Crowding Out Effect of Expenditure on Tobacco in Zambia: Evidence from the Living Conditions Monitoring Survey.

Grieve Chelwa[1]

25th August, 2013

Abstract: Tobacco consumption is widely recognised as a leading cause of preventable death. In addition to direct mortality and morbidity costs, tobacco use has been shown to crowd-out expenditure on other commodities especially among income constrained households. This paper formally tests the crowding-out hypothesis in Zambia using data from the 2006 round of the Living Conditions Monitoring Survey (LCMS). We estimate conditional Engel curves using the Quadratic Almost Ideal Demand System (QUAIDS) developed by Banks, Blundell and Lewbel (1997) where expenditure on tobacco is the conditioning variable. We use instrumental variables techniques to control for the likely endogeneity of tobacco expenditure in our demand equations. We also control for likely differences in the structure of preferences between smoking and non-smoking households. Our results show that tobacco expenditure tends to crowd-out spending on food, transportation and transport equipment maintenance in urban areas and schooling, electricity, transportation, transport equipment maintenance and entertainment in rural areas. Tobacco spending crowds-in spending on alcohol in both rural and urban areas suggesting that the crowding-out effects of tobacco are likely to be amplified through their positive impacts on alcohol spending. Our results also reveal interesting socio-economic patterns.

JEL Codes: D10, D11, D12, D13

[1] PhD Candidate, School of Economics, University of Cape Town. Email: grievechelwa@gmail.com. Paper prepared for the Biennial Economic Society of South Africa (ESSA) conference in Bloemfontein, South Africa.

1.0. Introduction

Tobacco use is the leading cause of premature death in the world. According to the World Health Organization, 5 million people per annum die from tobacco related deaths (WHO, 2010) including about 600,000 from passive smoking (Oberg, et al., 2010). There is another aspect to the cost of tobacco use that is separate from the mortality and morbidity costs, namely that tobacco use tends to crowd-out the consumption of other commodities. This additional cost has received little attention in the economics of tobacco control literature. This paper adds to the literature on the crowding out effect of tobacco in the context of Zambia. Our paper shows that spending on tobacco reduces the shares of total monthly expenditure dedicated to food, transportation and transport equipment maintenance in urban areas. In rural areas, tobacco expenditure crowds-out expenditure on schooling, electricity, transportation, transport equipment maintenance and entertainment. Further, our analysis shows that spending on tobacco tends to crowd-in (increase) the share of monthly expenditure dedicated to alcohol suggesting that the crowding-out effects of tobacco are likely to be amplified through tobacco's positive impact on alcohol spending. Our analysis also reveals interesting social-economic patterns. For instance, in urban areas, food is only crowded-out among poorer households and not among rich households. Further, the crowding-in effect of alcohol disappears when one looks at poor households in both urban and rural areas suggesting that tobacco and alcohol likely compete for resources among the poor. Our results also suggest that the degree of crowding-out is likely to be more severe among rural households than among urban households. To the best of our knowledge, this is the first paper to investigate whether tobacco expenditure has a crowding-out effect in Zambia.

Zambia's per capita cigarette consumption has recently started rising after declining over much of the 1990s (Chelwa, 2012). For the period 1990 to 2001, per capita cigarette consumption declined by 75%. From the 2002 to 2009, per capita cigarette consumption increased by 37% (ibid). The expectation is that per capita cigarette consumption will continue to increase in the current decade driven by population growth and favourable economic performance (ERC, 2010). Similarly, the prevalence of smoking has increased over the last decade. The 2002 Zambia Demographic and Health Survey (ZDHS) reported male and female adult smoking prevalence at 14% and 0.5% respectively (ZDHS, 2002). By 2007, male adult smoking prevalence had increased to 24% while that of females had increased, though marginally, to 0.7% (ZDHS, 2007). These numbers suggest that tobacco consumption has increasingly become important in the expenditure decisions of households. In addition, the fact that male smoking prevalence is significantly higher than female prevalence has implications for intra-household resource allocation. Deaton (1997) points out that in developing countries, the consumption interests of children and women are usually subordinated to those of men. This is likely to be the case with tobacco consumption where adult males in a household might predetermine the budget allocation to tobacco before allocations are made to other commodities. In this way, tobacco consumption might crowd-out the consumption of other commodities.

The paper investigates the crowding-out effect of tobacco consumption by estimating a number of demand systems following the method used in John (2008) and using data from the 2006 round of Zambia's Living Conditions and Monitoring Survey. The demand systems are estimated using the Quadratic Almost Ideal Demand System (QUAIDS) pioneered by Banks et al (1997). We prefer the QUAIDS as it allows commodities to be modelled as luxuries and necessities at different levels of

household expenditure. The QUAIDS also nests the popular Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980). The demand equations in our system are conditional on monthly tobacco expenditure and in this way capture the possibility that decisions pertaining to tobacco expenditure are made before decisions to spend on other commodities. In any case, we control for possible simultaneity bias (or reverse causality) between tobacco expenditure and the demand for other commodities by using instrumental variables techniques. Our estimation technique also controls for possible preference heterogeneity between smoking and non-smoking households using the method developed by Vermeulen (2003). Vermeulen's method is in essence a test for whether households reporting zero tobacco expenditure do so because of corner solutions or abstention. In the former case, households would want to purchase tobacco but are constrained from doing so by prices and/or income. That is, tobacco consumption enters the utility function of the household regardless of the household's tobacco consumption status. In the latter case of abstention, tobacco is not an argument in a household's utility function. Further, Vermeulen's test for preference heterogeneity also delivers a test for weak separability that is closely related to the concept of demographic separability of Deaton et al. (1989). If two commodities are weakly separable, then re-allocating expenditure to one of the commodities only generates an income effect. If two commodities are not weakly separable then re-allocating expenditure to one of the commodities generates both an income and substitution effect.

The rest of the paper is structured as follows: Section 2 gives an overview of the relevant literature. Section 3 discusses the theoretical framework behind conditional demand functions as well as defining what is meant by preference heterogeneity. Section 4 discusses our choice of empirical strategy and the data is described in section 5. We present the empirical results in section 6 and section 7 concludes.

2.0. Literature review

Efroymson, et al (2001) was one of the first studies to highlight the potential crowding-out effect of tobacco consumption using several datasets from Bangladesh. Theirs was not so much an econometric study as a simple comparison of the expenditure profiles of smoking households versus non-smoking. One of the study's main findings was that male cigarette smokers spent more than twice as much on cigarettes as per capita expenditure on clothing, housing, health and education combined. Further, the typical smoker could add over 500 calories to the diet of one or two children with his/her daily cigarette consumption. Efroymson et al, however, only used descriptive statistics in drawing out their conclusions and in this way did not account for possible preference heterogeneity between smoking and non-smoking households. Secondly, it is possible that the decision to consume tobacco and not to consume milk, say, is driven by some other factor independent of the two such that the quantities of both commodities are determined endogenously. If preference heterogeneity and endogeneity are not controlled for, then it is not directly evident that reducing a smoking household's expenditure on tobacco will elevate that household's consumption profile to match that of a non-smoking household. Busch et al (1994) use a more sophisticated econometric method to investigate whether tobacco expenditure has crowding-out effects using data from the Consumer Expenditure Survey (CES) in the United States. Specifically, they test for crowding out by estimating own and cross-price elasticities for cigarettes and other

budget categories using an Almost Ideal Demand System while controlling for socio-economic variables and state-level fixed effects. They find that in comparing smoking and non-smoking households, cigarette consumption tends to crowd-out other household expenditures. Specifically, they find that food and cigarettes are substitutes and so are apparel and cigarettes. That is, the consumption of the two tends to decline when cigarette consumption rises. Busch et al recognise that their analysis suffers from two limitations: failure to account for preference heterogeneity and possible endogeneity in their demand system.

More recently, Block and Webb (2009) investigate the crowding-out effect of tobacco on food consumption and by consequence child malnutrition using a detailed household survey in rural Java, Indonesia. The first part of their paper conducts a series of non-parametric analyses where it is established that households with at least one smoker allocate 68% on average of their budget towards food while non-smoking households allocate 75% on average of their budget towards food. After establishing the *prima facie* evidence, the authors proceed to formally test whether tobacco crowds-out food. The authors acknowledge the likely endogeneity of cigarette consumption in structural equations with food on the left-hand side and cigarettes expenditure on the right-hand side. In the absence of appropriate instruments, the authors estimate a series of reduced form equations for food expenditure, cigarette consumption and child height against a common set of covariates. The idea is that if a set of common covariates reduces the allocation to food and reduces child height but increases the allocation to cigarettes, then this is suggestive of crowding-out. Their indirect empirical strategy "demonstrate[s] that the same exogenous covariates that are associated with improved dietary quantity and quality are also associated with reduced allocation of household resources to tobacco" (ibid, p. 18). Block and Webb, however, do not to control for possible preference heterogeneity between smoking and non-smoking households.

The one study whose empirical strategy controls for both endogeneity and preference heterogeneity is John (2008). John estimates conditional demand functions for different commodities conditioning on tobacco expenditure using survey data from the 1999/2000 round of India's National Sample Survey. He augments his conditional demand functions with an indicator variable indicating whether a household reports positive expenditure on tobacco or not. The indicator variable tests and controls for preference heterogeneity between smoking and non-smoking households. John's econometric analysis shows that tobacco spending households have lower consumption of milk, education, clean fuels and entertainment, goods likely to have a bearing on the well-being of women and children in the household. He also finds that tobacco expenditure has a negative effect on per capita nutritional intake. His results on crowding-out are similar for rich and poor households.

3.0. **Theoretical framework**

In classical utility theory, a consumer (or household) chooses the quantity of a good to be consumed, x_i, as a result of the maximization problem:

$$Max\ U = u\ (x_1, \dots, x_n; \boldsymbol{a}\ ) \qquad (1)$$

subject to the budget constraint $\sum_{i=1}^{n} p_i\ x_i = Y$, where p_i is the price of the ith good, $\boldsymbol{a}$ is a vector of individual (or household) characteristics and Y is total individual (or household) income. The

maximization problem yields standard Marshallian demand functions of the form: $x_i = x_i(p_i, \dots, p_n; Y; \boldsymbol{a})$. The Marshallian demand functions imply that the quantity consumed of a commodity, x_i, is a function of its own price, p_i, the prices of the other $n-1$ commodities, total income Y and characteristics $\boldsymbol{a}$.

Now suppose that an individual (or household) that consumes the nth commodity first decides on the quantity to be consumed of that commodity before deciding on the quantities of the other $n-1$ commodities. In such a situation, the individual's (or household's) maximization problem for the other $n-1$ commodities becomes:

$$Max\ U = u\ (x_1, \dots, x_{n-1};\ \widehat{x_n}\ ;\ \boldsymbol{a}\ ) \qquad (2)$$

subject to $\sum_{i=1}^{n-1} p_i\ x_i = M$ where $M = Y - \ p_n \widehat{x_n}$ (or the income left-over after purchasing the nth commodity) and $\widehat{x_n}$ is the already determined quantity of the nth commodity. This maximization problem results in $n-1$ conditional demand functions, with the quantity demanded for each x_i $(i = 1,2, \dots, n-1)$ dependent on $\widehat{x_n}$. More specifically, the conditional demand functions take on the form: $x_i = g_i(p_i, \dots, p_{n-1}; M; \boldsymbol{a};\ \widehat{x_n}\ )$ for $i \neq n$. Remark that the quantity of the nth commodity, even though already determined, might be the result of a separate maximization problem, albeit one that happens earlier. A complete demand system, therefore, might contain a demand equation for the nth commodity where its determinants (regressors) need not be identical with those of the other $n-1$ commodities[2]. It also important to point out that the utility function of a household that does not consume the nth commodity would take the following form:

$$u\ (x_1, \dots, x_{n-1}; \boldsymbol{a}\ ) \qquad (3).$$

Any comparison of the consumption profiles of households with utility functions represented by (2) with those of households with utility functions represented by (3) should control for the inherent difference in the structure of preferences between the two types of households.

Pollak (1969) formally introduced and discussed the properties of conditional demand functions. Browning and Meghir (1991) argue that estimating conditional demand systems is especially useful in those instances where the commodity of interest, such as tobacco, is not consumed by many households.

4.0. Empirical Strategy

We conduct our empirical analysis in two parts. In the first part, we compare the mean expenditure shares for various commodities between smoking and non-smoking households. In particular, the comparisons are conducted for the following commodities: food, alcohol, healthcare, schooling, water, housing, electricity, alternative energy sources, transportation, transport equipment

[2] In estimating demand systems, it is not necessary that all the equations in the system have the same regressors.

maintenance (boats, cars, motor bikes and bicycles), telephone usage, entertainment, house care, personal care and a final category we call "other"[3].

The second part of our empirical strategy formally tests the crowding-out hypothesis of tobacco expenditure for the commodities listed in the previous paragraph. To do so we estimate conditional Engel curves using the Quadratic Almost Ideal Demand System (QUAIDS) developed by Banks, et al (1997). The QUAIDS has the advantage of not only being consistent with utility theory, but also nesting the popular Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980) and further allows commodities to be modelled as luxuries at some income levels and necessities at others. To account for likely preference heterogeneity between smoking and non-smoking households and to simultaneously test for whether zero tobacco expenditures arise from corner solutions or abstention, we augment our conditional Engel curves with an indicator variable *d* which takes on the value of 1 if a household reports positive tobacco expenditure and zero otherwise. More specifically, we estimate a system of conditional Engel curves with each Engle curve taking-on the following form:

$$w_i = (\alpha_{1i} + \alpha_{2i}d + \alpha_{3i}Tobacco + \boldsymbol{\delta_i'a}) + (\beta_{1i} + \beta_{2i}d)\, lnM + (\lambda_{1i} + \lambda_{2i}d)(lnM)^2 \qquad (4)$$

where w_i is the monthly expenditure share of commodity i in total monthly expenditure and lnM is the natural logarithm of total monthly expenditure (in Zambian Kwacha) on the other commodities except tobacco. $Tobacco$ is the monthly expenditure on tobacco in Zambian Kwacha. *d* is an indicator variable that takes on the value of one if a household reports positive tobacco expenditure and zero otherwise. $\boldsymbol{a}$ is a vector of household-specific characteristics that includes the natural logarithms of household size, age of household head, average age of adults in a household, average age of children in a household, years of schooling for household head and years of schooling for the most educated member of the household. Other household characteristics in $\boldsymbol{a}$ include the proportion of adults in a household (household structure), the number of employed persons in a household and a dummy variable for whether the household head receives a wage (or regular) income. We define adults as those who are 18 years old or older. $\boldsymbol{a}$ also includes a number of indicator variables for the type of household as classified by the local authority in which the household is located. Rural households are classified as small-scale agriculture, medium-scale agriculture, large-scale agriculture or non-agriculture households. Urban households are classified as low-cost household, medium-cost household and high-cost household. We conduct the analysis separately for urban and rural households.

Crowding-out is established if α_{3i} (the coefficient on $Tobacco$) in equation (4) is negative and significant. The hypothesis that zero tobacco expenditure arises from corner solutions is rejected if all the coefficients associated with the indicator variable *d* in equation (4) are jointly significant. We would then conclude that the zeroes are a result of abstention. The joint significance of these coefficients (i.e. rejection of the hypothesis of corner solutions) would suggest that smoking and non-smoking households behave differently and augmenting Engle curves with the *d* variable possibly controls for such preference heterogeneity. As stated by Vermeulen (2003), failure to reject the hypothesis of corner solutions, however, is not sufficient for ruling out abstention. It is possible

[3] This category contains goods and services that are difficult to classify or small enough to stand alone. The category contains expenditure shares on postage stamps, house repairs, batteries, laundry service and candles.

that smokers and non-smokers have the same preferences for the rest of the commodity bundle with the exception that tobacco enters the full utility function of smokers but does not enter that of non-smokers. In other words, we are on less certain ground as to the source of the zeroes once we fail to reject the hypothesis of corner solutions.

There is a possibility that $Tobacco$ and M in equation (4) are endogenous, in the sense that they are both choice variables. Such a situation would preclude our giving a causal interpretation to our demand system estimates. In this case, it is desirable to use instrumental variables to ensure consistent and unbiased estimates. John (2008) and Pu et al (2008) instrument tobacco expenditure with the adult-sex ratio in a household. We instead use the proportion of adult males in adult house size as an instrument for tobacco expenditure. Our instrument and that used in John (2008) and Pu et al (2008) are qualitatively similar and produce qualitatively similar results. We prefer the proportion of adult males in adult house size as our instrument because it is a "stronger" instrument. This choice of instrument is driven by the fact that tobacco consumption is largely an adult male affair in Zambia. According to the most recent round of the Zambia Demographic and Health Survey (ZDHS, 2007), smoking prevalence among adult males was estimated at 24% while among adult women it was estimated at 0.7% (for the 2002 round of the survey, adult male and adult female smoking prevalence was estimated at 15% and 0.5% respectively). We expect the proportion of adult males to be positively correlated with a household's expenditure on tobacco but not necessarily correlated with expenditure shares dedicated to other commodities. We follow John (2008), Pu et al (2008), Vermeulen (2003) and Keen (1986) by instrumenting M, the residual expenditure, with total household expenditure.

To implement the instrumenting technique outlined above, we estimate the system in equation (4) by Three Stage Least Squares (3SLS) combined with Seemingly Unrelated Regression (SURE). The SURE allows us to account for any within-household correlation of error terms by exploiting the structure of the covariance matrix of the errors (see Zellner, 1962). In addition, in estimating the system by SURE, one is required to arbitrarily drop one of the demand equations in the system otherwise the covariance matrix of error terms is singular and therefore not invertible (see Takada, et al, 1995). We opt to drop the tobacco equation because: (i) we have not articulated an explicit theoretical framework for the demand for tobacco and more importantly (ii) our interest is in assessing the impact of tobacco expenditure on other commodities. To ensure that our estimated coefficients are invariant to the choice of equation that is dropped, we estimate the system using the *ireg3* command in STATA which provides maximum likelihood estimates.

5.0. Description of the data

The data for this paper comes from the 2006 round of the Living Conditions Monitoring Survey (LCMS) conducted by the Central Statistical Office (CSO) in Zambia. The survey was nationally representative and used a two-stage stratified cluster sample design whereby 988 clusters were selected in the first stage. The second stage saw the selection of 18,662 households from the 988 clusters distributed as 9,530 households in urban areas and 9,132 households in rural areas. The urban households were classified as low-cost, medium-cost or high-cost according to the local authority's classification of residential areas. In rural areas, households were classified as either small-scale, medium-scale, large-scale or non-agriculture households. The survey collects a rich set

of data on the living conditions of households in the areas of education, health, economic activities and employment, child nutrition, death in the households, income sources, income levels, food production, household consumption expenditure, access to clean water and sanitation among other variables[4].

The household expenditure section of the survey asks each household to recall and report on the total expenditure allocated to a particular commodity over a reference period. In most cases the reference period is the month prior to the survey but for some commodities, such as expenditure on health care or schooling, the recall period is a year. In such cases, the annual expenditure is converted to monthly expenditure by dividing it by 12. As stated in section 4, we focus on the following commodities: food, alcohol, healthcare, schooling, water, housing, electricity, alternative energy sources, transportation, transport equipment maintenance, telephone usage, entertainment, house care, personal care and a final category we call "other". Below we briefly discuss the characteristics of the sample.

10% of the households in our sample reported spending some amount on tobacco in the previous month. In urban areas the proportion was 7% and 12% in rural areas. On average, a tobacco consuming household in urban areas spent USD5 on tobacco in the month prior to the survey. This represented an expenditure share of 3.8% of total household expenditure. For rural households, the comparable amount was USD1.60 which represents a share of 5%. These shares fall in the ranges found in other studies (John, 2008). Further, the expenditure shares dedicated to tobacco decrease as one moves to the right of the total expenditure distribution. For instance, urban households in the bottom 20% allocated 5% of their total expenditure to tobacco whereas households in the top 20% only allocated 2%. For rural households the comparable shares were 8% for households in the bottom 20% and 3% for households in the top 20% (see Table 1). This suggests that crowding-out is likely to be severe among poorer households, a hypothesis that we formally test below. The other interesting summary statistic is total household expenditure. In urban areas, households on average spent a total of USD217 in the month prior to the survey and in rural areas the comparable amount was USD45. This also suggests that crowding-out is likely to be more severe in rural areas. Other summary statistics of interest to this paper are contained in Table 2.

6.0. Empirical Results

We present the results of our empirical analysis in two parts: part one conducts difference of means tests for expenditure shares between smoking and non-smoking households. We repeat this process for the different types of foods that comprise a household's food budget. In part two we present the results of the econometric implementation of equation (4). Throughout the results are discussed separately for urban and rural households.

6.1. Expenditure shares

Table 3 contains our motivation for investigating whether expenditure on tobacco leads to crowding-out. In the table, for 11 commodities out of a total of 18, non-smoking households in urban areas allocate a greater share of expenditure to those commodities than smoking households. For rural

[4] Additional information on the Survey can be accessed here: http://catalog.ihsn.org/index.php/catalog/2258

areas the comparable number is 13. It is possible, however, that these expenditure share differences are driven by random chance. That is, repeating the exercise on a different data set might result in different results. In what follows, therefore, we only focus on those cases where the share differences are statistically significant and therefore less likely due to chance.

Food

Non-smoking urban households spend 47% of total monthly expenditure on food whereas smoking households spend 44.6%. The difference in food shares is statistically significant at the 1% significance level. The same is true in rural areas where non-smoking households allocate 54.7% of monthly expenditure on food whereas smoking households allocate 48.7%. This difference is also statistically significant at 1%.

Since food is a broad category encompassing many types of sub-commodities, it might be useful to break it down into its constituent parts to understand the sources of the overall food differences noted in Table 3. Table 4, therefore, looks at the expenditure shares allocated to different types of food across smoking and non-smoking households. In the table, smoking households allocate a significantly smaller share than non-smoking households for six food types in urban areas: vegetables, vegetable oil, fruit, dairy, bread and sugar. In rural areas, smoking households allocate significantly less than non-smoking households for the following food types: rice, fish, vegetables, sugar and a broad food category "other" which includes items such as potatoes (Irish and sweet), groundnuts and honey.

The above suggests that smoking households allocate a smaller share of their expenditure to those food types that are likely to have an impact on the nutritional status of children in the household. This pattern has been found elsewhere (John, 2008; Block and Webb, 2009). Interestingly, in urban areas, smoking households allocate a larger share to the purchase of the staple food (maize) than non-smoking households, although the difference is only statistically significant at 10%. Block and Webb (2009) find a similar result in Indonesia and view this as evidence of substituting quantity for quality: a kilogram of the staple costs less than a kilogram of beef, say, and results in more calories.

Alcohol

Table 3 shows that smoking households in both rural and urban areas allocate more of their monthly expenditure towards alcohol than non-smoking households. In both cases the differences are quite large: 1.5% versus 6.5% in urban areas and 1.3% versus 7.4% in rural areas. In both cases, the differences are statistically significant at 1%. This result is not surprising as many studies show that tobacco and alcohol are complementary goods (Guindon, et al, 2011).

School

Smoking households in both urban and rural areas allocate less of their monthly expenditure towards meeting school expenses which includes buying books, school uniforms and making contributions to the Parents and Teachers Association (PTA)[5]. Urban non-smoking households spend

[5] In Zambia, public education (primary and secondary) is provided free of charge by the state. As can be expected, the grant that public schools receive from the state is not sufficient to cover all running costs. And

5.6% of monthly expenditure on school related expenses whereas smoking households spend 3.7%. In rural areas the differences are 5.5% versus 4% for non-smoking and smoking households respectively. The differences are significant at 1%. This finding is related to the earlier discussion on food. Spending on school primarily "benefits" children and since children do not have control on monthly budget allocations, their needs are relegated to those of adults (Deaton et al, 1989; Deaton, 1997).

Clothing

Non-smoking households in rural areas allocate 12.5% of their monthly budget towards clothing expenses whereas smoking households allocate 11.3%. The difference between the two is statistically significant at the 1% level. In urban areas, the reverse is the case where smoking households allocate a larger share. This difference is however not statistically significant.

Housing

Smoking households under allocate their share of expenditure to housing when compared with non-smoking households in urban areas. The former spend 5.1% of monthly expenditure on housing versus 7% for the latter. The difference is significant at 1%. In rural areas, there is no discernible difference between smoking and non-smoking households. This is probably due to the fact that a large majority of people in rural area live in owner-occupied and owner-built houses.

Electricity and alternative energy sources

In urban areas, non-smoking households allocate a larger share of expenditure towards electricity than smoking households and the difference is statistically significant at 1%. Interestingly, smoking households in urban areas allocate a larger share of monthly expenditure on alternative energy sources such as kerosene, firewood and charcoal than smoking households (although the difference is only statistically significant at 10%). It is quite likely that this pattern exhibits some trade-off behaviour where resource constrained smoking households consider electricity and electricity alternatives as substitutes. There are neither discernible differences nor similar patterns in rural areas.

Water

In both urban and rural areas, smoking households allocate a smaller share of expenditure towards water than non-smoking households. The mean shares are 1.7% against 1.2% for non-smoking and smoking households respectively in urban areas. In rural areas the shares are 0.1% and 0.09% for non-smoking and smoking households respectively. The differences are statistically significant in both cases

Daily Transport

In rural areas, non-smoking households allocate a larger share of expenditure on daily transportation (to and from work and to and from school) than smoking households. The difference is statistically

---

Parents and Teachers Associations (PTAs) were formed to mobilise private resources from parents to supplement the grants coming from the government. Donations to the PTAs are voluntary.

significant at 1%. Even though we found qualitatively similar results in urban areas, the difference was not statistically significant.

Telephone

Telephone expenditure is a relatively important component of monthly household expenditure for urban households than rural households. Telephone expenditure mostly consists of a household's expenditure on "top-up" vouchers for their mobile phones. According to Table 3, non-smoking households in urban areas allocate 3.3% of monthly expenditure towards their telephones compared with 2.6% for smoking households. The difference is significant at 10%. There are no discernible patterns in mean shares in rural areas.

Personal Care

In both urban and rural areas, non-smoking households allocate a larger share towards personal care expenditure (toiletries, cosmetics and hairdressing) than smoking households. The differences in both cases are statistically significant at 1%.

6.2. Econometric results

In section 6.1 we compared the mean expenditure shares dedicated to different commodities by smoking and non-smoking households. We focussed on those instances where the differences were statistically significant and not due to random chance. This section tries to investigate whether the differences detailed in 6.1 can be given a causal interpretation. That is, are the expenditure share differences between smoking and non-smoking households caused by tobacco? To answer this question we proceed in a series of steps. We begin by investigating the "strength" of our instrumental variables, and then test for corner solutions and eventually discuss the demand system results.

6.2.1. Testing for instrument strength

Table 5 and Table 6 show the "first-stage" results of regressing monthly tobacco expenditure on the proportion of adult males in adult house size in rural and urban areas. Recall that adult male proportion is our choice of instrument for tobacco expenditure. We also include in these regressions the vector $\boldsymbol{a}$ of household characteristics. In both tables, the proportion of adult males in adult house size has a significant and positive causal effect on monthly tobacco expenditure after controlling for household characteristics. That is, an increase in this proportion results in a significant increase in a household's expenditure on tobacco. The coefficients on the instrument in both regressions have F-values equal to 13.7 (t-value = 3.7) in rural areas and 10.2 (t-value = 3.2) in urban areas. Both satisfy the standard rule of thumb that an instrument is considered "strong" if its associated F statistic is at least equal to 10 (Stock, Wright and Yogo, 2002).

Table 7 and Table 8 repeat the above exercise and test the strength of the log of total expenditure as an instrument for the log of residual expenditure M, which is defined as total expenditure minus expenditure on tobacco. In both urban and rural areas the log of total expenditure has a positive and significant effect on the log of M. In both cases the associated F statistics on the coefficients of the instruments are significantly greater than 10.

6.2.2. Testing for corner solutions

Table 9 reports the results of the test for corner solutions described in section 4. Recall that we can reject the hypothesis of corner solutions if the variables associated with *d*, the indicator variable, are jointly significant. In our case, we reject the hypothesis of corner solutions for 12 goods and services in urban areas and for 7 goods and services in rural areas. We therefore conclude that the zeroes that we observe in our dataset are likely driven by abstention implying that smoking and non-smoking households have different preferences in the sense made precise in section 3[6]. Augmenting our Engel curves with *d*, therefore, serves the simultaneous purpose of testing for corner solutions as well as controlling for preference heterogeneity once it is established.

6.2.3. Demand system results.

Tables 10, 11, 12 and 13 present the results of estimating equation (4) by Three Stage Least Squares on different subsets of the data.

Table 10 shows the results for urban areas. Recall that a negative coefficient on the $Tobacco$ variable implies crowding-out. Negative coefficients on the $Tobacco$ variable are observed in 14 of the demand equations although the coefficients are only statistically significant in 3 of the equations. That is, a household that spends on tobacco reduces its share of total expenditure dedicated to food, other transportation and transport equipment maintenance in urban areas. Crowding-in is observed in 4 of the demand equations with only 2 of the cases being statistically significant. That is, a household that spends on tobacco tends to increase its share of total expenditure dedicated to alcohol and daily transportation in urban areas. One implication of this result is that the crowding-out effect of tobacco is likely to be amplified through its positive effect on alcohol.

In Table 12, we estimate the demand system on high and low-income subsets of the urban sample. High-income is defined as the richest 50% and low-income is defined as the poorest 50% in terms of total monthly expenditure. It would be ideal to split-up the urban dataset into ever finer income subgroups such as quintiles. The fact that tobacco is only consumed in a handful of households cautions against doing so: some subgroups are likely to have a few number of households consuming tobacco and hence running the risk of estimates being influenced by small sample sizes. Broader subgroups, therefore, insure that we have a non-negligible number of households reporting positive tobacco expenditure in our sample. The results from Table 12 show some interesting patterns: food is the only commodity that is crowded-out in the bottom 50% of households. The only commodity crowded-out in the top 50% of households is transport equipment maintenance. Interestingly, alcohol is crowded-in in the richer households and crowded-out in the poorer households. That is, alcohol seems to "compete" with tobacco for resources among poorer households. Only the crowding-in effect is, however, statistically significant.

[6] In rejecting the hypothesis of corner solutions, it is not necessary that the chi-square statistics be large in all the demand equations in the system. It is enough that they be large in a handful of the demand equations. In the case of Vermeulen (2003), the hypothesis of corner solutions was rejected in 7 out 11 cases. In Pu et al (2008), the hypothesis was rejected in 15 out of 27 cases. In the case of John (2008), the hypothesis of corner solutions was rejected in all of the 9 demand equations.

Table 11 shows results of the demand system estimates on the entire rural sample. Tobacco expenditure is associated with negative coefficients in 15 of the 18 demand equations. In only 5 cases are the negative coefficients statistically significant. That is, tobacco expenditure results in a decrease in expenditure shares dedicated to schooling, electricity, daily transportation, other transportation and transport equipment maintenance. Crowding-in is observed in 3 cases and is significant for alcohol and telephones. Rural households that spend on tobacco also increase their share of monthly expenditure dedicated to alcohol and telephone usage. The fact that tobacco crowds-out more commodities in rural areas than in urban areas partly confirms our hypothesis that crowding-out is likely to be more prevalent among income-constrained households (rural households are generally poorer than urban households, with urban households spending on average 4 times more per month than rural households. See Table 2). Interestingly, tobacco does not crowd-out food in rural areas (the tobacco coefficient however has a negative sign). One possible explanation for this is that a lot of the transactions in rural areas are either non-market transactions or barter transactions and do not directly compete with transactions involving tobacco which is largely purchased using money. It is not unusual, in rural areas, for one to grow their own food or for one to exchange a bag of maize for a number chickens, for instance[7]. The 5 commodities that are crowded-out in rural areas are likely to be purchased using money.

Table 13 looks at demand system estimates for upper-income and lower-income households in rural areas. Tobacco crowds-out daily transportation, other transportation and transport equipment maintenance among the rural rich. Among the rural poor, the only commodity that is crowded-out is entertainment. Crowding-in is only present for alcohol and telephone among rich households. Just as in urban areas, the crowding-in effect of alcohol is not statistically significant among low-income households. In both upper-income and low-income households, the crowding-out effect of food is not statistically significant even though the magnitude of the coefficient on $Tobacco$ is bigger for low-income households.

7.0. Concluding remarks

The mortality and morbidity costs of tobacco use are widely documented. For instance, it is known that tobacco kills about 5 million people per annum (WHO, 2010) with a significant number dying from passive smoking. Not much is known, however, about the crowding-out effects of tobacco expenditure, if at all there are any such effects. That is, does tobacco expenditure displace expenditure on other commodities within households? A number of academic studies over the last decade have documented the crowding-out effects of tobacco expenditure in several parts of the world. Our paper is a contribution to this literature within the context of Zambia, a country where smoking prevalence has recently began to increase after declining for most of the 1990s. Specifically, our findings suggest that households that spend on tobacco reduce their shares of monthly expenditure on food, transportation and transport maintenance equipment in urban areas and on schooling, electricity, transportation, transport maintenance equipment and entertainment in rural areas. Further, tobacco expenditure tends to crowd-in the share dedicated to alcohol in both rural

[7] In instances where a commodity is obtained by barter exchange or is own-grown, the Central Statistical Office converts the quantities consumed into monetary units using estimated or actual market prices. See http://catalog.ihsn.org/index.php/catalog/2258 for additional information.

and urban areas suggesting that tobacco's crowding-out effects are likely amplified via their positive impact on alcohol spending. Our results also suggest that crowding-out is likely to be more severe in rural areas and the pattern of crowding-out and crowding-in effects have a socio-economic pattern in both rural and urban areas. For instance, food is not crowded-out among the urban rich while it is crowded-out among the urban poor. In addition, alcohol is only crowded-in among the rich in urban and rural areas whereas there seems to be no relationship between the two among the poor in both rural and urban areas. The latter suggests that tobacco and alcohol likely compete for resources among the poor in both rural and urban areas. Tobacco expenditure does not crowd-out food expenditure in rural areas a result that we suspect is due to the fact that food is largely home-grown in rural parts of Zambia.

Quintile	1	2	3	4	5	Total
Urban	4.78%	4.32%	4.21%	3.26%	2.37%	3.90%
Rural	8.47%	5.68%	4.87%	3.67%	2.85%	4.82%

Table 1: Tobacco expenditure shares across expenditure quintiles for smoking households only

Notes: Quintile 1 refers to the poorest 20% while quintile 5 refers to the richest 20%

Independent Variables	**Urban**	**Rural**
Monthly tobacco expenditure[8]	ZMK22,994 (USD5)	ZMK6,785 (USD1.60)
Monthly household expenditure	ZMK917,789 (USD217)	ZMK189,462 (USD45)
Proportion of adult males	0.48	0.46
Household size	5.3	5.2
Proportion of adults in household (household structure)	0.59	0.54
Age of household head	40	43
Average age of adults in household	33	36
Average age of children in household	8.9	8.2
Years of schooling of household head	10	7
Years of schooling for most educated h/hold member	11	8
Total household employment	1.4	1.6
Proportion of household heads with wage(regular) income	0.50	0.11
Proportion of low-cost households (small scale agric. rural)	0.67	0.75
Proportion of medium-cost households (medium scale agric. rural)	0.18	0.11
Proportion of high-cost households (large-scale agric. rural)	0.14	0.04
Non-agrciulture household	Not applicable	0.14
Dependent Variables	**Urban**	**Rural**
Food Share	0.47	0.54
Alcohol Share	0.02	0.02
Health Share	0.008	0.01
School Share	0.06	0.05
Clothing Share	0.06	0.12
Housing Share	0.06	0.002
Water Share	0.02	0.001
Electricity Share	0.02	0.002
Alternative Energy Share	0.04	0.05
Daily Transport Share	0.03	0.008
Other Transport Share	0.02	0.02
Equipment Maintenance Share	0.004	0.01
Entertainment Share	0.009	0.001
Telephone Share	0.04	0.007
Remittance Share	0.009	0.005
House Care Share	0.01	0.003
Personal Care Share	0.10	0.12
Other Share	0.02	0.02

Table 2: Summary statistics for independent and dependent variables between urban and rural households

[8] Zambian Kwacha converted to United States Dollar using the end-of-year exchange rate in 2006 obtained from www.oanda.com.

	Urban				Rural			
Expenditure Share on:	**n**	**Non-Smoking**	**Smoking**	**Difference**	**n**	**Non-Smoking**	**Smoking**	**Difference**
Food	9,514	0.47	0.446	**0.024*** **	9,105	0.547	0.487	**0.060*** **
Alcohol	9,512	0.015	0.065	**-0.05*** **	9,095	0.013	0.074	**-0.061*** **
Health	9,514	0.008	0.008	**0**	9,105	9,105	0.013	**0**
School Expenditure	9,514	0.056	0.037	**0.019*** **	9,105	0.055	0.04	**0.015*** **
Clothing	9,514	0.057	0.059	**-0.002**	9,105	0.125	0.113	**0.012*** **
Housing	9,510	0.07	0.051	**0.019*** **	9,092	0.0019	0.0019	**0**
Water	9,512	0.017	0.012	**0.005*** **	9,093	0.0014	0.0009	**0.0005** **
Electricity	9,514	0.023	0.013	**0.01*** **	9,099	0.0011	0.0013	**-0.0002**
Alternative Energy	9,514	0.046	0.051	**-0.005***	9,105	0.048	0.045	**0.003**
Daily Transport	9,514	0.032	0.031	**0.001**	9,105	0.0052	0.003	**0.002*** **
Other Transport	9,501	0.021	0.019	**0.002**	9,092	0.019	0.017	**0.002**
Equipt. Maintenance	9,514	0.003	0.004	**-0.001**	9,105	0.01	0.01	**0**
Entertainment	9,514	0.007	0.009	**-0.002**	9,105	0.002	0.0023	**-0.0003**
Telephone	9,507	0.033	0.026	**0.007***	9,095	0.0063	0.0044	**0.0019**
Remittance	9,510	0.008	0.007	**0.001**	9,095	0.005	0.005	**0**
House Care	9,514	0.014	0.013	**0.001**	9,105	0.002	0.003	**0.001**
Personal Care	9,514	0.098	0.084	**0.014*** **	9,105	0.123	0.11	**0.013*** **
Other	9,514	0.019	0.026	**-0.007*** **	9,105	0.023	0.021	**0.002**

Table 3: Expenditure Shares between smoking and non-smoking households, urban and rural areas

Notes: *, ** and *** refer to whether "mean expenditure share differences" are statistically significant at 10%, 5% and 1% levels respectively. A negative difference means Smoking Households allocate, on average, a larger share to that commodity than Non-Smoking Households.

	Urban				Rural			
Expenditure Share on:	**n**	**Non-Smoking**	**Smoking**	**Difference**	**n**	**Non-Smoking**	**Smoking**	**Difference**
Staple	9,514	0.04	0.047	**-0.007***	9,105	0.083	0.077	**0.006**
Rice	9,513	0.017	0.016	**0.001**	9,097	0.016	0.013	**0.003***
Meat	9,514	0.041	0.04	**0.001**	9,105	0.043	0.04	**0.003**
Poultry	9,514	0.042	0.038	**0.004**	9,105	0.025	0.022	**0.003**
Fish	9,514	0.056	0.052	**0.004**	9,105	0.099	0.09	**0.009****
Vegetables	9,514	0.099	0.093	**0.006***	9,105	0.078	0.069	**0.009*****
Vegetable Oil	9,514	0.039	0.036	**0.003****	9,104	0.061	0.05	**0.011*****
Fruit	9,513	0.005	0.004	**0.001****	9,100	0.002	0.002	**0**
Dairy	9,514	0.019	0.014	**0.005*****	9,105	0.007	0.006	**0.001**
Bread	9,512	0.043	0.037	**0.006*****	9,096	0.021	0.021	**0**
Sugar	9,509	0.031	0.028	**0.003****	9,102	0.053	0.044	**0.009*****
Non Alcoholic Bev	9,514	0.013	0.013	**0**	9,105	0.005	0.005	**0**
Other Foods	9,514	0.026	0.028	**-0.002**	9,105	0.055	0.048	**0.007*****

Table 4: Expenditure shares on food types between smoking and non-smoking households, urban and rural areas

Notes: *, ** and *** refer to whether "mean expenditure share differences" are statistically significant at 10%, 5% and 1% levels respectively. A negative difference means smoking Households allocate a larger share to a particular food type than non-Smoking Households. The shares in each column should add-up to the total share of Food in Table 3.

Independent Variables:	**Dependent Variable: Tobacco Expenditure**
Proportion of Adult Males	809.2732***
	(218.9579)
Log of Household Size	506.8729**
	(222.3545)
Log of Years of Schooling for H/hold Head	9.2551
	(138.6426)
Log of most educated household member	-4.6251
	(189.6907)
Total household employment	-66.43427
	(55.5247)
Dummy for whether household head has wage income	642.308**
	(297.0095)
Household Structure	1425.297***
	(532.4491)
Log age of household head	-747.0083
	(475.1585)
log of average adult age	1159.194**
	(523.2616)
log of average child age	-147.2257*
	(85.26003)
Dummy 1 for household type	404.4444
	(326.8703)
Dummy 2 for household type	873.1506
	(1269.976)
Dummy 3 for household type	89.62362
	(237.0464)
Constant term	-2101.299**
	(943.8988)
n = 6,444 , Model F = 2.86, p-value = 0.00	

Table 5: Regression Output for Tobacco Expenditure on proportion of adult males and other household characteristics, rural areas.

Notes: *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Robust standard errors in parentheses.

Independent Variables:	**Dependent Variable: Tobacco Expenditure**
Proportion of Adult Males	2153.553***
	(678.2834)
Log of Household Size	432.5268
	(529.8626)
Log of Years of Schooling for H/hold Head	480.6315
	(491.0744)
Log of most educated household member	85.7899
	(520.7656)
Total household employment	456.8468*
	(278.3551)
Dummy for whether household head has wage income	206.7248
	(307.754)
Household Structure	1214.468
	(1422.982)
Log age of household head	-727.0766
	(1045.301)
log of average adult age	3023.481**
	(1363.084)
log of average child age	46.3626
	(157.505)
Dummy 1 for household type	-439.1221
	(476.9734)
Dummy 2 for household type	488.6073
	(596.7942)
Constant term	-10 621.21***
	(3309.733)
n = 7,298 , Model F = 3.59, p-value = 0.00	

Table 6: Regression output for tobacco expenditure on proportion of adult males and other household characteristics, urban areas.

Notes: *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Robust standard errors in parentheses.

Independent Variables:	**Dependent Variable: Log M**
Log of Total Expenditure	1.0003***
	(0.0002)
Log of Household Size	-0.0003
	(0.0006)
Log of Years of Schooling for H/hold Head	0.001
	(0.001)
Log of most educated household member	0.0007
	(0.001)
Total household employment	-0.0002
	(0.0002)
Dummy for whether household head has wage income	0.0001
	(0.0004)
Household Structure	-0.0016
	(0.0014)
Log age of household head	0.002*
	(0.0014)
log of average adult age	-0.004***
	(0.002)
log of average child age	0.0001
	(0.0004)
Dummy 1 for household type	0.001**
	(0.0004)
Dummy 2 for household type	0.0003
	(0.0005)
Constant term	-0.003
	(0.004)
n = 7,296 , Model F = 3.8, p-value = 0.00	

Table 7: Regression output for Log M on Log of Total Expenditure and other household characteristics, urban areas

Notes: *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Robust standard errors in parentheses.

Independent Variables:	**Dependent Variable: Log M**
Log of Total Expenditure	1.0000***
	(0.0003)
Log of Household Size	-0.002***
	(0.0008)
Log of Years of Schooling for H/hold Head	0.0001
	(0.0008)
Log of most educated household member	0.004***
	(0.001)
Total household employment	0.0002
	(0.0002)
Dummy for whether household head has wage income	0.0007
	(0.0007)
Household Structure	-0.007***
	(0.002)
Log age of household head	0.002
	(0.002)
log of average adult age	-0.006***
	(0.002)
log of average child age	0.001
	(0.0005)
Dummy 1 for household type	0.0008
	(0.001)
Dummy 2 for household type	0.005***
	(0.001)
Dummy 3 for household type	0.001*
	(0.0007)
Constant term	0.006
	(0.005)
n = 6,437 , Model F = 3.6, p-value = 0.00	

Table 8: Regression output for Log M on Log of Total Expenditure and other household characteristics, rural areas

Notes: *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Robust standard errors in parentheses.

Commodity	Chi Square - Urban	Chi Square - Rural
Food	28.59***	54.96***
Alcohol	320.89***	485.96***
Health	0.20	1.16
School Expenditure	13.44***	20.04***
Clothing	8.71**	2.94
Housing	15.15***	12.21***
Water	1.73	2.91
Electricity	10.43**	1.42
Alternative Energy	19.89***	5.72
Daily Transport	7.39*	9.7**
Other Transport	113.23***	3.44
Equipment Maintenance	13.31***	7.03*
Entertainment	1.02	1.01
Telephone	0.89	18.77***
Remittance	3.07	6.02
House Care	2.04	4.59
Personal Care	12.63***	4.68
Other	8.48**	0.57

Table 9: Tests for consumer separability and corner solutions for urban and rural households.

Notes: Values in each column are Chi-Square statistics from a Wald Test for the joint significance of the three parameters associated with the tobacco dummy *d* in equation (4). *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively

Regressors	Food	Alcohol	School	Health	Clothing	Hsing	Water	Electricity	Alt. Energy	Daily T/port	Other T/port	Equ. Maint.	Entertain.	T/phone	Hse. Care	Personal Care	Remittance	Other
Tobacco	-.0003**	.0001***	-.00009	-.000006	-.00004	-.00003	-.00006	-.00001	-.00003	.0001*	-.0002***	-.00005**	.00002	.00003	-.00003	-.00006	-.00003	-.00002
LnM	.4***	.013	-.125***	-.003	-.111***	.064***	.024***	.077***	.014	-.289***	.05***	-.059***	-.092***	.037***	-.017***	.04**	-.017***	-.005
$(LnM)^2$	-.018***	-.0003	.005***	.0001	.004***	-.002**	-.0009***	-.0028***	-.001***	.012***	-.002***	.002***	.004***	-.0008*	.0009***	-.002***	.0008***	-.00008
d	-.821	-.448**	-.803**	-.047	-.707***	.326	.024	.195	1.031***	-.720**	2.1***	.213***	-.102	-.077	-.169	-.102	-.174	.221
d LnM	.094	.078**	.121**	.007	.110***	-.04	-.004	-.031	-.153***	.110**	-.33***	-.03*	.015	.01	.025	.008	.027	-.033
d $(LnM)^2$	-.003	-.003***	-.005**	-.0003	-.004***	.001	.0002	.001	.006***	-.004**	.013***	.001*	-.0005	-.0003	-.0009	-.0001	-.001	.001
Hhold Size	.022***	-.002	.032***	.0007	.00004	-.032***	.00001	.004**	.002	-.005*	-.007***	.002*	-.005***	-.003	-.005***	.002	-.005***	-.00006
Sch Head	-.006	-.003	.011***	-.004***	.003	.005	.003	.005**	-.001***	-.003	.0009	.002*	-.001	.006**	-.001	-.002	-.0001	-.003**
Mst Educ.	-.021**	-.003	.01*	.004**	-.003	-.01	.005**	.014***	-.01***	-.003	-.006*	-.002	.002	.01***	.007***	.004	.004***	-.001
Total Emply	-.002	.002***	-.008***	.00005	.004***	-.003**	-.0006	-.003***	-.0002	.001	.0006	.0006*	.0007	.003***	.0008*	-.00005	.002***	.001**
Wage	-.011***	.002**	.002	-.002***	-.0009	.019***	-.002**	-.00005	-.003***	-.007***	-.005***	-.003***	.001	.004***	.00009	.009***	.001*	-.003***
H/hold str.	.021	.01**	-.009	.002	-.016***	-.013	.009***	.014***	-.006	.007	-.008	.001	-.004	.001	-.0003	-.0008	-.009***	.00005
Age head	-.044***	-.013***	.09***	-.002	-.007	-.036***	.013***	.016***	.003	-.007	.007	-.0005	.003	-.011**	-.004	.005	-.0004	-.004
Adult age	.044***	.008	-.079***	.009***	-.005	.003	.001	-.003	.011**	.006	.002	.001	.0002	.005	.006**	-.014*	.0004	.002
Child age	-.006**	-.001	.024***	-.0006	-.001	-.008***	.0004	.003***	-.003***	.0009	-.002*	.0001	.0002	.0006	.00008	-.004***	-.0008*	-.001*
Hhold 1	-.003	-.005***	.012***	-.0007	-.0006	.0008	.013***	.008***	-.007***	-.01***	-.00008	-.001**	-.001	-.0006	.00006	.003	-.003***	-.004***
Hhold 2	-.009*	-.005***	.008***	-.002**	-.004*	-.007**	.011***	.012***	-.002	-.015***	-.006***	.0008	.005***	.001	.011***	.007***	-.002**	-.002**
Cons	-1.57	-.07	.652***	-.004	.88***	-.307**	.22***	-.608***	.077	1.73***	-.366***	.358***	.569***	-.33***	.065*	-.073	.09**	.115**

Table 10: Demand system results conditioning on monthly tobacco expenditure, residual expenditure *M* and other regressors, urban areas.

Notes: Parameters (and standard errors) associated with the Tobacco variable are multiplied by 1000. Standard errors not reported because of space considerations. *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Variable names are abbreviated for space considerations. Refer to Table 3 and Table 5 for names in full. Sample size is 7,275.

Regressors	Food	Alcohol	School	Health	Clothing	Hsing	Water	Electr.	Alt. Energy	Daily T/port	Other T/port	Equ. Maint.	Entertain.	T/phone	Hse. Care	Personal Care	Remit	Other
Tobacco	-.0004	.0007***	-.00005*	-.000009	-.0003	-.00006	-.000006	-.0001**	-.00009	-.0006***	-.0007***	-.0002**	.00005	.0003***	-.000007	-.0005	-.000004	-.00008
LnM	.475***	.01	.06***	-.006	-.07***	-.01***	.003**	-.03***	.01*	-.15***	-.06***	-.02***	-.007***	-.07***	-.02***	.02	-.03***	.001
$(LnM)^2$	-.02***	-.0003	.003***	.0002	.002***	.0004***	.0001**	.001***	-.001**	.007***	.003***	.001***	.003***	.003***	.0008***	-.002**	.001***	-.00008
d	-0.85	-.15	-.302	-.15	-.005	-.12*	-.0007	-.02	.004	.62***	.5	.28**	-.03	-.06	.05	.34	-.14	-.1
***d* LnM**	.1	.04	.045	.02	-.002	.02*	.0005	.005	-.007	-.10***	-.08*	-.05**	0.005	.015	-.007	-.06	.03	.016
***d* $(LnM)^2$**	-.003	-.002	-.002	-.001	.0001	-.001**	-.00004	-.0002	.0006	.004***	.004*	.002**	-.0002	-.0008	.0003	.003	-.001	-.0006
Hhold Size	-.06***	-.002	.05***	.003	.02***	-.001	.0002	-.0006	-.00008	.003	-.01***	.009***	.0003	.0008	-.001	-.009**	-.002	.002
Sch Head	.0001	-.002	.01***	-.001	-.001	-.0004	.0002	.0009	.002	-.0006	-.004	.001	.0002	.0007	.001**	-.006	-.0009	0.002
Mst Educ.	-.02**	-.001	.001	-.003*	.01**	.001	-.0005	.0004	-.0003	-.004*	.0002	-.001	-.0005	.002	-.0008	.01**	.02*	.0002
Total Emply	.005*	.0005	-.01***	.0005	.0006	-.0002	-.0003***	-.0002	.0006	.0007	0.0001	-.0004	.00005	-.0006	.0001	.003**	.0005	.0002
Wage Dummy	.01	-.0006	-.0007	-.003*	-.001**	.008***	.0004	.002**	-.003	-.01***	-.01***	-.01***	.002***	.007***	.002***	.02***	.003**	-.003
H/hold str.	-.06***	.006	.03**	.004	.02**	.002	.002	.003	.003	.0006	-.01	.008**	-.0001	.007**	.003*	-.02	-.007**	.008
Age head	-.008	-.008	.11***	-.007*	-.05***	-.003*	-.0004	-.0004	-.007	-.006	.005	-.01***	-.0005	-.006*	-.0006	-.008	-.002	-.003
Adult age	.018	0.006	-.09***	.01**	.02	.0002	-.0001	.0007	.007	.009*	-.0005	.01***	.0007	.003	.0007	.007	.001	.003
Child age	-.004	-.003**	.03***	.0008	-.01***	.0009***	.0001	.0009**	-.003**	-.002**	-.001	-.0008	-.0003	.0004	.0007*	-.004*	-.001**	-.002**
Hhold 1	-.02***	-.007***	.01***	.002	-.01***	-.002***	-.001***	-.0009	-.004*	.01***	-.004	.004***	-.001**	.002	.002***	.0008	.0005	-.001
Hhold 2	-.014	-.019*	-.008	-.002	-.008	.005*	-.003*	-.002	-.0001	.06***	-.06***	.02***	.0007	.03***	.01***	-.013	.01***	-.007
Hhold 3	.05***	-.007***	-.007*	.004**	-.02***	.001**	.0002	-.0003	.0001	.002	.003	-.005***	-.0002	-.0004	.0004	-.017***	-.002*	-.003
Cons	-2***	-.05	0.18**	.05	.74***	.05***	.02*	.14***	-.008	.77***	.27***	.13***	.03***	-.35***	.09***	.1	.14***	.014

Table 11: Demand system results conditioning on monthly tobacco expenditure, residual expenditure *M* and other regressors, rural areas.

Notes: Parameters (and standard errors) associated with the Tobacco variable are multiplied by 1000. Standard errors not reported because of space considerations. *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Variable names are abbreviated for space considerations. Refer to Table 3 and Table 5 for names in full. Sample size is 6,404.

Upper-income households

Regressors	Food	Alcohol	School	Health	Cloth	Hsing	Water	Elect	Alt. Energy	Daily T/port	Other T/port	Equ. Maint.	Entertain.	T/phone	Hse. Care	Pers. Care	Remit	Other
Tobacco	-.0002	.0002***	-.0001	-.000003	-.00003	.00001	-.00005	-.00001	-.00002	.00008	-.0001	-.00006*	-.00003	.00002	-.00004	-.00002	-.00004	-.00003
LnM	.06	.05	.06	.004	.1*	.44***	.008	.1**	-.15***	-.69***	-.07	-.17***	-.09**	.088	.04	.17**	.06**	-.032
$(LnM)^2$	-.006	-.002	-.002	-.00009	-.003*	-.015***	-.0004	-.004***	.005***	.026***	.003	.006***	.004***	-.003	-.001	-.006***	-.002**	.001
d	3	-.93	-1.04	-.29	.7	2.8*	-.52	.012	.04	-7***	7.3***	-1.89***	-1.3*	-1.5	-.56	.78	.26	-.47
d LnM	-.42	.15	.15	.04	-.09	-.39*	.07	-.008	-.002	1.02***	-1.05***	.26***	.18*	.2	.079	-.12	-.033	.064
d $(LnM)^2$	.02	-.005	-.006	-.001	.003	.01	-.002	.0004	-.00005	-.035***	.038***	-.009***	-.006	-.007	-.003	.004	.001	-.002

Lower-income households

Regressors	Food	Alcohol	School	Health	Cloth	Hsing	Water	Elect	Alt. Energy	Daily T/port	Other T/port	Equ. Maint.	Entertain.	T/phone	Hse. Care	Pers. Care	Remit	Other
Tobacco	-.001**	-.0001	-.00005	-.00008	-.0002	-.0002	-.0002	-.00002	-.00009	.0008	-.0002	-.00003	-.00003	.00002	-.00004	-.0004	.00002	-.00004
LnM	.44***	-.001	-.17***	.0008	-.23***	-.06	.002	-.12***	.34***	-.068***	-.038	.002	-.004	-.16***	-.043***	.035	-.015	.08***
$(LnM)^2$	-.02***	.0002	-..007***	-.00002	.009***	.003	-.00003	.005***	-.014***	.003***	.002*	-.0001	.0002	.007***	.002***	-.002	.0007*	-.004***
d	-2.4	-1.7***	.43	-.32	-2.2**	-.4	--.33	-.23	4.9***	-.07	.179	1.94***	.25	-.17	-.094	-.25	-.13	1.24**
d LnM	.33	.29***	-.08	.053	.35**	.09	.05	.044	-.8***	.012	-.038	-.31***	-.045	.03	.016	.03	.022	-.2**
d $(LnM)^2$	.-.01	-.012***	.003	-.002	-.014	-.005	-.002	-.002	.03***	-.0006	.002	.012***	.002	-.001	-.0007	-.0007	-.001	.008**

Table 12: Demand system results conditioning on monthly tobacco expenditure, residual expenditure *M* and other regressors between expenditure groups, urban areas.

Notes: Upper panel shows the results from upper-income households and lower panel shows the results from lower-income households. Upper-income defined as the top 50% of households and lower-income defined as the bottom 50% of households. Parameters (and standard errors) associated with the Tobacco variable are multiplied by 1000. Standard errors not reported because of space considerations. *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Variable names are abbreviated for space considerations. Refer to Table 3 and Table 5 for names in full. Sample size for upper panel (upper-income households) is 3,847 and that of the lower-panel (lower-income households) is 3,428.

Upper-income households

Regressors	Food	Alcohol	School	Health	Cloth	Hsing	Water	Elect	Alt. Energy	Daily T/port	Other T/port	Equ. Maint.	Entertain.	T/phone	Hse. Care	Pers. Care	Remit	Other
Tobacco	-.0003	.0008***	-.0004	-.000004	-.0002	.00007	.000009	-.0001	-.00008	-.0007***	-.0007***	-.0003**	.00006	.0003**	-.00003	-.0004	-.0000007	-.00008
LnM	.56***	.04	.049	-.005	.056	.014	-.002	-.05***	.05	-.67***	.16**	-.22***	-.02*	.033	-.03**	.06	-.063***	.08***
$(LnM)^2$	-.03***	-.002	-.002	.0001	-.003	-.0004	.0001	.002***	-.002	.028***	-.005**	.009***	.0009**	.0006	-.001***	-.003	.003***	-.003***
d	.045	-1.6**	-.17	-.102	.038	-.019	-.08	-.39	.62	.72	1.8	.05	-.22	-.22	.45**	-.73	-.76*	-.11
d LnM	-.04	.27**	.024	.017	-.0006	.008	.013	.06	-.101	-.12	-.3*	-.02	.04	.041	-.07*	.103	.12**	.015
d $(LnM)^2$	.002	-.01**	-.0009	-.0007	-.0003	-.0005	-.0005	-.003	.004	.005	.012*	.0009	-.001	-.002	.003*	-.004	-.005**	-.0005

Lower-income households

Regressors	Food	Alcohol	School	Health	Cloth	Hsing	Water	Elect	Alt. Energy	Daily T/port	Other T/port	Equ. Maint.	Entertain.	T/phone	Hse. Care	Pers. Care	Remit	Other
Tobacco	-.003	.0002	-.002	-.0003	-.003	-.0000003	.000009	-.00000005	.0005	-.00007	-.001	-.0005	-.0002*	.00001	.0002	-.003	-.00002	.0002
LnM	.15	.006	.001	-.05**	.005	-.007**	.002	-.0002	.003	-.01*	-.033	.019	-.002	-.01*	-.004	.04	-.005	-.1***
$(LnM)^2$	-.005	-.0001	-.0004	.002**	-.002	.0004**	-.0008	.00001	-.0004	.0005**	.002*	-.0009	.00009	.0005*	.0002	-.002	.0003	.005***
d	.21	-1.16	.48	-.93	-.013	-.029	-.008	-.0004	1.74	.03	.458	.411	-.12	-.053	-.02	-1	-.053	-.23
d LnM	-.08	.242	-.11	.18	.007	.006	.002	.0001	-.34	-.006	-.09	-.08	.022	.01	-.021	.18	.01	.045
d $(LnM)^2$	.005	-.012	.006	-.008	-.0006	-.0004	-.0001	-.0007	.016	.0003	.005	.004	-.001	-.0006	.001	-.01	-.0001	-.002

Table 13: Demand system results conditioning on monthly tobacco expenditure, residual expenditure *M* and other regressors between expenditure groups, rural areas.

Notes: Upper panel shows the results from upper-income households and lower panel shows the results from lower-income households. Upper-income defined as the top 50% of households and lower-income defined as the bottom 50% of households. Parameters (and standard errors) associated with the Tobacco variable are multiplied by 1000. Standard errors not reported because of space considerations. *, ** and *** refer to statistical significance at 10%, 5% and 1% respectively. Variable names are abbreviated for space considerations. Refer to Table 3 and Table 5 for names in full. Sample size for upper panel (upper-income households) is 3,618 and that of the lower-panel (lower-income households) is 2,786.

References

Banks, J, Blundell, R and Lewbel, A, 1997. "Quadratic Engel Curves and Consumer Demand". *The Review of Economics and Statistics*, 79(40): 527 – 539.

Block, S & Webb, P., 2009. "Up in smoke: Tobacco use, expenditure on food, and child malnutrition in Developing Countries". *Economic Development and Cultural Change*, 58(1): 1 – 23

Browning, M and Meghir, C, 1991. "The effects of male and female labour supply on commodity demands". *Econometrica*, 59(4): 925 – 951.

Busch, S.H., Jofre-Bonet, M, Falba, T.A. & Sindelar, J.L., 2004. "Burning a hole in the budget: Tobacco spending and its crowd-out of other goods". *Applied Health Economics and Health Policy*, 3(4): 263 – 272.

Chelwa, G, 2012. "The Tobacco Story in Zambia: A Demand, Supply and Tax Analysis". Unpublished manuscript. University of Cape Town.

Deaton, A & Muellbauer, J, 1980a. "An Almost Ideal Demand System". *The American Economic Review*, 70(3): 312 – 326.

Deaton, A, Ruiz-Castillo, J and Thomas, D, 1989. "The influence of Household Composition on Household Expenditure Patterns: Theory and Spanish Evidence". *Journal of Political Economy*, 97(1): 179 – 200

Deaton, A, 1997. *Analysis of household surveys*. Baltimore: Johns Hopkins University Press on behalf of the World Bank.

Efroymson, D, Ahmed, A, Townsend, J, Alam, S.M., Dey, A.R., Saha, R, Dhar, B, Sujon, A.I., Ahmed, K.U. & Rahman, O, 2001. "Hungry for tobacco: an analysis of the economic impact of tobacco consumption on the poor in Bangladesh". *Tobacco Control*, 10: 212 – 217.

ERC Group, 2010. *World Cigarettes 2010*.

Guindon, E.G., Nandi, A, Chaloupka, F and Jha, P, 2011. "Socioeconomic differences in the impact of smoking tobacco and alcohol prices on smoking in India". NBER Working Paper Series, WP 17580.

John, R.M, 2008. "Crowding out effect of tobacco expenditure and its implications on household resource allocation in India". *Social Science and Medicine*, 66(2008): 1356 – 1367

Keen, M, 1986. "Zero expenditures and the estimation of Engel curves". *Journal of Applied Econometrics*, 1(3): 277 – 286.

Oberg, M, Jaakola, M, Woodward, A, Peruga, A & Pruss-Ustun, A, 2010. "Worldwide burden of disease from exposure to second-hand smoke: a retrospective analysis of data from 192 countries". *The Lancet*, 377(9760): 139 – 146

Pollak, R.A, 1969. "Conditional demand functions and consumption theory". *Quarterly Journal of Economics*, 83(1): 60 – 68.

Pu, C, Lan, V, Chou, Y-J & Lan, C, 2008. “The crowding-out effects of tobacco and alcohol where expenditure shares are low: Analyzing expenditure data for Taiwan”. *Social Science and Medicine*, 9:1979-1989

Stock, J.H, Wright, J.H. & Yogo, M, 2002. “A survey of weak instruments and weak identification in Generalized Method of Moments”. *Journal of Business and Economics Statistics*. 20(4): 518-529

Takada, H, Ullah, A & Chen, Y-M, 1995. “Estimation of the seemingly unrelated regression model when the error covariance matrix is singular”. *Journal of Applied Statistics.* 22(4): 517-530.

Vermeulen, F, 2003. “Do smokers behave differently? A tale of zero expenditures and separability concepts”. *Economics Bulletin*, 4(6): 1 – 7.

World Health Organisation (WHO), 2010. *WHO Technical Manual on Tax Administration.* WHO Press, Geneva.

Zambia Demographic and Health Survey, 2002. *Country Report*. Available at http://www.measuredhs.com/ (accessed July 2012).

Zambia Demographic and Health survey, 2007. *Country Report*. Available at http://www.measuredhs.com/ (accessed July 2012).

Zellner, A, 1962. “An efficient method for estimating Seemingly Unrelated Regressions and tests for Aggregation Bias”. *Journal of the American Statistical Association,* 57(298): 348-368.