from friends and family, take out a loan using collateral, or beg.

Another reaction to illness is to modify consumption. This may be very detrimental to households if the goods that are reduced are concomitant or even necessary to escaping poverty. As (Gertler, Levine & Moretti, 2002) point out, the degree to which consumption is reduced to pay for OOP depends on the severity of the illness. Severe illness significantly reduces household consumption. In a study of health shocks in rural Ethiopia found if the household head becomes sick, weekly food expenditure is expected to decrease by 24 percent on average (Asfaw & Von Braun, 2004). Non-food expenditure is predicted to decrease by 28 percent(Carrin, Gray & Almeida, 1999). found similar results. In a review, Russell argued that education was the main area of investment under threat in the case of illness in several studies (Russell, 1996). In a study of the impact of HIV/AIDS mortality on households in Thailand, Pitayanon estimated that consumption on food and beverages decreased by 42 percent on average as a result of illness (Pitayanon, Kongsin & Janjareon, 1997). In some cases, children of the deceased were taken out of school to work. For those households with a non-HIV related death, the same percentage cut food expenditure, but by a smaller amount**.** Pryer concluded that in the Dhaka slums in Bangladesh, not many households changed labour allocation or reduced housing expenditure, but more households reduced expenditure on other goods(Pryer, 2003).

Other literature suggests consumption modification is a transitory or seldom used coping strategy. Wilkes found that in rural China, several households coping with severe illness modified consumption patterns by reducing expenditure on food and education, but this was not employed in the subsequent years of the study(Wilkes, Hao, Bloom & Xingyuan, 1997). In rural Bangladesh, Desmet found that health expenditure caused households to moderately reduce food items that were not necessities while protecting staple food items(Desmet, 2000). In addition, relocating to another residence due to health expenditure constituted only two percent of households who moved. Education expenditure was not reduced while clothing expenditure decreased in the lowest income quintile and for those with the highest health expenditure level. Desmet concluded that foregoing consumption of essential commodities could hardly be established. In a study of low-income countries, Townsend found that the percentage of the year an male adult is ill does not influence consumption, but as Gertler and Gruber mention, this may be due to measuring only small changes in health status(Townsend, 1995; Gertler & Gruber, 2002).

The literature demonstrates the existence of an integrated network of coping strategies used by households. Different strategies are employed in different settings and depend on the severity of illness. While the degree to which expenditure on essential commodities is reduced is not clear, the empirical evidence suggests that households modify consumption in the face of severe illness, at least in the short term.

This paper will examine short-term changes in consumption patterns due to OOP. It adds value to the current knowledge of coping strategies by empirically examining how households in African countries adjust their expenditure patterns on food, education, housing, and other goods with rising OOP levels. The use of a standardized survey questionnaire in each country improves the ability to systematically determine whether a similar pattern exists in the consumption modification strategies across these countries. Countries included in this study are Burkina Faso, Chad, Kenya, Senegal, Zambia, and Zimbabwe. The next section describes the available data and methodology. The results are then presented along with discussion and conclusions.

DATA AND VARIABLES

The data used in this analysis come from the 2003 World Health Survey (WHS) in Burkina Faso, Chad, Kenya, Senegal, Zambia, and Zimbabwe. The WHS is a clustered random sample designed to be representative of the entire population. Probability weights are used to correct for unit non-response and reconstruct population estimates. The sample sizes for the six countries are listed in Table 1. The questionnaire includes a number of expenditure and health-related modules as described below.

Expenditure items

Information is available on household expenditure within the last month for food, housing, education, OOP, health insurance premiums, and other goods and services. Values are reported in local currency, whether paid in cash or in kind. Food expenditure (FOOD) includes the value of any food produced and consumed by the household and excludes alcohol, tobacco, and restaurant meals. Education expenditure (EDUCATION) contains school fees and money spent on supplies. Housing expenditure (HOUSE) includes rent or payment for housing and costs for gas, electricity, water, telephone, and heating fuel. Health insurance premiums will be combined with the other goods and services category (OTHER). Total expenditure (TEXP) is defined as the sum of FOOD, EDUCATION, HOUSE, and OTHER. Total expenditure is defined without OOP in order to isolate the effects of OOP on expenditure share adjustments for different commodities.

The key variable of interest will be OOP as a proportion of a household's capacity to pay (CTP). Capacity to pay is defined as the household's non-subsistence expenditures and uses the methodology outlined in (Xu, Aguilar, Carrin & Evans, 2005), which calculates subsistence spending equal to the product of a relative, food based poverty line and the equivalence-scaled household size. The equivalent scaled household size is used to account for economies of scale in household consumption and is calculated by raising the household size to the power 0.56, which has been estimated from Xu et al 2003. The rationale for constructing this ratio is to reflect the notion that OOP does not necessarily have to be high in absolute terms to be considered catastrophic for a household. This ratio appears as categorical variable with the base category of OOP less than 10 percent of non-subsistence spending (CATA1) followed by OOP between 10 and 20 percent (CATA2), 20 and 40 percent (CATA3), and greater than 40 percent (CATA4).

Sources used to finance OOP

There is a series of questions regarding the financial sources used by the household to pay for OOP. These include current income (INC), savings (SAVE), reimbursement from an insurance plan (INS), the sale of assets (SELL), money borrowed (BORROW), and other (COPE-OTHER). These appear as dummy variables, with INC as the omitted variable, in order to control for the effects of these strategies on household consumption. The COPE-OTHER dummy variable could theoretically correspond to the case where some members of the household previously not working begin to work in order to finance OOP. The literature suggests intra-household labour substitution is a commonly employed strategy to cope with both the direct and indirect costs of illness.

Socioeconomic and demographic controls

A household roster contains information on the age, education level, marital status, and insurance coverage of each person living in the household. It also identifies the person providing the main economic support of the household, known as the household head. Regressors for demographic variables include a dummy variable for urban location of the household (URBAN), equivalence-scaled household size (EQSIZE), and the proportion of household members between the ages of 5 and 18 (SAGE518). The education level, age, and insurance status (INS-HEAD) of the household head are used to measure socioeconomic status. Education level is measured using three dummy variables. EDU1 corresponds to less than primary school completed and is the base category. EDU2 corresponds to secondary school or high school completed and EDU3 university or higher. The age of the household head appears as a second-order polynomial (AGE and AGE2 = $AGE^2/100$) to account for non-linearity in the effect of age.

Labels, definitions, and averages for variables used in the analysis are found in Table 2.

The problem of zero expenditure

The expenditure items in each country have a non-ignorable proportion of zero values. For insurance premiums, OOP, and other goods and services, zero values may seem possible, especially for poorer households. However, it is unlikely that households spend nothing on food since the value of anything produced by the household is supposed to be included. The only case a zero would be a true value would be if the household received a government subsidy. For housing, it is possible that the household did not purchase a home or pay rent during the survey period, but less likely they spent nothing on the associated costs. Nevertheless, since many of the households surveyed may come from very poor areas, this may be possible. For education, households without any children are not likely to have any expenditure. For those with children, it is possible that the survey period did not cover the time when school fees were collected. This may be the case for many households if school tuition is charged just once a year, for example. Table 3 below presents the proportion of observations with zero values for different expenditure items.

Methods of dealing with zero values differ depending on the nature of the zeros. The assumption is made that zeros for EDUCATION, HOUSE, and OTHER are due to purchase infrequency.

METHODOLOGY

Engle curves

In order to determine how households adjust their consumption patterns to cope with higher levels of OOP, a system of Engel curves is estimated in the form of budget shares. An Engel curve explains the relationship between consumer expenditure on a particular good and his total income or expenditure, holding prices constant. This analysis will estimate Engel curves in the general Working-Leser form:

$$w_{ih} = \alpha_{ih} + \beta_{ih} \log x_h \quad (1)$$

where w_i is the expenditure share of good i for household h, x is total expenditure, and α_i and β_i are parameters to be estimated. In each of the six countries, the system contains four equations corresponding to the expenditure shares (excluding OOP) of FOOD, EDUCATION, HOUSE, and OTHER. For any country, a typical equation can be specified as follows:

$$s_{ih} = \alpha_{ih} + \beta_1 \log(\text{TEXP}_h) + \beta_2\text{CATA2}_h + \beta_3\text{CATA3}_h + \beta_4\text{CATA4}_h + C_h\gamma + X_h\lambda + u_{ih} \quad (2)$$

where s is the expenditure share of good i for household h, C is the set of coping strategy dummy variables (SAVE, BORROW, SELL, COPE-OTHER), X is the set of socioeconomic and demographic controls, and u is a random error term. The prior hypothesis is that households protect food expenditure when they spend more of their capacity to pay on OOP and sacrifice both education and housing but particularly other goods. If this occurs, the coefficients should be positive on CATA2, CATA3, and CATA4 in the food share equations and negative in the others.

The equations are estimated simultaneously in order to account for possible correlation between the error terms. The system as defined in (2) is estimated both by seemingly unrelated regression (SURE) and three-stage least squares (3SLS). Instrumental variable (IV) estimation is used because TEXP may be endogenous due to measurement error or because it is jointly determined with the expenditure shares(Lewbel, 1996; Blundell, 1988). Education level of the household head is used as instruments for TEXP for each country except in Senegal and Zambia, where insurance status of the household head is also added to the instrument set in order to meet the orthogonality conditions. The instruments pass tests for overidentification using Hansen's *J*-statistic (Table 4) and relevance (the *F*-statistic of the regression of the endogenous variable, TEXP, on the excluded instruments is above 10 in each case(Table 5)[1](Hansen, 1982; Baum, Schaffer & Stillman, 2003).

[1] However, it is worth noting that for some of the equations, the p-values are only slightly larger than 0.05, which suggests interpreting these equations with caution due to the uncertainty surrounding the orthogonality of the instruments. This occurs for 2 of the countries in the education equation, which may not be surprising as education level of the household head is used as instruments.

3SLS estimates are compared to the SURE estimates with a Hausman specification test to determine if TEXP is endogenous (Table 6). SURE is rejected at the 5% significant level as a consistent estimator in all countries, except in Chad and Zimbabwe. Since estimation by 3SLS is less efficient than SURE if expenditure is actually exogenous, SURE is used to estimate the system for Chad. In Zimbabwe, 3SLS is still used as a precaution although the *p*-value of 0.090 passes the conventional significance level. Finally, parameter estimates for the countries estimated using 3SLS are compared to estimates from 2SLS to assess any misspecification in the systems. In each share equation for every country, the *p*-values from the Hausman test are zero and so 3SLS are used for efficiency reasons.

Two-part models The second part of the analysis estimates two-part models (2PM) for food, education, and housing expenditure in each country to determine if households decrease absolute spending on these goods. In a 2PM, separate processes are used to model the censoring mechanism and the outcome of interest. Such a specification is useful when there is strong reason to believe that certain observed values occur with either too large or too small a frequency for a simpler model to estimate (Cameron & Trivedi, 2005). Following conditional probability theory, the general form of a 2PM can be written as follows:

$$E[\, y \mid x] = \Pr(y > 0) \times E[\, y \mid y > 0, x] \quad (3)$$

where y is the dependent variable of interest and x is a set of covariates.

The first part of the model, also known as the hurdle specification, is estimated using a logit regression to determine the probability of observing a positive expenditure. The dependent variable b_{ih} is a binary outcome corresponding to whether positive expenditure on a good is observed or not. This can be given a latent variable interpretation, where

$$b_{ih} = 1 \quad \text{iff } b_{ih}^{*} > 0 \quad (4)$$
$$= 0 \quad \text{otherwise}$$

$$\text{and} \quad b_{ih}^{*} = x_{ih}\beta_{ih} + \varepsilon_{ih} \quad (5)$$

where i corresponds to the commodity (either FOOD, EDUCATION, or HOUSE) for household h, x is a vector of all regressors used in the Engel curve analysis except now the variable TEXP subtracts expenditure on commodity i, in addition to subtracting OOP, from total expenditure, β is a vector of parameters, and ε is the error term with a standard logistic distribution.

The second part of the model, also known as the level specification, uses either OLS or Generalized Linear Modelling (GLM) on the set of positive observations. For estimation by OLS, the logarithm of FOOD, EDUCATION, or HOUSE is the dependent variable for those households reporting positive expenditure for that commodity. A typical equation can be written as,

$$\log(y_{ih}) = x_{ih}\beta_{ih} + u_{ih} \qquad \text{if } i > 0 \tag{6}$$

where y_{ih} is expenditure for household h on commodity i, x is the same set of regressors as in the logit specification, β is a vector of coefficients, and u is an additive error term but where normality and constant variance y_{ih} are not assumed. The second part of the model gives information about E [log(y_{ih}) | y_{ih} > 0, x], not E[y_{ih} | y_{ih} > 0, x]. In the absence of normality, retransformation from the log scale back to the raw scale involves using Duan's nonparametric smearing estimator (Duan, 1983). This is calculated by taking the mean of the exponentiated residuals from the logged equation over the positive observations.

$$S = \frac{1}{n+}\sum_{i=1}^{n_+}\exp(\hat{u}_i) \tag{7}$$

where $\hat{u}$ are the estimated residuals and n_+ is the set of positive observations. So the expected value of the level equation can then be written as,

$$E[\, y \mid y > 0,\, x] = \exp(x\beta + ½\, \sigma^2) \times S \tag{8}$$

This estimator is only consistent under homoskedasticity. For the equations with heteroskedastic log-scale errors, GLM estimation is used to prevent yielding biased results. Chad passes the White test for heteroskedasticty for every equation, while Senegal passes for FOOD and EDUCATION, and Burkina Faso for HOUSE (Table 7). In order to achieve both consistency and efficiency, modified Park-tests as suggested by Manning and Mullahy are used to determine the best choice of GLM estimator (Manning & Mullahy, 2001).(Tables 8-10)[2] In each case, the gamma distribution is chosen with a log-link between the variance and the mean.

In the 2PM, the partial effects from the OOP dummy variables (CATA2, CATA3, and CATA4) are once again of primary interest. The formula for the partial effect for a discrete variable, x_d, is

$$\frac{\partial E[y \mid x]}{\partial x_d} = E[\, y \mid x_d = 1] - E[\, y \mid x_d = 0] \tag{9}$$

$$= (\Pr(\, y > 0 \mid x_d = 1) - \Pr(\, y > 0 \mid x_d = 0)) \times E[\, y \mid y > 0 \mid x_d = 0] + \Pr(\, y > 0 \mid x_d = 0) \times (E[\, y \mid y > 0 \mid x_d = 1] - E[\, y \mid y > 0 \mid x_d = 0]) \tag{10}$$

Partial effects are computed for every household in the sample and are evaluated at the observed values of the regressors. The average partial effect (APE) is simply the mean of

2 This test is performed by running OLS on the logarithm of the squared residuals from the raw scale on a constant and the logarithm of the raw-scale predicted values:
$\ln(y - \hat{y})^2 = \theta_0 + \theta_1 \ln(\hat{y}) + u$, where $\hat{y} = \exp(x\beta)$. The choice of which GLM specification to choose then depends on the estimate of θ; Manning and Mullahy suggest choosing the NLS model if θ = 0, the Poisson class model if θ = 1, the gamma model if θ = 2, or the inverse Gaussian model if θ = 3.

the individual partial effects over the whole sample. It is expected that for FOOD, EDUCATION, and HOUSE, an increase in the OOP share of non-subsistence expenditure would correspond to a reduction in expenditure for each commodity.

In order to determine the adequacy of the 2PMs, several specification checks are performed. First, for the hurdle specification, Ramsey's RESET test is carried out on the logit models for each expenditure equation in the six countries. To check the specification of both the level equation and the full 2PM, a modified Hosmer-Lemeshow type test is performed (Hosmer & Lemeshow, 1995)[3].

EMPIRICAL RESULTS

Results from expenditure share analysis

For each country, the coefficients on CATA2, CATA3, and CATA4 from the expenditure share regressions appear in Table 11. The other coefficients are presented in Table 12.

The results in all countries, except Burkina Faso, are consistent with the prior hypothesis that households protect the amount of minimum food spending when OOP is greater than 40% of households' capacity to pay (CATA4) and the coefficient in Burkina Faso is not statistically significant. The coefficients for CATA2 and CATA3 are negative in Chad and Burkina Faso. In each country, the coefficient of CATA4 is larger than that of CATA2. It is also worth noting that the results support Engel's law, which states that as total expenditure rises, less is spent proportionately on food, except in Chad where the coefficient is positive, but not significant.

For the education share, the coefficients on CATA2 are negative in four countries, but none of these are statistically significant. They are also quite small in magnitude. In four countries, four of the CATA4 coefficients are negative, but only significantly so in Chad. In Burkina Faso and Senegal, CATA4 is significantly positive. The absolute size of the coefficients across countries tends to be smaller than those for the food share. There does not seem to be clear evidence of any type of gradient across countries. In Burkina Faso and Senegal, the coefficients increase with rising CATA levels, while the opposite occurs in Chad. The other countries do not display any constant pattern across OOP levels.

The coefficients on CATA2 and CATA3 are positive in the housing share equations for all countries but Senegal. They are statistically significant in about half of cases. However, the sizes are not very large. For CATA4, three other countries now have negative coefficients although the absolute values still tend to be small and the only statistically significant estimate is Chad. The most notable result is that CATA4 is 0.100 in Burkina

[3] The test is computed as suggested by Deb, Manning, and Norton(Deb, Manning & Norton, 2005) and Buntin and Zaslavasky (Buntin & Zaslavsky, 2004). First, the residuals are computed on the raw scale. Then, predicted values are grouped into deciles and the residuals are regressed on these groups. An *F*-test of the joint significance of all the groups provides a test of whether the mean residuals are different from 0.

Faso and significant at the 5% level. There is some indication of a negative gradient in Chad, Zambia, and Zimbabwe, but the estimates on CATA4 are not statistically significant in the latter two countries.

For the share of expenditure spent on other goods, the coefficients on CATA2 are positive in four countries and statistically significant in two. However, moving to CATA3, the estimates are significantly negative in Zambia and Zimbabwe, but are positive in the other four countries. Only Chad has a statistically significant positive coefficient for CATA3, though, and the magnitude of the others is quite small. For each country, the coefficients on CATA4 are negative and statistically significant at the 1% level. The absolute size of the coefficient also increases with nearly all larger than those for the food, education, and housing share equations. The presence of a negative gradient and the size and significance of the coefficients indicates that households facing higher OOP burdens reduce their share of expenditures spent on goods other than food, housing, and education.

Results from 2PM analysis

APEs for CATA2, CATA3, and CATA4 for 2PMs in each country are presented in Table 13. The APE can be interpreted as the average change in food expenditure for those households in CATA2, CATA3, or CATA4 compared to the base group, holding other variables fixed. Standard deviations appear in parentheses and illustrate the variability of the partial effects across households. The estimates are measured in local currency units and so are not directly comparable across all countries.

Most of the APEs for the CATA variables are negative in the FOOD 2PM, but only statistically significant at the 5 % level or better in Burkina Faso for CATA3 and CATA4 and in Kenya for CATA1. The APEs of CATA3 are negative for all countries. For CATA4, the estimates are once again negative in all countries except Zambia.

There are more positive APEs for education expenditure, and these are all found for CATA2 and CATA3. Only two countries have negative estimates for the APEs of CATA2 and four countries do for CATA3. None of the underlying coefficients for CATA2 or CATA3 are statistically significant for the cases with positive APEs. For each country, the APEs of CATA4 are negative and are statistically significant in most countries. The estimates sequentially decrease with rising CATA levels in five countries. In Chad, the estimate on CATA3 is only barely larger than that for CATA2 and neither of the two underlying coefficients is statistically significant.

Similarly, about half of the APEs for housing expenditure are positive for CATA2 and CATA3. The underlying coefficients are only statistically significant and positive for CATA2 in Burkina Faso, Chad, and Senegal. Similar to EDUCATION, the partial effects for CATA4 are negative in all countries and statistically significant once again in all but Zimbabwe. The negative gradient is most pronounced with housing expenditure as the APEs decrease with subsequent rising CATA levels in each country.

Results of specification tests appear in Tables 14-16 and indicate that the full 2PMs for education expenditure are generally not mispecified, but those for food and especially housing expenditure suffer from various degrees of misspecification.

DISCUSSION

The results of the Engle curve and 2PM analyses provide some important indications of the impact of OOP on household expenditure patterns. The Engel curve analysis indicates relative changes in expenditure shares for a particular good, but not absolute increases or decreases. However, knowledge of whether or not households actually reduce expenditure on a particular commodity is perhaps of more interest to policy makers. While households protect food when OOP comprises high shares of non-subsistence spending by allocating larger proportions of their budget, the 2PMs suggest that the absolute level of food still decreases. Small changes in budget shares for education and housing occur despite high OOP shares, but absolute expenditure on these items also decreases in absolute terms.

An important message to bring to policy makers is that across nearly each country, households with higher OOP shares of non-subsistence spending reduce expenditure on goods that are arguably essential to development and poverty alleviation. While households begin to cope at different levels in different countries, households spending more than 40 percent of non-subsistence expenditure on OOP demonstrate strong signs of consumption modification. The estimates for the coefficients in the Engel curve analysis and the APEs in the 2PMs for CATA4 tend to suggest a similar pattern across countries; food shares increase while shares for education and housing decrease. However, the absolute level of each commodity decreases. The results indicate it is at the level of 40% of non-subsistence spending on health care that households demonstrate the strongest coping behaviour.

The analysis highlights the vulnerability of households with high OOP spending levels and it is the implications of these health shocks that policy-makers should be especially concerned with. Ideally, households would be covered by some form of health insurance or able to access government subsidized services. However, most sub-Saharan African countries do not have these means due to high levels of informal sector employment and limited ability to raise tax revenues, the difficulties with establishing or expanding social health insurance and volatile and uncoordinated donor funds. A number of African governments subsidize public providers and have rolled back their systems of formal user fees. However, primarily as a result of insufficient government payments to providers, informal charges are often levied at the point of service. Furthermore, the patient also pays to non-state providers for services and drugs. As a result, the direct financial costs of illness are still sizeable. Although the ultimate goal for African countries is to reach universal coverage that provides a sufficient benefits package, this is a long-term solution.

This analysis suggests the importance of linkages with welfare protection in other social sectors. High-income countries provide safety nets for the poor that include subsidies for food, education, and housing. In low-income countries, this is more difficult; however, if development requires the combination of proper nutrition, living standards, and increasing human capital through education, the best strategies for financial protection need to be multi-faceted.

Additionally, it is important to recognize the value added of each model in analyzing patterns of consumption modification, but to be cautious in interpreting them jointly. For example, the Engle curve results indicate that households increase their budget share of food relative to other goods when they face the highest CATA level. However, the 2PM suggests that they also reduce food expenditure in CATA4. This is possible since the simultaneous reduction in education, housing, or other goods could be larger. Moreover, the absolute reduction in food expenditure could also be smaller than the reductions in education, housing, and other goods, and the budget share of food could still increase.

This helps to shed some light on certain results that appear to be outliers. For example, the coefficient on CATA4 in Burkina Faso in the Engel curve housing share equation is positive, large (0.100), and statistically significant. However, the APE on CATA4 in the housing 2PM is negative. This is most likely a result of both the APEs for both food and education in the country to be 4 and 2 times as large, respectively, as the housing APE. This may also help to explain the only negative coefficient of CATA4 in the food share equations is in Burkina Faso. Nevertheless, more detailed research would be necessary to draw strong conclusions regarding exactly why this is the case. For each country, the same set of regressors was used in both models for simplicity and generality. However, the presence of any institutional and social idiosyncrasies, in addition to differences in basic household characteristics, would offer compelling reason for different specifications for each country.

Additionally, the difficulty in drawing strong conclusions with regards to education and housing share is partly caused by the zero expenditure problem. Perhaps households spend nothing on education because they do not have any children to send to school and so the zero represents a corner solution.[4] It is also possible that a household does routinely spend on education, but the survey period did not cover the time when expenditure was incurred. Finally, education expenditure may have been reduced to zero in order to cope with high medical expenditure. For housing, expenditure may be more difficult to adjust than expenditure on other commodities. Moving residence is oftentimes not a common activity but interest lies in the cases where such relocation and reduction in housing expenditure occurs to finance medical care. The other housing components like the costs of gas, electricity, water, telephone, and heating fuel are easier to adjust; however, it is possible that many households surveyed in these countries tend not to spend much on these items.

It is clear that households with the highest OOP spending generally reduce spending on other goods and tend to protect food consumption, at least relative to other commodities;

[4] The coefficient on SAGE518 is statistically significant in every country.

however, there is more heterogeneity in the predicted effects of OOP among households with between 10 and 40 percent of their capacity to pay. This may be due to the fact that rising CATA levels may not necessarily represent households that are less financially protected. For example, for any two households, 30 percent of each of their capacity to pay may imply two very different residual absolute spending capacities based on what their total expenditure is. The same could be true for spending more than 40 percent as well, though.

It would be necessary to determine to what extent surpassing a 40 percent threshold indicates greater financial vulnerability, especially if policy-makers seek to effectively target those with the least protection from health shocks. Similarly, the base group, which spends less than 10 percent of its capacity to pay, may be heterogeneous, especially among those reporting zero OOP spending; a household may report a zero because it did not fall ill or because it was too poor to purchase any medical care. It is possible that this is affected by institutional factors of the health system of each country, distance travelled to a health facility, or other factors unaccounted for, which may suggest why CATA2 and CATA3 do not exhibit exactly the same patterns as for CATA4 across countries.

The study methodology could be improved in several ways. For the 2PMs, examining plots of the residual versus predicted values shows signs that systematic departure of the predicted values from the residuals often lies in the higher groups, i.e. those predicted to have a higher value of expenditure. For the log-OLS models, it would be worth considering using a different smearing estimator for these groups. Searching for suitable exclusion restrictions for the level equation in the 2PM may also improve specification.

Additionally, different non-parametric specifications of Engel curves could be explored. The coping strategy variables INS, SAVE, BORROW, SELL, and COPE-OTHER could be interacted with demographic and socioeconomic variables, such as URBAN, to reflect differences in availability of financial sources to pay for OOP. The indirect costs of illness, such as lost income and the cost of travel to a health facility, could also be included to give a more complete picture of the financial cost of illness.

Finally, panel data would enable greater insight into the coping mechanisms used by households and the degree to which grouping households based on OOP spending as a share of capacity to pay (rather than, say, their distance from some poverty line) adequately reflect differences in financial protection among the population could be explored.

CONCLUSION

This analysis has shown that households in the six African countries examined are likely to reduce expenditure on basic goods, including food, education, and housing, when faced with high OOP. Food tends to be protected while spending on other goods is sacrificed, at least relatively. The results of this analysis offer a snapshot of household expenditure adjustments to cope with OOP and further research could build upon this analysis in several ways. From a policy perspective, if the objective is to design mechanisms that help people escape poverty, a dynamic research approach is necessary to determine what programs would best shield the negative impact of high medical costs. Nevertheless, this research highlights the importance of a multi-pronged approach towards development policy if adequate amounts of food, housing, and education are necessary conditions to escape poverty.

REFERENCES

Asfaw,A., & Von Braun,J. (2004). Is consumption insured against illness? Evidence on vulnerability of households to health shocks in rural Ethiopia. *Economic Development and Cultural Change 53*(1), 115-129.

Baum,C., Schaffer,M., & Stillman,S. (2003). Instrumental Variables and GMM: Estimation and Testing. *The Stata Journal 3*(1), 1-33.

Blundell,R. (1988). Consumer Behaviour: Theory and Empirical Evidence – A Survey. *The Economic Journal 98*(389), 16-65.

Buntin,M., & Zaslavsky,A. (2004). Too much ado about two-part models and transformation? Comparing methods of modelling Medicare expenditures. *Journal of Health Economics 23* 525-542.

Cameron,A., & Trivedi,P. Microeconometrics: Methods and Applications. 2005. New York, Cambridge University Press.

Carrin,G., Gray,E., & Almeida,J. (1999). Coping with Ill Health in a Rickshaw Puller's Household in Chittagong, Bangladesh. *Southeast Asian Journal of Tropical Medicine and Public Health*(30), 136-148.

Deb,P., Manning,W., & Norton,E. Modeling Health Care Costs and Counts. iHEA Barcelona Conference . 2005.

Desmet,M. Coping with the costs of illness in slum households in Bangladesh: an empirical analysis of the relationship between income distribution and household behaviour. 2000. Universiteit van Amsterdam.

Duan,N. (1983). Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association 78* 605-610.

Gertler,P., & Gruber,J. (2002). Insuring Consumption Against Illness. *American Economic Review 92*(1), 51-70.

Gertler,P., Levine,D.I., & Moretti,E. (2002). Do Microfinance Programs Help Families Insure Consumption Against Illness?

Hansen,L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica 50*(3), 1029-1054.

Hosmer,D., & Lemeshow,S. (1995). *Applied Logistic Regression.* New York: John Wiley & Sons .

Lewbel,A. (1996). Demand Estimation with Expenditure Measurement Errors on the Left and Right Hand Side. *The Review of Economics and Statistics 78*(4), 718-725.

Manning,W., & Mullahy,J. (2001). Estimating log models: to transform or not to transform? *Journal of Health Economics 20*(4), 461-494.

Pitayanon,S., Kongsin,S., & Janjareon,W. (1997). The economic impact of HIV/AIDS mortality on households in Thailand. In D.Bloom, & P.Godwin (Eds.), *The Economics of HIV and AIDS: The Case of South and South East Asia* Delhi: Oxford University Press.

Pryer,J. Strategies to cope with the cost of work disabling illness among household heads. 2003. Ashgate. Poverty and Vulnerability in Dhaka slums: the urban Livelihood study.

Russell,S. (1996). Ability to pay for health care: concepts and evidence. *Health Policy and Planning 11(3)* 219-237.

Townsend,R. (1995). Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies. *Journal of Economic Perspectives 9*(3), 83-102.

Wilkes,A., Hao,Y., Bloom,G., & Xingyuan,G. Coping with the Costs of Severe Illness in Rural China. (58). 1997. Brighton, University of Sussex. IDS Working Paper .

Xu,K., Aguilar,A., Carrin,G., & Evans,D. Distribution of health payments and catastrophic expenditures-Methodology. (2). 2005. WHO Health Financing Policy discussion paper.

Table 1 - Sample Sizes for WHS Datasets

Country	Sample Size
Burkina Faso	4,928
Chad	4,797
Kenya	4,594
Senegal	3,355
Zambia	4,157
Zimbabwe	4,149

Table 2 – Variable Labels, Definitions, and Sample Means

Variable Label	Definition	Burkina Faso	Kenya	Chad	Senegal	Zambia	Zimbabwe
TEXP	**Total expenditure without OOP**	41,278	7,478	40,119	102,861	223,372	76,912
FOOD	Food expenditure	25,002	3,346	33,084	61,974	140,371	46,023
EDUCATION	Education expenditure	3,527	1,554	1,653	5,753	19,059	8,741
HOUSE	Housing expenditure	5,225	1,653	2,042	26,707	26,622	7,878
OTHER	Expenditure on all other goods	7,228	926	3,339	8,427	37,320	14,271
OOP	OOP expenditure	5,137	640	2,811	9,592	3,934	3,444
CATA2	= 1 if 0.1 ≤ OOP/CTP < 0.2 = 0 otherwise	0.09	0.08	0.03	0.07	0.04	0.03
CATA3	= 1 if 0.2 ≤ OOP/CTP < 0.4 = 0 otherwise	0.11	0.09	0.05	0.09	0.04	0.03
CATA4	= 1 if 0.4 ≤ OOP/CTP = 0 otherwise	0.21	0.13	0.17	0.14	0.06	0.04
BORROW	= 1 if household borrowed to pay for OOP = 0 otherwise	0.18	0.13	0.07	0.14	0.03	0.07
SAVE	= 1 if household used savings to pay for OOP = 0 otherwise	0.07	0.07	0.01	0.04	0.01	0.04
SELL	= 1 if household sold assets to pay for OOP = 0 otherwise	0.25	0.1	0.1	0.05	0.05	0.08
INS	= 1 if household used insurance to pay for OOP = 0 otherwise	0.01	0.03	0.01	0.03	0.01	0.03
COPE-OTHER	= 1 if household used another coping strategy to pay for OOP = 0 otherwise	0.09	0.03	0.07	0.04	0.02	0.02
URBAN	= 1 if household lives in urban area = 0 otherwise	0.41	0.32	0.25	0.48	0.41	0.36
EQSIZE	**(Household size)$^{0.56}$**	2.57	2.15	2.37	3.11	2.49	2.37
SAGE518	Share of household between age 5 and 18	0.31	0.27	0.31	0.30	0.34	0.32
AGE	Age of household head	43.18	42.7	41.51	51.13	42.57	44.22
AGE2	$AGE^2/100$	21.04	20.77	19.88	28.34	20.3	22.24

Table 3 - Percentage of Observations with Zero Values by Expenditure Item

Country	FOOD	EDUCATION	HOUSE	OTHER	OOP
Burkina Faso	1%	81%	50%	52%	46%
Chad	5%	83%	83%	67%	72%
Kenya	1%	54%	21%	56%	56%
Senegal	4%	63%	27%	63%	59%
Zambia	1%	73%	44%	47%	77%
Zimbabwe	3%	58%	39%	57%	82%

Table 4 - Overidentification Tests Using Hansen's J-Statistic

	Food		Education		Housing		Other	
Country	$\chi^2_{(1)}$	*p*-value	$\chi^2_{(1)}$	*p*-value	$\chi^2_{(1)}$	*p*-value	$\chi^2_{(1)}$	*p*-value
Burkina Faso[1]	0.074	0.786	0.293	0.589	0.428	0.513	0.306	0.580
Chad[1]	0.394	0.530	0.400	0.527	0.903	0.342	0.806	0.369
Kenya[1]	1.192	0.551	1.080	0.583	4.053	0.132	0.243	0.886
Senegal[2]	2.688	0.101	3.551	0.060	0.083	0.774	0.832	0.362
Zambia[2]	0.016	0.901	3.635	0.057	1.667	0.197	3.399	0.065
Zimbabwe[1]	3.697	0.055	1.402	0.236	1.466	0.226	0.377	0.539

1. excluded instruments for total expenditure include dummy variables for education level of household head

2. excluded instruments for total expenditure include dummy variables for education level and insurance status of household head. For Senegal, there are 3 excluded instruments and so the test statistics are compared to a $X^2_{(2)}$.

Table 5 - *F*-statistics from OLS Regressions of TEXP2 on excluded instruments

Country	*F*-statistic
Burkina Faso	F(2, 4915) = 275.84
Kenya	F(2, 4589) = 392.18
Senegal	F(2, 3312) = 79.50
Chad	F(2, 4614) = 44.18
Zambia	F(2, 4135) = 398.40
Zimbabwe	F(2, 4104) = 162.59

II.1. Table 6 - Hausman Test of 3SLS vs SURE

Country	$\chi^2_{(16)}$	*p*-value
Burkina Faso	132.42	0.000
Chad	9.10	0.909
Kenya	96.16	0.000
Senegal	67.82	0.000
Zambia	64.84	0.000
Zimbabwe	23.97	0.090

3SLS is the always-consistent estimator. SURE is inconsistent under the alternative hypothesis and efficient under the null

Table 7 - White Test for Heteroskedasticity in Logarithm-scale Error

	Food		Education		Housing	
Country	X^2	*p*-value	X^2	*p*-value	X^2	*p*-value
Burkina Faso	1594.73	0.000	144.26	0.002	126.03	0.079
Chad	105.99	0.482	115.6	0.247	68.13	0.998
Kenya	249.97	0.000	199.72	0.000	334.34	0.000
Senegal	63.72	1.000	107.75	0.434	160.13	0.000
Zambia	1736.84	0.000	275.75	0.000	652.57	0.000
Zimbabwe	176.41	0.000	124.67	0.104	131.65	0.046

Table 8 - Results of modified Park test for food expenditure model

		P-values from tests on θ				
Country	$\hat{\theta}$	θ = 0	θ = 1	θ = 2	θ = 3	GLM family chosen
Burkina Faso	1.699	0	0	0.004	0	Poisson or Gamma
Chad	2.218	0	0	0.099	0	Gamma
Kenya	1.773	0	0	0.002	0	Poisson or Gamma
Senegal	1.639	0	0	0.005	0	Poisson or Gamma
Zambia	1.972	0	0	0.737	0	Gamma
Zimbabwe	1.694	0	0	0	0	Poisson or Gamma

Table 9 - Results of modified Park test for education expenditure model

		P-values from tests on θ				
Country	$\hat{\theta}$	θ = 0	θ = 1	θ = 2	θ = 3	GLM family chosen
Burkina Faso	2.152	0.000	0.000	0.056	0.000	Gamma
Chad	2.000	0.000	0.000	0.999	0.000	Gamma
Kenya	1.975	0.000	0.000	0.643	0.000	Gamma
Senegal	1.828	0.000	0.000	0.065	0.000	Gamma
Zambia	2.156	0.000	0.000	0.005	0.000	Gamma or inverse Gaussian
Zimbabwe	2.220	0.000	0.000	0.000	0.000	Gamma or inverse Gaussian

Table 10 - Results of modified Park test for housing expenditure equation

Country	$\hat{\theta}$	P-values from tests on θ θ = 0	θ = 1	θ = 2	θ = 3	GLM family chosen
Burkina Faso	1.975	0.000	0.000	0.595	0.000	Gamma
Chad	2.051	0.000	0.000	0.755	0.000	Gamma
Kenya	2.012	0.000	0.000	0.75	0.000	Gamma
Senegal	1.593	0.000	0.000	0.000	0.000	Poisson or Gamma
Zambia	1.973	0.000	0.000	0.607	0.000	Gamma
Zimbabwe	1.776	0.000	0.000	0.014	0.000	Poisson or Gamma

Table 11 - Coefficients on CATA variables by share equation

	Food			Education			Housing			Other		
Country	CATA2	CATA3	CATA4	CATA2	CATA3	CATA4	CATA2	CATA3	CATA4	CATA2	CATA3	CATA4
Burkina Faso	-0.037**	-0.039**	-0.006	-0.002	0.001	0.027**	0.001	0.029**	0.100**	0.039**	0.011	-0.120**
Chad	-0.121**	-0.094**	0.057**	0.035**	0.034**	-0.014**	0.025*	0.009	-0.014*	0.061**	0.051**	-0.029**
Kenya	0.005	-0.029*	0.034*	-0.014	0.004	-0.013	0.022*	0.024*	0.008	-0.013	0.002	-0.030**
Senegal	0.022	-0.001	0.032	-0.011	0.002	0.017*	-0.016	-0.002	-0.018	0.005	0.001	-0.030**
Zambia	0.006	0.023	0.131**	0.008	0.015	-0.004	0.006	0.023*	-0.012	-0.020	-0.060**	-0.115**
Zimbabwe	-0.019	0.022	0.127**	-0.019	-0.004	-0.013	0.035*	0.038*	-0.002	0.003	-0.056**	-0.111**

*** = significant at 1% level ** = significant at 5% level * = significant at 10% level

Table 12 - Coefficients from 3SLS and SURE Regressions for Food Share

Country	TEXP2	URBAN	EQSIZE	SAGE518	BORROW	SAVE	SELL	INS	COPE-OTHER	AGE	AGE2	Constant
BFA	-0.23**	0.021	0.076**	-0.037*	-0.019	0.016	0.034**	0.082	-0.007	0.002	-0.001	2.924**
TCD	0.005	-0.127**	-0.010*	-0.043**	0.003	-0.064	-0.012	-0.008	0.014	0.000	0.001	0.867**
KEN	-0.207**	0.046**	0.079**	-0.024	-0.010	0.034*	-0.011	0.061*	-0.014	0.001	0.000	2.129**
SEN	-0.208**	-0.024	0.067**	-0.048	0.028	-0.002	0.001	-0.019	-0.013	0.001	0.000	2.752**
ZMB	-0.097***	-0.008	0.017	-0.018	-0.012	-0.110**	0.029	-0.051	-0.061**	-0.002	0.002	1.886**
ZWE	-0.081**	-0.001	0.023*	-0.075**	-0.019	-0.017	-0.049**	-0.048	-0.029	0.001	-0.002	1.516**
Coefficients from 3SLS and SURE Regressions for Education Share												
BFA	0.076**	-0.014*	-0.015**	0.029**	0.000	-0.003	-0.017**	-0.039	-0.010	0.001	0.000	-0.732**
TCD	-0.011**	0.041**	0.021**	0.017*	-0.007	-0.004	0.010	0.015	0.011	0.000	0.000	0.077**
KEN	0.112**	-0.080**	-0.023**	0.101**	0.019*	0.006	0.041**	-0.058**	-0.008	0.003**	-0.003**	-0.843**
SEN	0.074**	-0.013*	-0.015**	0.048**	-0.020**	0.005	0.008	0.010	0.014	-0.001	0.000	-0.729**
ZMB	0.044**	-0.014*	-0.003	0.045**	0.004	0.067**	0.008	0.017	0.001	0.000	0.001	-0.489**
ZWE	0.050**	-0.024*	-0.006	0.106**	-0.012	-0.001	0.036**	-0.020	-0.007	0.001	-0.001	-0.503**
Coefficients from 3SLS and SURE Regressions for Housing Share												
BFA	0.164**	0.036**	-0.067**	0.020	-0.019**	-0.049**	-0.050**	-0.038	0.008	-0.001	0.001	-1.430**
TCD	-0.003	0.093**	-0.005*	0.020*	0.011	0.012	-0.012	-0.009	-0.006	0.000	0.000	0.063*
KEN	0.062**	0.080**	-0.040**	-0.042**	-0.004	-0.008	-0.033**	-0.014	-0.016	-0.004**	0.003**	-0.179**
SEN	0.134**	0.048**	-0.048**	0.007	-0.010	0.004	-0.011	0.006	-0.002	-0.002	0.001	-1.098**
ZMB	0.046**	0.046**	-0.014**	-0.011	0.006	0.034	-0.032**	0.005	0.015	-0.001	0.001	-0.409**
ZWE	0.021*	0.048**	-0.015**	-0.005	0.000	-0.014	-0.011	0.005	0.015	-0.002**	0.002**	-0.062
Coefficients from 3SLS and SURE Regressions for Other Share												
BFA	-0.002	-0.043**	0.006	-0.012	0.038**	0.036*	0.033**	-0.005	0.008	-0.002	0.000	0.235
TCD	0.009**	-0.008	-0.004	0.006	-0.006	0.055*	0.014	0.003*	-0.019	0.000	-0.001	-0.009
KEN	0.033**	-0.046**	-0.016**	-0.035**	-0.006	-.0.032**	0.003	0.012	0.037**	0.000	0.000	-0.106*
SEN	-0.001	-0.011	-0.003	-0.008	0.002	-0.007	0.003	0.003	0.001	0.001	-0.001	0.076
ZMB	0.007	-0.024*	0.000	-0.017	0.002	0.009	-0.005	0.028	0.045**	0.002**	-0.003**	0.012
ZWE	0.009	-0.024*	-0.002	-0.025	0.031	0.032	0.024	0.063*	0.022	0.000	0.000	0.048

** = significant at 1% * = significant at 5%

Table 13 - Average Partial Effects of CATA Variables

Country	CATA2	CATA3	CATA4
Food Expenditure 2PMs			
Burkina Faso	-108 (131) *	-2,591 (983) **	-5,344 (2,068) ***
Chad	-181 (70)	-1,302 (713) *	-7,467 (2,494)
Kenya	-61 (45) **	-485 (302)	-186 (120)
Senegal	6,133 (2,357)	-197 (335)	-1,810 (812) *
Zambia	-9,882 (5,082)	-17,151 (8,344)	1,135 (562)
Zimbabwe	-5,120 (2,908)	-13,565 (7,856)	-7,821 (4,597) *
Education Expenditure 2PMs			
Burkina Faso	329 (413)	-1,007 (2,045) ***	-3,164 (5,211) ***
Chad	1,741 (1,879)	1,852 (1,945)	-453 (1,051) ***
Kenya	-655 (1,414) ***	-970 (1,970) ***	-2,045 (3,644) ***
Senegal	2,078 (730)	1,498 (843)	-3,094 (3,183) *
Zambia	-18,328 (253,045) *	-23,534 (310,078)	-57,599 (700,355) ***
Zimbabwe	1,322 (1,668)	-1,925 (2,211) **	-5,811 (5,507)
Housing Expenditure 2PMs			
Burkina Faso	1,113 (1,555)**	470 (973)	-1,338 (1,731) **
Chad	1,806 (2,526) **	116 (1,492)	-1,353 (2,322) ***
Kenya	-42 (181)	-212 (533) **	-849 (1,610) ***
Senegal	2,397 (1,798) *	-4,535 (5,866)	-20,497 (11,035) ***
Zambia	-2,195 (6,759) *	-6,661 (10,666) ***	-16,766 (17,727) ***
Zimbabwe	3,575 (1,993)	1,808 (626)	-2,448 (1,380)

* = coefficient in level equation significant at 10% ** = coefficient in level equation significant at 5%
*** = coefficient in level equation significant at 1%

Table 14 - Specification Tests for FOOD 2PM equations

Country	Ramsey's RESET on hurdle p-value	Hosmer Lemeshow on level p-value	Hosmer Lemeshow on 2PM p-value
Burkina Faso	0.931	0.008	0.001
Chad	0.027	0.188	0.047
Kenya	0.644	0.004	0.001
Senegal	0.823	0.155	0.102
Zambia	0.052	0.003	0.000
Zimbabwe	0.276	0.018	0.046

Table 15 - Specification Tests for EDUCATION 2PM equations

	Ramsey's RESET on hurdle	Hosmer Lemeshow on level	Hosmer Lemeshow on 2PM
Country	p-value	p-value	p-value
Burkina Faso	0.028	0.195	0.177
Chad	0.093	0.645	0.158
Kenya	0.000	0.380	0.051
Senegal	0.601	0.013	0.211
Zambia	0.265	0.024	0.000
Zimbabwe	0.001	0.000	0.392

Table 16 - Specification Tests for HOUSE 2PM equations

	Ramsey's RESET on hurdle	Hosmer Lemeshow on level	Hosmer Lemeshow on 2PM
Country	p-value	p-value	p-value
Burkina Faso	0.047	0.007	0.000
Chad	0.226	0.312	0.001
Kenya	0.060	0.000	0.000
Senegal	0.273	0.000	0.005
Zambia	0.394	0.126	0.000
Zimbabwe	0.000	0.000	0.000