Università Degli Studi di Siena

Ph.D. in Economics, XXIII cycle

Doctoral Dissertation

A Microeconometric Investigation into Issues of Development and Poverty in Mozambique

Vincenzo Salvucci

Doctoral Defence Committee	Supervisor
Prof. Gianni Betti	Prof. Francesca Bettio
Prof. Stephan Klasen	Ph.D. Director
Prof. Mariacristina Rossi	Prof. Ugo Pagano

Acknowledgements

Vorrei ringraziare tutte le persone che mi sono state vicino in questi anni di dottorato. Ci sono stati tanti momenti difficili e ho pensato di cambiare strada così tante volte che non riesco a credere di essere arrivato alla fine.

Ringrazio la professoressa Francesca Bettio in primo luogo, perché senza il suo aiuto e il suo incoraggiamento non sarei andato in Africa e non mi sarei appassionato alla ricerca che ora sto portando avanti.

Grazie ai miei familiari, che hanno sopportato molti dei momenti bui di questa esperienza, standomi accanto, ma lasciandomi tutta la libertà che avrei potuto desiderare nello scegliere la direzione della mia vita.

Grazie ai miei amici di sempre, che tra un'esperienza e l'altra sono sempre pronti ad accogliermi e a trovare il tempo per divertirci insieme. Grazie ai colleghi di dottorato, agli amici del Collegio Santa Chiara, agli amici Erasmus, grazie alle suore Comboniane di Alua, grazie a Bruna Ingrao e Marilena Giannetti, grazie a Laura e Valentina, grazie a Eva, grazie a Vasco Nhabinde, Joe Hanlon, Maureen MacKintosh, Rebecca Hanlin, Gianni Betti, Pierpaolo Pierani.

Ma soprattutto grazie di cuore a Novi. Perché è soprattutto grazie a lei che questi sono stati anche anni bellissimi. Love you.

Grazie.

Contents

Introduction

The economic literature focusing on developing countries benefited in recent years from the increasing availability of richer and more detailed survey data. This brought about a greater utilisation of microeconometric techniques in this field, and allowed researchers to investigate more deeply the causes and consequences of poverty and development at household level.

This dissertation attempts to make a contribution to the microeconometric literature on development economics for Mozambique by shedding new light on the multidimensional nature of poverty in the country as well as on households' decisions in matters of agriculture and health.

The dissertation includes three stand-alone essays which, however, share the focus on developmental issues, the microeconometric approach, and the use of Mozambican data sources.

The first essay asks why only a small part of the country's staple production is exchanged on the agricultural market despite the high number of producers. The second essay delves into poverty measurement in the specific attempt to identify which items should be included in the measurement in order to get a more complete mapping of deprivation. Within a more general analysis of the determinants of health care seeking in the Mozambique, the last essay specifically investigates whether implicit rationing in health care based on social status is in place in the country.

In the first essay, *"Staple Crop Marketing Decisions of Rural Households in Mozambique"*, participation of small-holders in the agricultural market and quantity of staples sold are investigated using selection and hurdle econometric models. The paper presents estimates at national and regional level using the agricultural survey *Trabalho do Inquérito Agricola* (TIA) 2005. Also, the impact of price, household income and other assets is estimated at the 10%, 25%, 50%, 75%, and 90% of the marketed staple distribution using quantile analysis. Finally, the relation between marketed surplus and poverty is investigated on the basis of household income and other assets. The paper offers three main results which bring some novel insights to the existing literature. First, income, assets, technology and information are found to play a significant role in determining the marketing decisions of staple food commodities; second, significant differences in these decisions are uncovered at regional level and between larger and smaller sellers; thirdly the price of crops is found to weakly influence the quantity sold at low quantiles of the distribution of staple crops sold, while an increase in household income and assets has a highly non-linear effect along this distribution.

The second essay, *"Multidimensional and Fuzzy Measures of Poverty and Inequality at National and Regional Level in Mozambique"* provides a step-by-step account of how fuzzy measures of non-monetary deprivation as well as monetary poverty may be constructed based on survey data such as those from the Mozambican Household Budget Survey 2008-09 (*Inquérito aos Agregados Familares sobre Orçamento Familiar 2008-09*, IOF08). Estimates are provided at national level and broken down at provincial and urban/rural level with standard errors computed using a recent methodology based on Jack-knife Repeated Replications (JRR). Monetary and non-monetary levels of deprivation are found to have very different distribution patterns, especially when analysed at sub-national level and by area of residence.

The third essay, *"Health Provider Choice and Implicit Rationing in Health Care: Evidence from Mozambique"* first conducts an analysis of health provider choice using Mozambican Household Survey on Living Conditions (*Inquérito aos Agregados Familares sobre Orçamento Familiar* (IAF) 2002-03 and 2008-09). In contrast to earlier results reported in the literature income appears to have played an important role throughout the 00s. Specific analysis of how socioeconomic disparities can give rise to implicit rationing in health care is carried out next by developing a theoretical model which is then tested on the latest the IAF survey. "Implicit rationing" means here that some people do not seek care because they anticipate they will not be granted access to quality health care or treatment. Some evidence that such behaviour is in place in Mozambique is found for 2008-09.

Staple Crop Marketing Decisions of Rural Households in Mozambique

Vincenzo Salvucci[a]

[a] Dipartimento di Economia Politica e Statistica - Facoltà di Economia "Richard M. Goodwin"
Università degli Studi di Siena - Piazza S. Francesco, 7 - 53100 Siena, Italy

salvucci3@unisi.it

Abstract

The paper analyses staple crop marketing decisions of Mozambican rural households with regard to market participation and quantities sold. The estimation is conducted at a national and regional level, using selection and hurdle models. Then analysed is whether the impact of price, household income and other assets differs for larger and smaller sellers. The estimation is carried out at 10%, 25%, 50%, 75%, and 90% of the marketed staple distribution using quantile analysis. Finally, a series of qualitative analyses on the relations between marketed surplus and poverty are conducted on the basis of household income and other assets. The dataset used (*Trabalho do Inquerito Agricola*, TIA 2005) refers to the 2004/2005 agricultural season, and is representative of all the provinces and agro-ecological zones of Mozambique. The findings are that, first, income, assets, technology, and information may play a significant role in determining the marketing decisions of staple food commodities in Mozambique; second, significant differences exist regarding these decisions at a regional level and between larger and smaller sellers; and third, market participation and marketed quantities are strongly associated with higher income and greater asset endowment.

Keywords: agriculture, market participation, Mozambique.
JEL classification: Q13, Q18, O13

Introduction

The vast majority of households in Mozambique live in rural areas (64%), and agriculture represents the main economic activity, accounting for more than 27% of GDP (World Bank, 2008). With regard to the poorest 80% of rural households, income from farm resources accounts for an average of 66.7% to 80.7% of their total income (Mather et al., 2008).

Notwithstanding the high level of market liberalisation in Mozambique, agricultural products are rarely exchanged or sold on the market (Boughton et al., 2007), and most rural households only produce goods for their own subsistence. Moreover, agriculture is mainly rain-fed, which increases vulnerability to climate shocks.

This paper analyses the staple crop marketing decisions made by rural households regarding their participation and the quantities sold, using selection and hurdle models. The dataset used (*Trabalho do Inquerito Agricola 2005*, henceforth TIA 2005) refers to the 2004/2005 agricultural season, and is representative of all the provinces and agro-ecological zones of Mozambique (Ministério da Agricultura, 2005).

An understanding of what influences staple crop marketing decisions is crucial in developing countries because agricultural market participation is acknowledged to be a poverty-reduction strategy and a productivity-enhancing mechanism for rural households (Barrett, 2008). In addition, staple crop markets usually exhibit distinctive characteristics, such as a variable responsiveness to price changes, low entry barriers, high transaction costs, and a lack of regulated contract schemes (Boughton et al., 2007).

Price-based interventions or imposed market liberalisations alone have rarely brought about large-scale shifts from subsistence to market agriculture (Barrett, 2008). There is, however, an ongoing debate on the relative importance of providing better infrastructures, governmental support, management skills, credit, and technology, especially in a country like Mozambique (Bingen et al., 2003; Tostão and Brorsen, 2005; Tschirley et al., 2006; Cirera and Arndt, 2008; Markelova et al., 2009; Bachke, 2010; Cunguara and Darnhofer, 2011). In recent years, the Mozambican government and a number of donor partners have committed to increasing agricultural productivity in the country, but with modest results so far (Jayne et al., 1999; Tschirley et al., 2006).

This study attempts to provide new insights into what influences staple crop marketing decisions in Mozambique. In line with a significant part of the literature on agricultural markets in developing countries, the decision to sell and the determination of the quantity to be sold are treated as separate but interrelated processes (Goetz, 1992; Key et al., 2000; Heltberg and Tarp, 2002; Holloway et al., 2005; Bellemare and Barrett, 2006; Boughton et al., 2007). Therefore analysed are both the probability of participation in the market and the determinants of the amount of marketed surplus for staple crops in Mozambique. Since the non-negligible differences that exist among the northern, central, and southern regions of the country are considered to be crucial, the estimation is conducted at a national level, and then disaggregated at a regional level.

Also analysed is whether the determinants of marketing decisions - or the effects of these determinants - vary between larger and smaller sellers. Estimation is therefore also made of the impact of price, household income, and other assets at 10%, 25%, 50%, 75%, and 90% of the marketed staple distribution, using quantile analysis. It is shown that the effect of these determinants does indeed vary if the household sells small or large quantities of staple crops. This evidences that agricultural policies need to be well-designed, because different outcomes may arise depending on the characteristics of the targeted population.

Finally, since the levels of income and possession of household and farm assets are found to affect market participation and the quantities of staple crops sold to a considerable degree, close examination is made of the role of these variables, and a series of qualitative analyses are conducted on the relation among market participation, marketed surplus, and poverty.

The article is organised as follows: Section 1 presents some of the literature on market participation, focusing on staple markets and Mozambique. In Section 2 and Section 3, the conceptual and empirical frameworks used for studying market participation and quantities sold are described. Section 4 presents the data and descriptive statistics for the variables considered in the subsequent analyses. Section 5 sets out the econometric results and discusses them. Section 6 presents a number of qualitative analyses of the link between marketing decisions and income/other assets. Section 7 draws policy implications and concludes.

1. Market participation in Mozambique

The market participation of rural households in developing countries is a complex issue because households are simultaneously producers and consumers of the same goods. After the early studies by Singh et al. (1986), De Janvry et al. (1991), and Fafchamps (1992), rural household models were developed so as to take account of the dual role of rural households - as producers and consumers - together with the pervasive market imperfections that exist in developing countries. At the same time, more sophisticated empirical approaches attempted to provide better estimates of marketed surpluses by considering sample selection, transaction costs and market incompleteness.

Two benchmark studies influenced the subsequent literature on staple crop market participation: Goetz (1992), which analysed grain market participation in Senegal, and Key et al. (2000), which studied maize market participation in Mexico. A growing body of studies have also examined the commercialisation of livestock, animal products, and cash crops (Holloway et al., 2005; Bellemare and Barrett, 2006; Benfica et al., 2006)[1].

Certain common characteristics seem to emerge from the empirical studies on market participation of rural households in Africa (Barrett, 2008). Despite the fact that many rural households produce staple food crops, few of them participate in the market. Moreover, market participation is associated with greater asset endowments, facilitated access to market, and improved agro-ecological characteristics, although transaction costs are substantial, and under many circumstances greatly influence household behaviours relative to commercialisation and production.

Empirical evidence on market participation in Mozambique is meagre, but what there is confirms the above-mentioned characteristics: Heltberg and Tarp (2002) find that risk, technology, and infrastructures play a key role, while landholding and possession of other assets substantially increase the probability of market entry. Boughton et al. (2007) analyse whether participation in remunerative cash crop markets - cotton and tobacco - is associated with higher asset endowments compared with participation in the less remunerative maize market. They confirm that greater endowments of assets such as land, livestock, labour, and equipment are indeed associated with participation in more remunerative markets. A more recent study by Mather et al. (2011) has confirmed that these assets also have a considerable effect within the maize market. Access to land, livestock, the use of animal traction and fertilisers, and better roads, means of transport, and information are all found to facilitate the passage from semi-subsistence to market agriculture (Mather et al., 2011). Poor infrastructures and market institutions are especially recognised as being more detrimental than other market imperfections for to the development of an integrated Mozambican agricultural market (Tostão and Brorsen, 2005; Cirera and

[1] They also propose new techniques for the estimation of marketing behaviours. For example, Bellemare and Barrett (2006) use an ordered tobit approach, while Holloway et al. (2005) apply a double-hurdle estimator.

Arndt, 2008; Cunguara and Darnhofer, 2011). Bingen et al. (2003) suggest that market participation is boosted when communities are provided with increased management skills and human capacity. Investments in these sectors by public or private institutions may enable small farmers to secure sustained benefits from market participation. Organisations and associations of farmers can also help small producers to reduce the transactions costs deriving from market incompleteness, and stimulate production and market participation (Tschirley et al., 2006; Markelova et al., 2009; Bachke, 2010).

2. A model of agricultural supply

Following a standard farm household model, as in Singh et al. (1986), now described is the effect of prices, income, other assets, technology, and information on the marketed surplus of staple crops. In this model, the household maximises utility (*U*) as a function of leisure (*l*) and staple crop (*y*), while a vector of household characteristics affecting tastes (***B***) is included, $U = U(l, y; \boldsymbol{B})$. Since the household is at the same time both a producer and consumer of *y*, it is possible to describe the marketed surplus of *y* as follows:

$$y^s = y^f - y^c \quad , \tag{1}$$

where y^s is the marketed surplus of *y*, y^f is the output of *y* produced by the household, and y^c is household consumption of *y*. The production function for *y* can be modelled as:

$$y^f = y^f(p, z, k) \quad , \tag{2}$$

while for the consumption part we have:

$$y^c = y^c\left[p, \boldsymbol{B}, A + p^l T(m) + \pi(p, z, k)\right] \quad , \tag{3}$$

where *p* is the price of the staple, p^l is the price of labour, ***B*** is the vector of household characteristics affecting taste, *T* is the time available to the household for work and leisure, *m* is the household characteristics determining *T*, *z* is a vector of farm characteristics, including fixed inputs, *k* is a vector of production technology parameters, *A* is exogenous income, and π is profits. It is possible to show that solving for the reduced form of marketed surplus leads to the following (Strauss, 1984):

$$y^s = y^s(p, \boldsymbol{B}, m, z, k, A) \quad . \tag{4}$$

Hence, the estimation of marketed surplus should include staple and labour price, household and farm characteristics, and any exogenous income. Moreover, when markets are complete, separability holds (Taylor and Adelman, 2003), and the household first maximises farm profits by deciding the amount of labour to be used. The decision on consumption of staple crop and leisure follows. However, when markets are incomplete - as is likely to be the case in Mozambique - demographics and non-negligible transaction costs influence the decision on production inputs, and consequently on farm profits (De Janvry et al., 1991). A study of marketed surplus must therefore take into account the interaction among production, the decision to participate in the market, and the quantity to be supplied to the market. Rather importantly for subsequent analyses, Sadoulet and De Janvry (1995) showed that the restrictions attached to supply and demand functions do not apply to marketed surplus. With adequate data, we could estimate both the output supply and the consumption demand functions for staple crops, and subsequently derive the marketed surplus. However, the dataset used does not comprise comprehensive information on staple crop demand and the proper structural restrictions could not be set up to ensure identification. The only available information that tells us something about the staples demanded is a dummy for whether the household had bought staple food crops in the previous thirty days or during the preceding hungry season. However, no information about quantities demanded is collected. Hence, in what follows we do not set up a more complex structural model but merely estimate a reduced form equation for marketed surplus.

3. Empirical estimation

With regard to the econometric estimation of the model, we refer to the empirical evidence that the decision to sell and the determination of the quantity to be sold are likely to follow separate processes in developing countries (Goetz, 1992; Key et al., 2000; Heltberg and Tarp, 2002; Holloway et al., 2005; Bellemare and Barrett, 2006; Boughton et al., 2007). Hence, we estimate the following reduced-form equation:

$$y_i^s = \alpha + \beta\, p_i + \gamma\, w_i + \delta\, \boldsymbol{X_i} + \zeta\, R + \epsilon_i \quad , \tag{5}$$

using a standard Heckman model with sample selection[2], and a double-hurdle model as a robustness check[3]. In Equation 5, y^s represents sold quantities of staple food crops aggregated in value. Output prices for staple crops are represented by p, and w is the agricultural wage rate. Other factors that affect staple crop production and commercialisation, such as household demographics, income, assets, landholding and ownership of livestock are denoted by the vector $\boldsymbol{X}$. Fixed transaction costs, which are known to affect commercialisation, are also included in $\boldsymbol{X}$: these include the availability of information on prices, membership of an agricultural association, and information received from extension agents. The number of years of education of the head of household and whether agriculture or husbandry is the household's principal activity are also contained in $\boldsymbol{X}$, as a proxy for management ability. Median maize and tobacco production per district, whether the household also produced horticultural or cash crops, and dummies for whether the household incurred any losses due to flood or drought are also included in $\boldsymbol{X}$. Regional dummies, R, are considered as well.

In order to estimate a supply function and identify the relationships, we also need a variable that shifts the demand schedule while not affecting supply. Since we know whether the household had bought staple food crops in the preceding hungry season, we include this dummy among the explanatory variables. We also consider a variable that registers the scarcity of staple crop reserves: this is computed as the proportion of months in which the household had no available reserves of the staple food crop considered to be the most important for its food security.[4] Indeed, depending on the level of available reserves, it might become necessary to buy food, but it is generally not possible to produce it because of the change of season. Thus, fewer reserves during the year may be a proxy for higher demand for staple food crops, and help to identify the supply schedule.

[2] Previously used for Mozambique by Benfica et al. (2006) and Heltberg and Tarp (2002) among others.

[3] See for example Holloway et al. (2005), Langyintuo and Mungoma (2008), Ricker-Gilbert et al. (2011) for an application of the double-hurdle model to the agricultural sector; or Nielson (2009) for an application of the double-hurdle model to the market of maize in Mozambique.

[4] 45.7% of Mozambican households - mostly in the central part of the country - indicate maize as the most important staple crop for their food security. Another 43.7% - mostly in the north - instead indicate cassava.

3.1 Market participation and quantities sold

With reference to the estimation technique, we estimate a Heckman and a double-hurdle model. While in the Heckman regression we correct for the self-selection of the non-selling household group, in the double-hurdle model quantities sold are observed only if the household crosses two hurdles. In the context of staple crop marketing decisions, the first hurdle involves the decision whether or not to participate in the market (participation decision). This may also depend upon social and demographic factors that are independent of the quantity consumed. The second hurdle concerns the amount of staple crops to be sold (selling decision). While in the Heckman selection model zeros are not affected by the selling decision, in double-hurdle models, when a household is observed to have zero selling, this is the result of either participation or selling decisions, and potential sellers may have zero staple crop selling (Aristei and Pieroni, 2008).

As said, the decision to estimate a reduced form equation for marketed surplus using selection and hurdle models was also motivated by the absence of information on the quantities of staple crops demanded. This led to exclusion of more recent and promising marketing decision estimation procedures, such as the ordered probit (Bellemare and Barrett, 2006), because it was not possible to distinguish between autarchic households and net buyers.

Notwithstanding the theoretical differences between the Heckman and double-hurdle models, the two are similar in spirit, and in what follows we present both estimation results. As evidenced in Wooldridge (2002), the Heckman model is better identified with exclusion restrictions similar to instrumental variables[5]. To this end, Key et al. (2000) and Heltberg and Tarp (2002) used fixed transaction costs as the key variable for the identification of the two processes in their estimation procedures. The reason was that, while variable transaction costs affect both the participation and quantity processes, fixed transaction costs only influence participation decisions. However, transaction costs are only partly observable. In their study, Heltberg and Tarp (2002) used 'information' variables as proxies for fixed transaction costs.[6] The choice of these variables is usually difficult - and frequently debatable - because the decisions are usually interconnected. Nonetheless, 'information' variables appear to be good

[5] It is usually not advisable to rely only on the non-linearity of the model (Wooldridge, 2002; pp. 564-566).

[6] 'Information' variables in their article include information about prices obtained before going to the market; access to radio, TV, and other tools that provide price information; and the ability to process the information received.

candidates.[7] We shall therefore seek to identify the selection model using these variables as proxies for fixed transaction costs, although we are aware that they are not ideal.

The double-hurdle results, which are presented alongside the Heckman results, may thus provide a robustness check for our estimates. The double-hurdle model does not need exclusion restrictions to be identified. In fact, the two tiers of the model follow different processes: the first hurdle determines the dichotomous participation decision, while a tobit model - the second hurdle - determines the amount sold (Cragg, 1971; Moffatt, 2005; Aristei and Pieroni, 2008).

The fact that the estimated coefficients are very similar in sign and magnitude in both the Heckman and the double-hurdle estimation gives us greater confidence in our estimation strategy, and lesser concern about the choice of instrumental variables. Both models are estimated via maximum likelihood and using the sample probability weights provided in the dataset.

3.2 The differences between smaller and larger sellers

As a consequence of the differing effects of certain regressors at different quantiles of the distribution of staple crops sold, we also estimate the impact of prices, household income, and other assets at 10%, 25%, 50%, 75%, and 90% of marketed staple distribution, using quantile regression. The quantile regression technique was first proposed by Koenker and Bassett (1978), and allows estimation of the entire distribution of the response variable dependent on any set of regressors. Quantile regression results suggest that while most of the coefficients remain the same for the whole distribution, some key variables present different coefficients for smaller and larger sellers. The pattern displayed could not be fully understood using standard analyses based on the conditional distribution mean. The quantile regression results are then compared with conditional mean estimates from the Heckman and double-hurdle models in Section 5[8].

---

[7] As also explained in the sections that follow, the exclusion of these variables from the quantity equation makes it clear that some sort of selection bias exists, while not changing the estimation results. Moreover, these variables appear not to be significant in the quantity equation of the double-hurdle model, which reinforces our hypothesis that they influence market participation relatively more than they influence the quantity of marketed surplus.

[8] When running the quantile regression, we do not correct for sample selection. Sample selection issues are of less concern in the quantile regression framework, because all data observations are used to construct each quantile regression estimate. See Koenker and Hallock (2001) for a discussion on why quantile regressions avoid the Heckman-type sample selection bias.

4. Data and descriptive statistics

The data used in this study are taken from the *Trabalho do Inquerito Agricola 2005*, TIA 2005, which is a nationally representative survey of rural households in Mozambique. It refers to the 2004/2005 agricultural season, and contains detailed data on agricultural production and marketing representative for each province and agro-ecological zone of Mozambique. The TIA 2005 covers ten provinces (the province of Maputo City and the city of Matola are excluded), and 94 districts out of 128.[9] The data were collected by the Ministry of Agriculture, with technical support from Michigan State University (Ministério da Agricultura, 2005).

The survey contains information on household demographics, income, assets, land ownership[10], crop production and sales, services, and technology. With regard to agricultural market participation, information on prices and quantities produced and sold is available for staple and cash crops.

The dependent variable used in this study is the marketed surplus of staple crops. This is computed by taking the logarithm of the total value of all staple crops sold, since quantities could not be directly added.

In regard to the independent variables, we know from the theoretical model presented in Sections 2 and 3 that we should consider the price of the staple, the wage rate, household characteristics affecting taste, household characteristics determining the time available to the household for work and leisure, farm characteristics, landholding and livestock, and exogenous income. Given that markets are likely to be incomplete in Mozambique, we also include household demographics, management ability, and transaction costs in the regression.

With regard to staple crop prices, we use the price of maize as a proxy in the study because maize is the most important staple crop and food reserve in hungry seasons. The price of maize is calculated as Meticais per kilogram (Mt/Kg), and we take the median for each district. Where this measure is not available, we use the province median price. We also include the district/province median price of tobacco in the regression, in order to consider the opportunity cost of producing cash crops instead of staple crops[11].

[9] Mozambique now has about 150 districts, due to reclassification.

[10] With regard to land, a sub-sample (25%) of all plots was measured by enumerators using GPS, and these measures allowed for correction for self-reporting errors, also taking into account the education levels of heads of household and the differences among districts (Mather et al., 2008).

The measure of household income used is computed as described in Mather et al. (2008). It is composed of farm and non-farm income, and includes crop income, sales of livestock products, unskilled farm wage labour, non-farm wage labour, micro- and small enterprise activities, remittances, and pensions. In the subsequent estimations, we consider the logarithm of total income divided by the number of adult equivalents, computed according to the OECD-modified scale[12] (Haagenaars et al., 1994).

For household and farm assets, we include both the adult equivalent amount of land owned by the household and the amount of livestock in tropical livestock units (TLUs). Ownership of a radio is also considered: this variable can be seen as both a part of household wealth and as a means to gather information on prices and agricultural practices.

As introduced in Section 3, we also consider a variable that registers the scarcity of staple crop reserves. This is computed as the proportion of months in which the household did not have available reserves of the staple food crop that it considers to be the most important for its food security (mainly maize or cassava). This variable can be considered a shifter of the demand schedule of staple crops that affects supply only marginally, thus helping to identify the underlying econometric relation.

Technological variables which are included are the median provincial production of maize and tobacco, the use of fertiliser, and ownership of improved granaries. The use of median provincial production of maize and tobacco, in particular, is an attempt to proxy for the simultaneous effect of technology, climate, and past investment (Heltberg and Tarp, 2002).

With regard to the other household characteristics we include in the analysis the age of the head of household; years of education; a dummy if the head of household is a woman; the number of healthy members of the household (that is, household members without serious illnesses divided by total household members); a dummy if some member of the household is affiliated with an agrarian association; and a dummy if some member of the household earns a wage or a salary. This variable is also important for evaluating the effect of other income sources on market participation. Finally, whether any of the household members received price

[11] Tobacco is the most important cash crop in Mozambique (Benfica et al., 2006; Boughton et al., 2007).

[12] This equivalence scale assigns a value of 1 to the household head, a value of 0.5 to any other adult in the household, and a value of 0.3 to children.

information or agriculture-related information from extension agents is also included, together with other geographical controls such as regional dummies and whether the community has experienced natural disasters (drought or flood).

Table 1 presents some descriptive statistics for the staple crops considered, while Table 2 displays the descriptive statistics for all the variables used in the regression analysis.

Table 1 indicates that maize is the most important staple food crop: 79.1% of the households in the sample cultivate maize, and it represents the most important food reserve in hungry seasons. The central and northern regions of the country are those in which the most maize is produced, owing to more favourable climate conditions.[13] However, only 22.3% of producers sell at least some on the market, and 54.9% of sellers sell their entire production in a single day. Prices vary considerably among crops and across regions and districts, the median selling price for maize being 3 Mt/kg (Meticais/kilogram). However, in Maputo Province the median selling price of maize is much higher (6 Mt/kg). The median (seasonal) sales value of maize also varies across provinces, from 200 Meticais in Inhambane to 740 Meticais in Maputo Province.

The second most important staple crop is cassava, which is also recognised as being the most important staple crop for food security in the northern region. It is produced by 72.5% of households, and sold by just 18.3% of them. As in the case of maize, more than half of the sellers sell their entire production in a single day (51.3%).

5. Estimation results

The estimation results for staple crop marketed surpluses at a national level are set out in Tables 3 and 4. The first two columns of Table 3 present the results of the Heckman model - quantity and selection - while the double-hurdle results are reported in the third and fourth columns. In Table 4, the quantile regression results are shown for different quantiles of the distribution of staple crop sold.

The effect of the price of maize on quantities sold displays the expected positive sign, even though the magnitude of the coefficient varies along the distribution of staple crops sold. A price increase for maize seems to have a greater impact on the average quantities of staple crops sold - quantiles from 0.25 to 0.50 - than it does on

[13] See Tschirley et al. (2006) and Boughton et al. (2007) for a discussion on the production and marketing of maize in Mozambique.

larger quantities. Conversely, for very small sellers (up to the 0.15 quantile), the effect of the price of maize appears not to be significant. This group of sellers has a significantly lower income than the other market participants (*p-value* < 0.01). It is therefore likely that poorer households sell staple crops only when strictly necessary or in response to adverse shocks, and for this reason are not especially concerned about the price (distress sales). We do not have information on motivations for selling, so we cannot offer strong evidence to back this statement. However, if we consider sale of the whole production in one day as a proxy for a distress sale, we find that the percentage of smaller sellers (up to the 0.15 quantile) selling their entire maize production in one day is 78%, whereas the national average is about 55%. Also, the percentage of smaller sellers selling their entire production of cassava in one day is 87%, compared with an average for the overall sample of about 51%. Moreover, smaller sellers report exceptional sales of animals and other goods as a survival strategy more than the population as a whole (23% versus 18%).

The price of maize does not seem to influence the likelihood of entering the market (selection).[14]

Households with higher equivalent incomes sell significantly more than do poorer households: income influences the quantities sold and participation in the market, and its effects are even stronger at the very end of the distribution of staple crop sold. A one per cent increase in income stimulates staple crop sales by 0.51 per cent on average, by 0.41 at the median, and by 0.61 and 0.77 at the end of the distribution - respectively quantile 0.90 and 0.99.

Similar, but weaker, effects are observed for other assets: land (0.25), livestock (0.04), and improved granaries (0.39). In the case of livestock, the effect is significant only for households that sell medium to high quantities of staple crops, whereas improved granaries affect smaller sellers comparatively more (0.83 compared to a median effect of 0.20 and to a mean effect of 0.39). The possibility of storing agricultural products is consequently important, probably because the lack of adequate storage facilities prevents households not only from smoothing consumption but also from negotiating better trading conditions.

Whether agriculture or husbandry are a household's main activity indicates both increased management and marketing ability and greater agricultural knowledge: the

[14] We are cautious about presenting price elasticity, as price responsiveness is hard to identify in cross-sectional data.

coefficient associated with this variable is thus positive, as expected, and significant in both the selection and the quantity equation, especially for smaller sellers.

The use of fertilisers is also associated with a greater probability of participation in the market (13%): this generally denotes greater agricultural knowledge or more frequent contacts with associations, NGOs, and extension agents, which often provide fertilisers or give information about them directly. The use of part-time or full-time agricultural workers is also significant in the quantity equation, with a positive sign (0.27).

Whether a member of a household receives a wage or salary instead reduces both the probability of participating in the staple food market and the quantity sold. This is in line with previous findings, and probably reflects the existence of different coping strategies, as suggested by Heltberg and Tarp (2002).

Geography is also a key aspect of market participation: as stated above, and as emphasised by Boughton et al. (2007), the northern and central regions of the country are those where more staple food crops - and especially more maize - are produced. As a consequence, there is a greater likelihood that households in the north and centre will sell their products (33% and 27% respectively), and they also sell relatively larger amounts than those in the south. Households in areas that have experienced a drought are also observed to sell lesser amounts of staple food crops, though the effect is only significant at 10% significance level.

As proposed by Heltberg and Tarp (2002), possession of price information - and the ability to process it - may proxy for unobservable fixed transaction costs. These costs only affect the decision to commercialise staple crops, and not the quantity sold. Indeed, in our estimation, having information on prices has the expected positive sign in the participation equation, while not affecting the quantities sold.[15] Some caution is necessary when interpreting the results concerning selection, however. The identification of the model might be greatly improved with less weak exclusion restrictions or better proxies for fixed transaction costs.

A check for the robustness of our estimation procedure is provided by the double-hurdle results. The coefficients estimated via double-hurdle are very similar in sign and magnitude to those estimated using the Heckman model. Moreover, the correlation parameter (*rho*) for the two tiers of the double-hurdle model is positive, very significant, and comparable in magnitude with the one estimated in the

[15] As results from the double-hurdle estimations presented along with the Heckman results.

Heckman model. This makes us more confident about the existence of sample selection, and means that we are less concerned about our choice of instrumental variables.

Production of cash crops or horticultural crops both display a negative sign. These are more remunerative activities compared with staple crop cultivation, and it is thus reasonable to expect that households producing these crops will sell less in the staple crop market.

Finally, an increase in the proportion of months without reserves of staple food crops significantly discourages households from participating in the market (-24%), but also negatively influences the amount of staple crops sold (-0.43). The same negative effect on market participation is observed for those households that state that they have bought staple food crops in the previous hungry season (-6%). It seems that the variables that shift the demand schedule and help to identify the supply equation work in the expected direction: they are strongly significant with negative sign.

Furthermore, if the number of meals the household was able to afford in the previous hungry season is higher, this is associated with both a higher probability of market participation (4%) and a larger marketed surplus (0.24). This variable appears to be much more important at lower quantiles of the distribution of staples sold than it is at medium-high quantiles: its effect decreases from 0.44 at quantile 0.10 to 0 after quantile 0.75.

The decision to model agricultural decisions with sample selection and hurdle models seems to be justified in our analysis: as shown at the bottom of Table 3, both the Heckman and the double-hurdle model display a positive and significantly different from zero estimate for *rho*. This means that the lambda term (*rho***sigma*) has a positive sign, which suggests that the error terms in the selection and quantity equation are positively correlated. Hence, (unobserved) factors that make market participation (selection) more likely tend to be associated with greater quantities of staple crops sold. Households that sell lower quantities are relatively more likely not to participate in the market at all.

Tables 3 and 4 summarise all the estimation results.[16] Marginal effects for the probability of selection (*psel*) and for the expected amount of staple crops sold, given

---

[16] The double-hurdle estimates are obtained using the user-developed STATA program called "dhurdle", written by J. Fennema. This makes it possible to have correlated error terms in the two tiers and to test this hypothesis. The ado file is available at: http://www.sml.hw.ac.uk/staffpages/somjaf/Stata/.

that the household participates in the market (*ycond*), are presented in Appendix A (Table A.1).

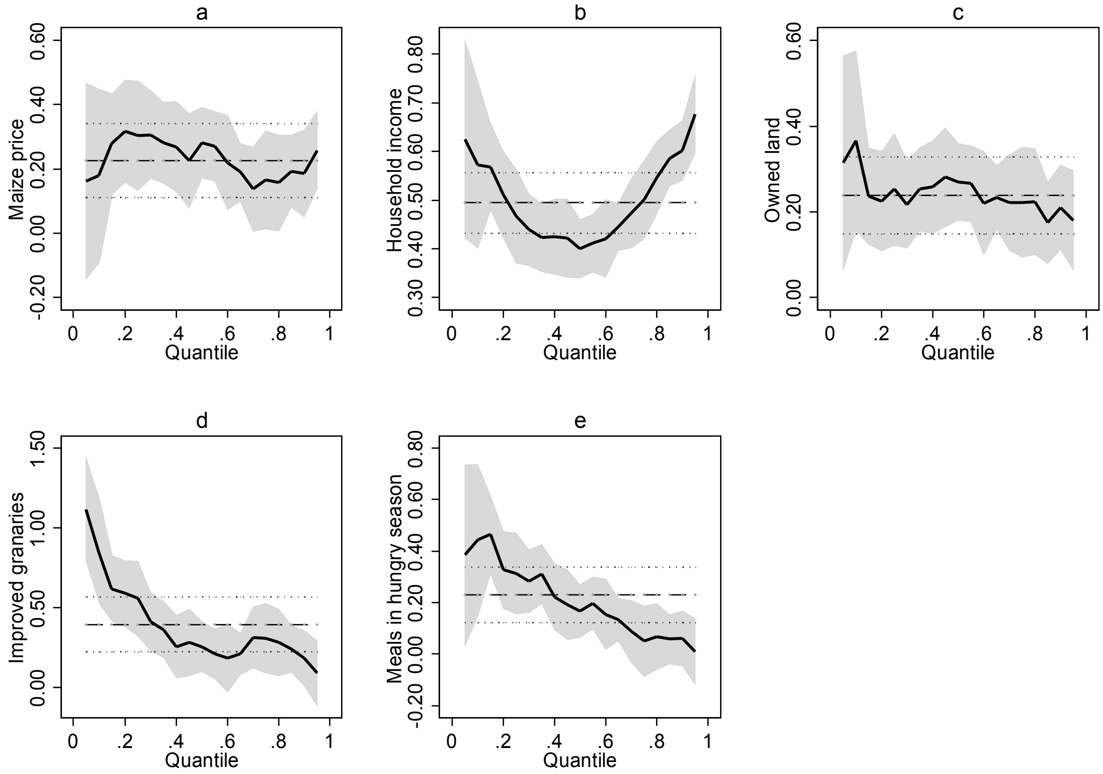
5.1 Analysis of quantile regression results

For a graphical analysis of the quantile regression results, we plot some of the estimated coefficients in Figure 1, ranging from quantile 0.05 to quantile 0.95. The interpretation of each point on the thick solid line is the effect of a unit change of the independent variable on the dependent variable, keeping all the other independent variables fixed. The thin dashed line is the OLS estimate, while the grey area represents the 90% confidence band for the quantile regression estimates (Koenker and Hallock, 2001). The OLS estimate is constant throughout the interval because it represents the estimate of the effect of the independent variable at the mean of the distribution of the dependent variable.

Most of the coefficients do not vary substantially when they are analysed at different points of the distribution of staple crop marketed surplus. However, the effect of the price of maize, household income, land, improved granaries, and the number of meals that could be afforded in the previous hungry season varies substantially at different quantiles (of sold quantities).

As shown by Figure 1, the effect of the price of maize on quantities sold is not significant at low quantiles of the distribution of staple crops sold (panel a). This suggests that for households selling very small amounts of staple crops, the price of maize might not be a key decision variable. Conversely, an increase in household income has a highly non-linear effect: the coefficient decreases up to the median of the distribution of staple crops sold. After this point, it peaks at 0.77 at quantile 0.99 (panel b). Higher income is, therefore, a crucial variable in determining the marketing behaviour of Mozambican rural households, the extent of which could not be observed by relying solely on conditional mean estimates. The effect of land per adult equivalent also differs at low quantiles of quantity sold (up to 0.15) with respect to higher quantiles (panel c), but not as much as it does for improved granaries (panel d) and the number of meals that could be afforded in the previous hungry season (panel e). For the latter two variables, the effect on quantity sold decreases steeply, which denotes that providing rural households with more meals or improved granaries might have significant effects, especially at low levels of marketing participation.

Figure 1 – Quantile regression for selected coefficients



Thick solid line: the effect of a unit change of the independent variable on the dependent variable, keeping all the other independent variables fixed.
Grey area: 90% confidence band for the quantile regression estimates.
Thin dashed line: OLS estimate. Thin dotted lines: 90% confidence band for the OLS estimate.

5.2 Regional analysis

Mozambique is divided into three macro-regions, the north, the centre and the south.[17] These are characterised by differing agricultural practices, production and marketing attitudes, rainfall patterns, and numbers of resident rural households (Boughton et al., 2007). It is thus reasonable to expect that marketing decisions and market participation will also differ among the macro-regions.[18]

In particular, it emerges from our analysis that income, asset endowments, the number of meals, and available reserves influence rural household marketing behaviours relatively more in the centre and the north, whereas 'information'

[17] The north includes the provinces of Niassa, Cabo Delgado, and Nampula, while the centre is formed of the provinces of Zambezia, Tete, Manica, and Sofala. The southern provinces are Inhambane, Gaza, Maputo Province, and Maputo City.

[18] Indeed, if a Wald test is performed for the national and regional coefficients, the results indicate that the estimated regional coefficients are different from the nationally computed ones. In particular, the estimated coefficients for the north are not statistically different from the nationally computed ones (*F = 1.10*), while those estimated for the centre are significantly different from the national coefficients only at 10% significance level (*F = 1.50**). Conversely, the coefficients estimated for the south are significantly different from the nationally computed ones (*F = 4.97****).

variables are more likely to affect rural household decisions in the south. Boughton et al. (2007) have already pointed out that the northern and central macro-regions have more similar characteristics with respect to staple crop markets compared with the south. Hence, a regional analysis is first run separately for each macro-region, and then the estimation is conducted by aggregating the centre and the north.[19]

Table 5 sets out the results of the empirical model used.[20] The main differences with respect to the analysis conducted at national level are the following:

- The coefficient for the price of maize is significant in the participation equation for the south, and in the quantity equation for the centre (and centre-north).
- While income maintains its positive coefficient in the three macro-regions, its effect appears to be much stronger in the north and centre than it is in the south.
- Land and livestock ownership are significantly associated with increased quantities sold only in the centre, while access to improved granaries boosts the amount of staple crops sold in both the centre and the north.
- The proportion of months without reserves has a strong negative effect on quantity and selection in the centre, while influencing only market participation in the north.
- Information on prices and information received from extension agents only affects participation in the south, while it is not significant in the other two regions.

Appendix A presents the expected amount of staple crops sold given that the household participates in the market (*ycond*) for the centre-north and the south (Table A.2). District dummies are also considered, but no major changes are apparent when they are included. With regard to the issue of sample selection, the only region where there is some evidence for this is the centre.

6. Staple food marketing decisions and poverty

In this Section, we analyse whether market participation and the quantity of staple crop sold differ between poorer and richer households. On the basis of the household

[19] The Wald test confirms that the coefficients for the north and centre are not statistically different ($F = 1.27$).

[20] Only the Heckman estimation is shown in Table 5.

equivalent income variable already described, we conduct a series of qualitative analyses on the relationship between poverty status and marketing behaviours.

Heltberg and Tarp (2002) analyse this issue in their study, finding that while poor and non-poor have very similar coefficients - and hence similar behaviours - they differ with respect to the likelihood of selling, with the non-poor being more likely to sell both staple and cash crops. By contrast, Boughton et al. (2007) suggest that owned land, livestock, animal traction, income sources other than farm income, and education are indeed associated with higher participation and/or higher sales value.

Descriptive statistics for the variables used in previous estimations are presented in Table 6, disaggregated for poor and non-poor.[21] Asterisks are displayed in the usual fashion when the mean for the two groups is significantly different. Notably, most of the variables considered in our study differ considerably between poor and non-poor households.

It emerges that poorer households are in general less endowed with farm and household assets, have fewer years of education, are headed by women, live in the northern and - to a lesser extent - central regions, participate less in the staple market, and sell smaller quantities at a lower price. Moreover, most poorer households engage in agriculture/husbandry as their main activity, but receive considerably less information from extension agents and information on prices.

A robustness check for the relation between marketing decisions and socioeconomic status is also carried out, using a constructed wealth index instead of household income. The results are presented in Appendix B.

6.1 Qualitative analyses

In the following figures, we present qualitative analyses regarding marketing decisions and income and asset endowments. First, we examine the relation between sold quantities of staple crops - the dependent variable in previous regressions - and income. In Figure 3, the result of a local polynomial smoothing (Epanechnikov kernel, bandwidth 0.8) of the – actual - (log) value of staple crops sold on household equivalent income is plotted (panel a). As expected, the result is an increasing curve. The relation increases at all levels of income, becoming slightly steeper at very high income levels.

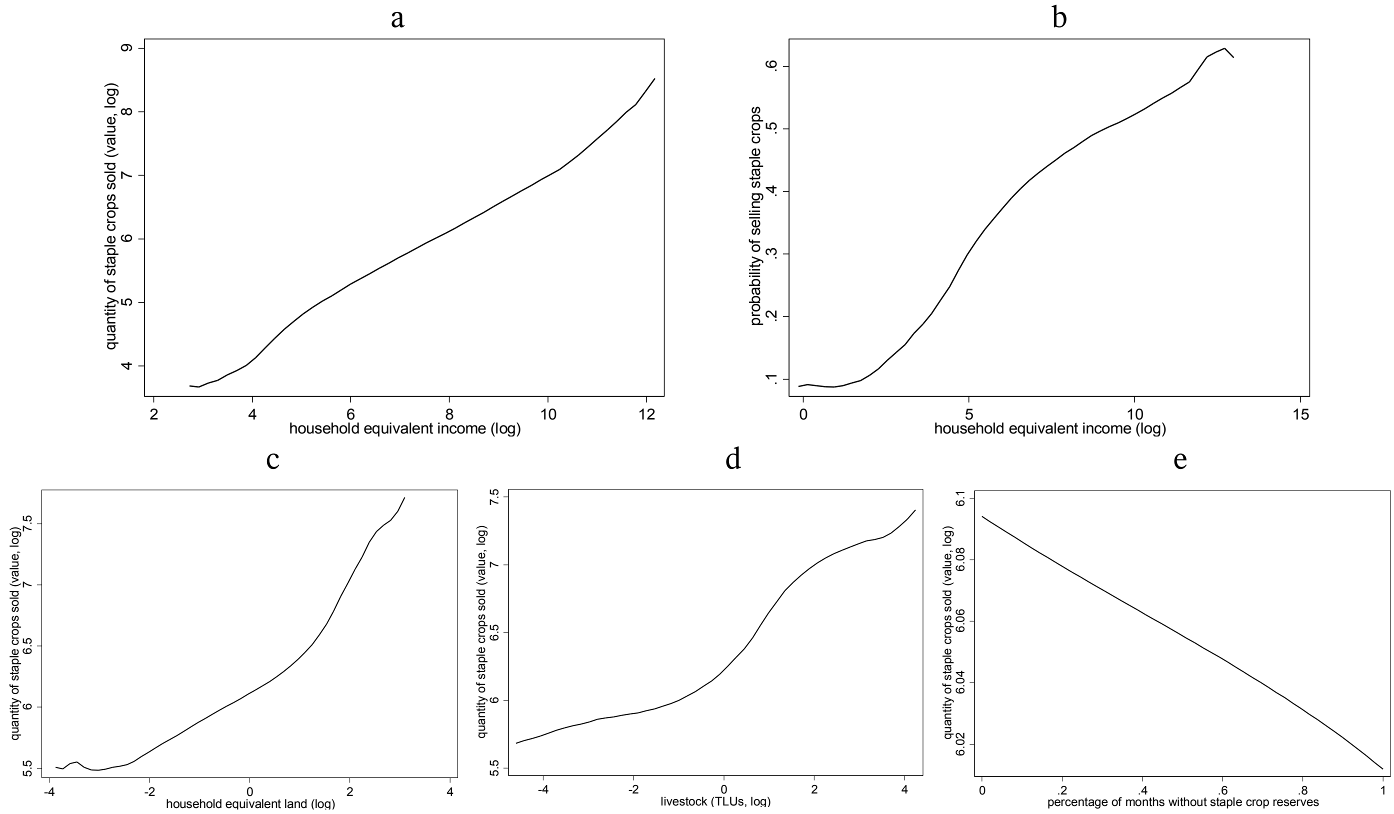
[21] Here we consider as poor those households who are in the first two quintiles of the income distribution (bottom 40%), and as non-poor those who are in the fourth and fifth quintile of the income distribution (top 40%).

We also explore whether market participation in staple food markets is associated with higher household income. The probability of selling staple crops (*psel*) is predicted after the Heckman model and placed in relation to household income. The local polynomial smoothing again displays that the probability of selling staple crops increases substantially with income, especially after that a threshold -(log) income close to 2.0- is passed (panel b).

In the bottom panels of Figure 3, we also plot a non-parametric relation via local polynomial smoothing for the quantity of staple crop sold and owned land (panel c), and for the quantity of staple crop sold and livestock owned (panel d). In both cases, we observe a sizeable increase in the quantity sold if households own more land or livestock, although the pattern is different. For livestock, the relation peaks after a value of (log) TLUs of about -1.0, while for land the relation presents a sort of exponential shape. It starts to increase at values of (log) hectares of owned land of about -3.0, and peaks at values greater than 1.0. Conversely, a decreasing curve for the relation between quantity sold and the proportion of months without reserves is observed (panel e).

Especially for panels b, c and d a highly non-linear effect is observed. Households who are above a certain threshold of income, owned land or livestock appear to participate in the staple crop market sensibly more, and to sell increasingly greater quantities. It may be that below a given threshold the possibility to escape from subsistence agriculture and enter into market agriculture is very remote. Although the extent of such threshold and the causality link between income/assets and market participation/selling would require some future investigation, the two phenomena appear to be strictly associated.

Figure 2 – Local polynomial smoothing of the quantity of staple crops sold on household income and other assets



7. Policy implications and conclusions

The paper has presented an analysis of the decisions made by rural households concerning staple food marketing in Mozambique. It has analysed both the decision to participate in the staple crop market and the quantity of marketed surplus, applying sample-selection and double-hurdle models. The role of income, asset endowments, technology used, information, and geography emerged as crucial in determining marketing decisions and outcomes.

Given the existing differences among the three macro-regions of Mozambique, estimation was also made of the determinants of staple crop marketing separately for the northern, central and southern regions. It emerged that income, asset endowment, the number of meals, and the available reserves in hungry seasons influence rural household marketing behaviours relatively more in the northern and central regions, whereas 'information' variables are more likely to affect rural household decisions in the south.

As a further step, maize prices, income, and other assets were studied in more detail by using quantile regression analyses. It was observed that the effect of these determinants varies for larger and smaller sellers. It emerged that income has a decreasing effect up to the median of the distribution of staple crops sold, but displays a growing effect at higher quantiles of staple crops sold. Also, the effect of land per adult equivalent differs at low quantiles of quantity sold with respect to higher quantiles, but not as much as it does for improved granaries and the number of meals that could be afforded in the previous hungry season. For the latter two variables, the effect on quantity sold decreases steeply, denoting that providing rural households with more meals or improved granaries might have noteworthy effects, especially for smaller sellers.

Instead, the price of maize does not seem to be a significant variable for small sellers (up to the 0.15 quantile). This group of sellers also has a significantly lower income than the rest of the market participants (*p-value* < 0.01). Hence, the possibility that poorer households sell staple crops only when strictly necessary or in response to adverse shocks, and are thus not especially concerned about price (distress sales), should be investigated more thoroughly.

All these variables, which are crucial in determining the marketing behaviour of Mozambican rural households, exhibit a pattern that could not be observed by relying solely on conditional mean estimates. The results obtained with quantile analysis evidence that agricultural policies need to be well-designed, as different outcomes may arise depending on the characteristics of the targeted population.

Then further examined was the relation among market participation, marketed surplus and income/assets. It turned out that most of the variables used in the regression are significantly different for poorer and richer households. Poorer households are generally less endowed with farm and household assets, have fewer years of education, are headed by a woman, live in the northern and central regions, and sell fewer staple crops at a lower price. Most of them engage in agriculture or husbandry as their main activity, but they receive considerably less information from extension agents or information on prices. Moreover, the quantity of staple crops sold in the market is associated with greater income, but also with greater farm and household asset endowments.

As stated in the introduction, and as often acknowledged by the Mozambican government, rural household participation in the agricultural market can be extremely important for poverty reduction and the overall development of the country (Tschirley et al., 2006). Nonetheless, the design of agricultural policies aimed at increasing marketing participation in rural environments should consider all the variables that affect this complex process. In particular, income, assets, technology, and information play a significant role in influencing marketing decisions regarding staple food commodities. Both market participation and marketed staple crop surpluses display a strong positive association with greater income and asset endowments. The evidence that this relation is highly non-linear and the possibility that one or more thresholds need to be passed in order to bring about a switch from subsistence to market agriculture emerge as a groundwork for future work on Mozambican data. Furthermore, the significant differences that exist concerning marketing decisions at regional level and between larger and smaller sellers should also be considered.

Further investigation of this problem might also shed light on the causation link between income, asset endowments, and marketing decisions. In any event, the association between income/assets and marketing behaviours evidenced here reinforces the argument that accompanying policies should be undertaken in order to obtain more effective outcomes in the field of agricultural market participation and rural poverty reduction.

The study has thus attempted to provide further insights into Mozambican agricultural market behaviours, and to contribute towards expanding the literature on staple crop marketing decisions for rural households in Mozambique. We believe that this will be of help in designing future agricultural policies and rural poverty reduction strategies.

Acknowledgements

The author is grateful to Vasco Nhabinde, Monica Magaua, Aurelio Mate, Francesca Bettio, Bruna Ingrao, Leonardo Ditta, Tushar Nandi, Pierpaolo Pierani, Novella Maugeri, Valentina Gil, Laura Anselmi, and the Centro de Estudos de Economia e de Gestao (CEEG), Universidade Eduardo Mondlane, Maputo, Mozambique.

References

Aristei, D., Pieroni, L., 2008. A double-hurdle approach to modelling tobacco consumption in Italy. Applied Economics 40(19), 2463-2476.

Bachke, M., 2010. Do Farmers' Organizations Enhance The Welfare of Small-Scale Farmers?. Working Paper, Norwegian University of Life Sciences.

Barrett, C., 2008. Smallholder market participation: concepts and evidence from eastern and southern Africa. Food Policy 33(4), 299-317.

Bellemare, M., Barrett, C., 2006. An ordered Tobit model of market participation: Evidence from Kenya and Ethiopia. American Journal of Agricultural Economics 88(2), 324.

Benfica, R., Tschirley, D., Boughton, D., 2006. Interlinked transactions in cash cropping economies: the determinants of farmer participation and performance in the Zambezi river valley of Mozambique. International Association of Agricultural Economics, August, Gold Coast (Australia). East Lansing: Department of Agricultural Economics, Michigan State University.

Bingen, J., Serrano, A., Howard, J., 2003. Linking farmers to markets: different approaches to human capital development. Food Policy 28(4), 405-419.

Boughton, D., Mather, D., Barrett, C., Benfica, R., Abdula, D., Tschirley, D., Cunguara, B., 2007. Market participation by rural households in a low-income country: An asset-based approach applied to Mozambique. Faith and Economics 50(1), 64-101.

Cirera, X., Arndt, C., 2008. Measuring the impact of road rehabilitation on spatial market efficiency in maize markets in Mozambique. Agricultural Economics 39(1), 17-28.

Cragg, J., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39, 829-844.

Cunguara, B., Darnhofer, I., 2011. Assessing the impact of improved agricultural technologies on household income in rural Mozambique. Food Policy 36(3), 378-390.

De Janvry, A., Fafchamps, M., Sadoulet, E., 1991. Peasant household behaviour with missing markets: Some paradoxes explained. The Economic Journal 101(409), 1400-1417.

Fafchamps, M., 1992. Cash crop production, food price volatility, and rural market integration in the third world. American Journal of Agricultural Economics 74, 90-99.

Goetz, S., 1992. A selectivity model of household food marketing behavior in sub-Saharan Africa. American Journal of Agricultural Economics 74, 444-452.

Haagenaars, A., de Vos, K., Zaidi, M., 1994. Poverty Statistics in the Late 1980s: Research Based on Micro-data. Office for Official Publications of the European Communities. Luxembourg.

Heltberg, R., Tarp, F., 2002. Agricultural supply response and poverty in Mozambique. Food Policy 27(2), 103-124.

Holloway, G., Barrett, C., Ehui, S., 2005. The double hurdle model in the presence of fixed costs. Journal of International Agricultural Trade and Development 1, 17-28.

Jayne, T., Mukumbu, M., Chisvo, M., Tschirley, D., Weber, M., Zulu, B., Johansson, R., Santos, P., Soroko, D., 1999. Successes and challenges of food market reform: Experiences from Kenya, Mozambique, Zambia, and Zimbabwe. Food Security International Development Working Papers.

Key, N., Sadoulet, E., De Janvry, A., 2000. Transactions costs and agricultural household supply response. American Journal of Agricultural Economics 82, 245-259.

Koenker, R., Bassett Jr, G., 1978. Regression quantiles. Econometrica 46, 33-50.

Koenker, R., Hallock, K., 2001. Quantile regression. The Journal of Economic Perspectives 15(4), 143-156.

Kolenikov, S., Angeles, G., 2004. The use of discrete data in PCA: theory, simulations, and applications to socioeconomic indices. Technical Report, MEASURE/Evaluation project, Carolina Population Center, University of North Carolina, Chapel Hill.

Kolenikov, S., Angeles, G., 2009. Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer?. Review of Income and Wealth 55(1), 128-165.

Langyintuo, A., Mungoma, C., 2008. The effect of household wealth on the adoption of improved maize varieties in Zambia. Food Policy 33(6), 550-559.

Markelova, H., Meinzen-Dick, R., Hellin, J., Dohrn, S., 2009. Collective action for smallholder market access. Food Policy 34(1), 1-7.

Mather, D., Cunguara, B., Boughton, D., 2008. Household Income and Assets in Rural Mozambique, 2002-2005: Can Pro-Poor Growth Be Sustained?. Mozambique Ministry of Agriculture Research Paper Series 66.

Mather, D., Donovan, C., Jayne, T., Weber, M., 2011. Smallholder Heterogeneity and Maize Market Participation in Southern and Eastern Africa: Implications for Investment Strategies to Increase Marketed Food Staple Supply. MSU International Development Working Paper 113.

Ministério da Agricultura, 2005. Trabalho do Inquérito Agrícola 2005. Departamento de Estatística, Direcção de Economia, MINAG, República de Moçambique, Maputo, Mozambique.

Moffatt, P., 2005. Hurdle models of loan default. Journal of the Operational Research Society 56(9), 1063-1071.

Nielson, H., 2009. The role of cassava in smallholder maize marketing in Zambia and Mozambique. Unpublished M.S. Thesis, Michigan State University.

Olsson, U., 1979. Maximum likelihood estimation of the polychoric correlation coefficient. Psychometrika 44(4), 443-460.

Ricker-Gilbert, J., Jayne, T.S., Chirwa, E., 2011. Subsidies and Crowding Out: A Double-Hurdle Model of Fertilizer Demand in Malawi. American Journal of Agricultural Economics 93(1), 26-42.

Sadoulet, E., De Janvry, A., 1995. Quantitative development policy analysis. Johns Hopkins University Press, Baltimore and London.

Singh, I., Squire, L., Strauss, J., 1986. Agricultural household models: Extensions, applications, and policy. Johns Hopkins University Press, Baltimore.

Strauss, J., 1984. Marketed surpluses of agricultural households in Sierra Leone. American Journal of Agricultural Economics 66, 321-331.

Taylor, J., Adelman, I., 2003. Agricultural household models: genesis, evolution, and extensions. Review of Economics of the Household 1(1), 33-58.

Tostão, E., Brorsen, B., 2005. Spatial price efficiency in Mozambique's post-reform maize markets. Agricultural Economics 33(2), 205-214.

Tschirley, D., Abdula, D., Weber, M., 2006. Toward Improved Maize Marketing and Trade Policies to Promote Household Food Security in Central and Southern Mozambique. MINAG/DE/DAP Research Report 60E.

Walker, T., Tschirley, D., Low, J., Tanque, M., Boughton, D., Payongayong, E., Weber, M., 2004. Determinants of Rural Income in Mozambique in 2001-2002. MINAG/DE/DAP Research Report 57E.
Wooldridge, J., 2002. Econometric analysis of cross section and panel data. The MIT press.
World Bank, 2008. Mozambique at a glance. URL: http://devdata.worldbank.org/AAG/moz_aag.pdf, accessed on 01/06/2010.
Zadeh, L., 1965. Fuzzy sets. Information and control 8(3), 338-353.

Table 1 – Descriptive statistics for staple food crops

Crop	Production (%)	North (%)	Centre (%)	South (%)	Male hh head (%)	Female hh head (%)	Obs	Selling (%)	North (%)	Centre (%)	South (%)	Male hh head (%)	Female hh head (%)	Obs
Groundnut (big)	13.36	8.30	4.79	0.27	10.01	3.36	6149	24.52	18.23	6.29	0.00	20.06	4.47	816
Groundnut (small)	29.35	13.52	7.46	8.37	21.71	7.63	6149	22.91	17.68	4.53	0.71	19.05	3.87	1880
Rice	19.50	5.82	12.87	0.81	15.07	4.43	6149	10.45	3.54	6.22	0.70	8.85	1.60	873
Sorghum	28.98	13.68	14.39	0.92	22.55	6.43	6149	4.96	2.58	0.0237	0.02	3.90	1.07	1938
Millet	5.44	0.99	4.13	0.32	4.34	1.10	6149	4.23	0.98	0.0318	0.08	3.19	1.05	547
Maize	79.13	28.69	35.40	15.04	60.28	18.85	6149	22.29	8.22	13.35	0.73	18.55	3.74	5195
Bean (*boer*)	21.82	8.42	13.06	0.34	17.78	4.04	6149	15.39	3.90	0.1149	0.00	12.68	2.71	977
Bean (*jugo*)	12.32	8.19	2.62	1.51	9.34	2.98	6149	3.97	2.87	0.0107	0.03	3.57	0.40	757
Bean (*manteiga*)	11.11	2.69	6.18	2.24	8.08	3.03	6149	34.08	11.37	20.61	2.09	27.39	6.68	873
Bean (*nhemba*)	50.42	18.80	19.54	12.09	37.53	12.89	6149	9.07	4.47	3.81	0.79	7.52	1.55	3369
Bean (*oloko*)	2.88	1.71	0.55	0.62	2.08	0.80	6149	8.14	2.42	4.93	0.79	6.87	1.27	133
Bean (*outro*)	0.24	0.18	0.05	0.02	0.18	0.07	6149	21.53	5.92	9.42	6.19	21.39	0.14	24
Sweet potato (*alaranjada*)	2.36	0.49	0.92	0.96	1.89	0.47	6149	12.98	7.29	4.85	0.84	11.06	1.91	224
Sweet potato (*nao alaranjada*)	15.46	1.42	10.50	3.54	11.60	3.86	6149	14.62	2.01	10.88	1.72	11.98	2.63	1074
Cassava	72.48	31.56	27.02	13.89	55.29	17.18	6149	18.34	7.85	7.52	2.97	14.77	3.57	3813
Cashew nut	31.72	12.99	8.79	9.94	23.95	7.78	6149	40.11	27.44	7.38	5.28	31.71	8.40	1862
Coconut	16.02	3.74	5.73	6.55	11.83	4.18	6148	32.83	7.92	13.92	10.99	24.00	8.83	750

Crop	Selling in one day (%)	North (%)	Centre (%)	South (%)	Male hh head (%)	Female hh head (%)	Obs	Median selling price (Mt/Kg)	North	Centre	South	Male hh head	Female hh head	Obs
Groundnut (big)	71.51	57.73	13.78	0.00	53.81	17.70	189	8.94	8.69	9.37	-	8.12	9.41	180
Groundnut (small)	77.44	61.62	13.69	2.12	63.53	13.90	272	7.12	7.20	6.44	11.04	7.12	7.68	261
Rice	58.27	22.52	31.45	4.30	44.94	13.33	110	5.93	5.93	6.00	6.01	5.93	8.25	101
Sorghum	64.68	36.27	28.13	0.28	50.63	14.05	91	3.02	3.02	2.80	3.48	3.02	3.63	89
Millet	76.10	23.14	51.13	1.83	51.36	24.74	14	5.69	1.68	5.69	3.42	5.69	1.58	14
Maize	54.91	23.76	29.78	1.37	45.24	9.67	1018	3.00	3.00	2.87	4.60	3.00	3.00	987
Bean (*boer*)	79.84	23.43	0.564	0.01	65.07	14.77	132	3.00	2.52	3.02	17.50	3.02	2.52	126
Bean (*jugo*)	60.12	47.27	12.08	0.77	57.89	2.23	25	3.77	4.37	3.09	14.79	3.09	6.78	25
Bean (*manteiga*)	68.08	21.63	43.18	3.28	52.85	15.23	256	8.78	8.41	8.78	14.63	8.78	8.97	242
Bean (*nhemba*)	78.83	42.55	32.25	4.03	65.28	13.54	229	5.00	4.75	5.16	7.58	5.00	5.16	218
Bean (*oloko*)	84.35	29.71	44.87	9.76	68.70	15.65	7	5.63	6.00	4.68	8.04	5.63	2.46	7
Bean (*outro*)	73.79	27.50	43.75	2.54	73.79	0.00	10	-	-	-	-	-	-	0
Sweet potato (*alaranjada*)	67.13	36.63	27.11	3.40	57.99	9.14	28	1.10	1.10	1.93	4.00	1.10	11.90	16
Sweet potato (*nao alaranjada*)	44.58	7.69	32.89	3.99	37.38	7.20	142	2.48	3.57	2.38	2.76	2.38	3.57	115
Cassava	51.31	24.67	18.78	7.87	39.28	12.03	634	0.68	0.46	0.68	2.64	0.68	0.69	462
Cashew nut	78.77	59.74	12.89	6.14	62.36	16.41	596	6.69	7.00	5.82	5.82	6.86	6.02	560
Coconut	25.75	7.48	9.58	8.69	17.35	8.40	220	1.25	2.50	1.25	1.25	1.25	1.25	203

Table 2 – Descriptive statistics for the variables used in the analysis

Variable	Obs	Mean	Std. Dev.	Min	Max
Quantity of staple crop sold (value, log)	2376	5.80	1.51	-0.69	11.94
The household sold staple crops (1: yes, 0: no)	6149	0.47	0.50	0.00	1.00
Maize price (median per district/province, log)	6149	1.22	0.51	-1.46	3.30
Tobacco price (median per district/province, log)	6149	2.29	1.21	-2.08	4.09
Household equivalent income (Meticais, log)	6082	7.49	1.27	-0.14	12.99
Owned land per adult equivalent (ha, log)	5983	-0.62	0.81	-6.68	4.13
Age (log)	6147	3.73	0.35	2.71	4.60
Education years (1-5)	6149	0.48	0.50	0.00	1.00
Education years (6+)	6149	0.14	0.34	0.00	1.00
Female household head (1: yes, 0: no)	6149	0.25	0.43	0.00	1.00
Number of equivalent household members (log)	6149	0.91	0.40	0.00	3.05
Quantity of maize produced (median per province, Kg)	6149	133.62	123.17	4.88	400.20
Quantity of tobacco produced (median per province, Kg)	6149	214.17	216.83	0.00	641.52
The household produced cash crops (1: yes, 0: no)	6149	0.23	0.42	0.00	1.00
The household used fertiliser (1: yes, 0: no)	6149	0.04	0.19	0.00	1.00
Daily agricultural wage (median per district, log)	6149	1.45	0.40	0.64	2.85
Livestock owned (TLUs, log)	4617	-1.62	1.69	-4.61	4.29
The household lives in the north (1: yes, 0: no)	6149	0.38	0.49	0.00	1.00
The household lives in the centre (1: yes, 0: no)	6149	0.43	0.50	0.00	1.00
Someone in the household has a paid job (1: yes, 0: no)	6149	0.34	0.47	0.00	1.00
The household has improved granaries (1: yes, 0: no)	6149	0.13	0.33	0.00	1.00
The household used temporary or full-time agricultural workers (1: yes, 0: no)	6149	0.18	0.39	0.00	1.00
The household owns a radio (1: yes, 0: no)	6149	0.54	0.50	0.00	1.00
Proportion of healthy household members (household members without serious illnesses divided by total household members)	6149	0.95	0.13	0.00	1.00
The community where the household lives experienced a flood (1: yes, 0: no)	6149	0.02	0.14	0.00	1.00
The community where the household lives experienced a drought (1: yes, 0: no)	6149	0.88	0.32	0.00	1.00
Proportion of months without available reserves of the most important staple crop for household's food safety	6149	0.49	0.38	0.00	1.00
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	6149	0.62	0.49	0.00	1.00
The household produced horticultural crops (1: yes, 0: no)	6149	0.40	0.49	0.00	1.00
Number of meals the household could afford during the past hungry season	6142	1.98	0.59	0.00	5.00
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	6149	0.82	0.38	0.00	1.00
The household received information about agricultural prices (1: yes, 0: no)	6149	0.40	0.49	0.00	1.00
The household received information about agricultural from extension agents (1: yes, 0: no)	6149	0.15	0.35	0.00	1.00
The household belongs to an agricultural association (1: yes, 0: no)	6149	0.06	0.24	0.00	1.00

Table 3 – Heckman and double-hurdle estimation results, national level

Variable	Heckman				Double-hurdle			
	Quantity of staple crop sold (value, log)		**The household sold staple crops (1: yes, 0: no)**		**Quantity of staple crop sold (value, log)**		**The household sold staple crops (1: yes, 0: no)**	
Maize price (median per district/province, log)	0.23***	(0.08)	0.00	(0.06)	0.22***	(0.07)	0.00	(0.05)
Tobacco price (median per district/province, log)	0.05	(0.05)	0.03	(0.06)	0.05	(0.04)	0.03	(0.04)
Household equivalent income (Meticais, log)	0.51***	(0.04)	0.17***	(0.03)	0.51***	(0.04)	0.17***	(0.02)
Owned land per adult equivalent (ha, log)	0.25***	(0.07)	0.12***	(0.04)	0.25***	(0.06)	0.12***	(0.04)
Age of the household head (log)	-0.28**	(0.12)	-0.12	(0.09)	-0.28**	(0.12)	-0.12	(0.08)
Education years (1-5)	-0.05	(0.08)	0.08	(0.06)	-0.06	(0.08)	0.08	(0.06)
Education years (6+)	0.00	(0.13)	-0.12	(0.09)	0.00	(0.13)	-0.12	(0.09)
Female household head (1: yes, 0: no)	-0.11	(0.10)	0.00	(0.07)	-0.11	(0.10)	0.00	(0.07)
Number of equivalent household members (log)	0.52***	(0.12)	-0.07	(0.08)	0.52***	(0.12)	-0.07	(0.08)
Quantity of maize produced (median per province, Kg)	0.00*	(0.00)	-0.00***	(0.00)	0.00*	(0.00)	-0.00***	(0.00)
Quantity of tobacco produced (median per province, Kg)	-0.00***	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)
The household produced cash crops (1: yes, 0: no)	-0.28***	(0.08)	-0.12	(0.07)	-0.28***	(0.08)	-0.12*	(0.06)
The household used fertiliser (1: yes, 0: no)	0.01	(0.15)	0.33**	(0.13)	-0.01	(0.15)	0.33***	(0.13)
Daily agricultural wage (median per district, log)	-0.22	(0.14)	-0.19**	(0.09)	-0.23*	(0.12)	-0.19***	(0.07)
Livestock owned (TLUs, log)	0.04	(0.02)	-0.02	(0.02)	0.04	(0.02)	-0.02	(0.02)
The household lives in the north (1: yes, 0: no)	0.70***	(0.18)	0.83***	(0.15)	0.70***	(0.16)	0.83***	(0.12)
The household lives in the centre (1: yes, 0: no)	0.23	(0.16)	0.67***	(0.12)	0.22	(0.15)	0.67***	(0.09)
Someone in the household has a paid job (1: yes, 0: no)	-0.35***	(0.08)	-0.22***	(0.06)	-0.35***	(0.08)	-0.22***	(0.06)
The household has improved granaries (1: yes, 0: no)	0.39***	(0.08)	-0.09	(0.08)	0.39***	(0.08)	-0.09	(0.07)
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.26***	(0.09)	0.08	(0.06)	0.26***	(0.09)	0.08	(0.07)
The household owns a radio (1: yes, 0: no)	0.10	(0.08)	0.13**	(0.06)	0.10	(0.08)	0.13**	(0.06)
Proportion of healthy household members (household members without serious illnesses/total household members)	0.32	(0.34)	0.02	(0.24)	0.32	(0.33)	0.02	(0.24)
The community where the household lives experienced a flood (1: yes, 0: no)	-0.27	(0.25)	0.13	(0.17)	-0.26	(0.25)	0.13	(0.17)
The community where the household lives experienced a drought (1: yes, 0: no)	-0.22*	(0.12)	-0.06	(0.09)	-0.21*	(0.12)	-0.06	(0.09)
Proportion of months without available reserves of the most important staple crop for household's food safety	-0.43***	(0.13)	-0.61***	(0.08)	-0.43***	(0.13)	-0.61***	(0.08)
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	-0.12	(0.08)	-0.15***	(0.06)	-0.11	(0.08)	-0.15***	(0.06)
The household produced horticultural crops (1: yes, 0: no)	-0.21***	(0.08)	0.05	(0.06)	-0.21***	(0.08)	0.05	(0.05)
Number of meals the household could afford during the past hungry season	0.24***	(0.08)	0.10**	(0.05)	0.25***	(0.07)	0.10**	(0.05)
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.42***	(0.13)	0.31***	(0.08)	0.42***	(0.13)	0.31***	(0.08)
The household received information about agricultural prices (1: yes, 0: no)			0.13**	(0.06)	-0.04	(0.07)	0.12**	(0.06)
The household received information about agricultural from extension agents (1: yes, 0: no)			0.12	(0.08)	0.02	(0.09)	0.12*	(0.07)
The household belongs to an agricultural association (1: yes, 0: no)			0.02	(0.10)	0.08	(0.12)	0.03	(0.10)
Constant	1.53**	(0.77)	-0.94*	(0.56)	1.55**	(0.73)	-0.94*	(0.50)
athrho			0.16**	(0.07)			0.15***	(0.06)
lnsigma			0.21***	(0.02)			0.21***	(0.02)
Observations	4,406				4,406			
Censored observations	2593				2593			
F	19.22***							
Chi^2					600.40***			

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 – Quantile regression estimation results, national level

	Quantile regression									
	Quantity of staple crop sold (value, log)									
Variables	**0.1**		**0.25**		**0.5**		**0.75**		**0.9**	
Maize price (median per district/province, log)	0.20	(0.13)	0.26***	(0.07)	0.27***	(0.05)	0.14***	(0.05)	0.17***	(0.05)
Tobacco price (median per district/province, log)	-0.02	(0.08)	0.08*	(0.05)	0.03	(0.04)	0.06*	(0.03)	-0.04	(0.03)
Household equivalent income (Meticais, log)	0.57***	(0.08)	0.46***	(0.04)	0.41***	(0.03)	0.51***	(0.03)	0.61***	(0.02)
Owned land per adult equivalent (ha, log)	0.36***	(0.10)	0.26***	(0.05)	0.30***	(0.04)	0.21***	(0.04)	0.20***	(0.04)
Age of the household head (log)	-0.81***	(0.22)	-0.44***	(0.12)	-0.12	(0.09)	0.07	(0.08)	0.10	(0.07)
Education years (1-5)	-0.18	(0.15)	-0.07	(0.08)	0.07	(0.07)	-0.07	(0.06)	-0.01	(0.05)
Education years (6+)	-0.40*	(0.23)	-0.09	(0.13)	0.22**	(0.10)	0.17*	(0.10)	0.09	(0.08)
Female household head (1: yes, 0: no)	-0.46**	(0.19)	-0.08	(0.10)	0.00	(0.08)	-0.09	(0.08)	-0.02	(0.07)
Number of equivalent household members (log)	0.62***	(0.23)	0.58***	(0.11)	0.51***	(0.09)	0.60***	(0.09)	0.52***	(0.08)
Quantity of maize produced (median per province, Kg)	0.00*	(0.00)	0.00***	(0.00)	0.00***	(0.00)	0.00***	(0.00)	0.00***	(0.00)
Quantity of tobacco produced (median per province, Kg)	-0.00	(0.00)	-0.00**	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)
The household produced cash crops (1: yes, 0: no)	-0.40***	(0.15)	-0.34***	(0.08)	-0.25***	(0.06)	-0.26***	(0.07)	-0.17***	(0.05)
The household used fertiliser (1: yes, 0: no)	-0.40*	(0.23)	-0.23	(0.15)	-0.13	(0.12)	0.05	(0.12)	-0.08	(0.10)
Daily agricultural wage (median per district, log)	-0.28	(0.18)	-0.44***	(0.11)	-0.29***	(0.09)	-0.14	(0.09)	-0.10	(0.08)
Livestock owned (TLUs, log)	-0.01	(0.04)	0.01	(0.02)	0.08***	(0.02)	0.04**	(0.02)	0.05***	(0.02)
The household lives in the north (1: yes, 0: no)	0.72***	(0.27)	0.79***	(0.16)	0.82***	(0.12)	0.57***	(0.12)	0.44***	(0.11)
The household lives in the centre (1: yes, 0: no)	0.25	(0.26)	0.28**	(0.14)	0.35***	(0.11)	0.05	(0.10)	-0.08	(0.09)
Someone in the household has a paid job (1: yes, 0: no)	0.03	(0.15)	-0.36***	(0.08)	-0.32***	(0.06)	-0.40***	(0.06)	-0.44***	(0.05)
The household has improved granaries (1: yes, 0: no)	0.83***	(0.15)	0.50***	(0.09)	0.20**	(0.08)	0.34***	(0.08)	0.21***	(0.06)
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.10	(0.15)	0.27***	(0.09)	0.30***	(0.07)	0.34***	(0.07)	0.15***	(0.05)
The household owns a radio (1: yes, 0: no)	0.09	(0.14)	0.06	(0.08)	0.01	(0.06)	0.06	(0.06)	0.13**	(0.06)
Proportion of healthy household members (household members without serious illnesses/total household members)	0.58	(0.48)	0.76***	(0.28)	0.76***	(0.25)	0.05	(0.25)	-0.56***	(0.20)
The community where the household lives experienced a flood (1: yes, 0: no)	0.10	(0.35)	-0.57**	(0.24)	-0.88***	(0.18)	-0.21	(0.19)	0.36**	(0.17)
The community where the household lives experienced a drought (1: yes, 0: no)	-0.20	(0.20)	-0.25**	(0.12)	-0.16*	(0.09)	-0.21**	(0.10)	-0.08	(0.09)
Proportion of months without available reserves of the most important staple crop for household's food safety	-0.19	(0.23)	-0.24**	(0.12)	-0.37***	(0.09)	-0.21**	(0.09)	-0.47***	(0.09)
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	-0.07	(0.14)	-0.27***	(0.08)	-0.13**	(0.06)	-0.11*	(0.06)	0.07	(0.05)
The household produced horticultural crops (1: yes, 0: no)	-0.26*	(0.15)	-0.25***	(0.07)	-0.22***	(0.06)	-0.24***	(0.06)	-0.14***	(0.05)
Number of meals the household could afford during the past hungry season	0.44***	(0.14)	0.34***	(0.06)	0.16***	(0.05)	0.07	(0.05)	0.05	(0.04)
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.64***	(0.21)	0.40***	(0.11)	0.26***	(0.09)	0.28***	(0.09)	0.19**	(0.08)
Constant	0.60	(1.32)	1.38*	(0.71)	1.85***	(0.56)	2.02***	(0.52)	2.54***	(0.41)
Observations	1,813		1,813		1,813		1,813		1,813	
Pseudo R^2	15.87		15.14		17.03		20.03		25.34	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5 – Heckman estimation results, regional level

Variable	North				Centre				Centre-North				South			
	Heckman				Heckman				Heckman				Heckman			
	Quantity of staple crop sold (value, log)		The household sold staple crops (1: yes, 0: no)		Quantity of staple crop sold (value, log)		The household sold staple crops (1: yes, 0: no)		Quantity of staple crop sold (value, log)		The household sold staple crops (1: yes, 0: no)		Quantity of staple crop sold (value, log)		The household sold staple crops (1: yes, 0: no)	
Maize price (median per district/province, log)	0.05	(0.12)	-0.22	(0.16)	0.33***	(0.11)	-0.04	(0.09)	0.20**	(0.08)	-0.08	(0.07)	0.09	(0.25)	0.41**	(0.16)
Tobacco price (median per district/province, log)	0.03	(0.07)	0.06	(0.08)	0.18**	(0.07)	-0.21***	(0.08)	0.07	(0.05)	0.03	(0.06)	-0.24	(0.21)	-0.24*	(0.13)
Household equivalent income (Meticais, log)	0.62***	(0.07)	0.10*	(0.05)	0.53***	(0.07)	0.23***	(0.04)	0.55***	(0.05)	0.18***	(0.03)	0.30***	(0.09)	0.17***	(0.05)
Owned land per adult equivalent (ha, log)	0.18*	(0.10)	0.12	(0.07)	0.46***	(0.10)	0.16***	(0.06)	0.30***	(0.07)	0.13***	(0.05)	0.11	(0.12)	0.09	(0.07)
Age of the household head (log)	-0.16	(0.17)	-0.34**	(0.15)	-0.57***	(0.20)	0.03	(0.13)	-0.38***	(0.14)	-0.16	(0.10)	0.59**	(0.28)	0.12	(0.21)
Education years (1-5)	-0.06	(0.12)	-0.03	(0.10)	-0.01	(0.13)	0.08	(0.10)	-0.04	(0.09)	0.05	(0.07)	-0.23	(0.23)	0.16	(0.12)
Education years (6+)	-0.23	(0.25)	-0.42***	(0.15)	0.19	(0.17)	-0.01	(0.13)	0.01	(0.14)	-0.17*	(0.10)	0.32	(0.34)	0.05	(0.20)
Female household head (1: yes, 0: no)	-0.27	(0.18)	-0.16	(0.13)	0.07	(0.15)	0.20*	(0.11)	-0.11	(0.12)	0.01	(0.09)	-0.13	(0.20)	0.00	(0.13)
Number of equivalent household members (log)	0.54**	(0.22)	0.09	(0.15)	0.76***	(0.15)	0.01	(0.13)	0.63***	(0.13)	0.01	(0.10)	0.23	(0.22)	-0.21	(0.14)
Quantity of maize produced (median per province, Kg)	0.00	(0.00)	-0.00***	(0.00)	-0.01*	(0.00)	-0.01***	(0.00)	0.00**	(0.00)	-0.00**	(0.00)	0.03***	(0.01)	-0.01*	(0.01)
Quantity of tobacco produced (median per province, Kg)	-0.00***	(0.00)	-0.00***	(0.00)	0.00*	(0.00)	0.00***	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)	0.01	(0.01)	-0.02	(0.01)
The household produced cash crops (1: yes, 0: no)	-0.40***	(0.14)	-0.28**	(0.13)	-0.26**	(0.12)	-0.08	(0.09)	-0.33***	(0.09)	-0.18**	(0.08)	-0.17	(0.28)	0.50**	(0.20)
The household used fertiliser (1: yes, 0: no)	-0.14	(0.22)	0.10	(0.21)	-0.24	(0.25)	0.56***	(0.19)	-0.21	(0.17)	0.42***	(0.14)	0.95***	(0.34)	-0.16	(0.42)
Daily agricultural wage (median per district, log)					1.97**	(0.95)	1.52*	(0.87)	-0.59***	(0.23)	-0.40**	(0.20)				
Livestock owned (TLUs, log)	0.01	(0.04)	-0.02	(0.04)	0.09***	(0.03)	-0.00	(0.03)	0.05*	(0.03)	-0.02	(0.02)	-0.02	(0.06)	-0.01	(0.03)
The household lives in the north (1: yes, 0: no)									0.53***	(0.14)	0.19	(0.13)				
Someone in the household has a paid job (1: yes, 0: no)	-0.40***	(0.13)	-0.11	(0.11)	-0.35***	(0.12)	-0.27***	(0.09)	-0.41***	(0.09)	-0.21***	(0.07)	-0.12	(0.21)	-0.20*	(0.11)
The household has improved granaries (1: yes, 0: no)	0.46***	(0.15)	-0.27**	(0.13)	0.35***	(0.11)	0.02	(0.12)	0.43***	(0.09)	-0.07	(0.09)	-0.16	(0.22)	-0.11	(0.18)
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.33**	(0.14)	0.12	(0.12)	0.14	(0.12)	0.08	(0.10)	0.24**	(0.10)	0.09	(0.08)	0.20	(0.24)	-0.02	(0.13)
The household owns a radio (1: yes, 0: no)	0.04	(0.11)	0.15	(0.10)	0.12	(0.12)	0.14	(0.10)	0.09	(0.08)	0.15**	(0.07)	0.02	(0.20)	0.06	(0.13)
Proportion of healthy household members (household members without serious illnesses/total household members)	-0.08	(0.46)	-0.01	(0.41)	0.19	(0.62)	0.47	(0.42)	0.07	(0.36)	0.25	(0.29)	1.24*	(0.70)	-0.35	(0.40)
The community where the household lives experienced a flood (1: yes, 0: no)	-0.40	(0.35)	0.23	(0.28)	-0.03	(0.34)	0.22	(0.21)	-0.25	(0.24)	0.22	(0.17)	-0.17	(0.72)	-2.04***	(0.43)
The community where the household lives experienced a drought (1: yes, 0: no)	-0.10	(0.15)	-0.07	(0.11)	-0.59**	(0.24)	-0.14	(0.20)	-0.21	(0.13)	-0.09	(0.10)	0.24	(0.38)	0.30	(0.31)

Table 5 – Heckman estimation results, regional level (continued)

Variable	North		Centre		Centre-North		South	
	Heckman		Heckman		Heckman		Heckman	
	Quantity of staple crop sold (value, log)	**The household sold staple crops (1: yes, 0: no)**	**Quantity of staple crop sold (value, log)**	**The household sold staple crops (1: yes, 0: no)**	**Quantity of staple crop sold (value, log)**	**The household sold staple crops (1: yes, 0: no)**	**Quantity of staple crop sold (value, log)**	**The household sold staple crops (1: yes, 0: no)**
Proportion of months without available reserves of the most important staple crop for household's food safety	-0.32 (0.21)	-0.59*** (0.15)	-0.57*** (0.20)	-0.71*** (0.13)	-0.42*** (0.15)	-0.69*** (0.10)	-0.16 (0.28)	-0.24 (0.17)
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	-0.18 (0.13)	0.08 (0.11)	-0.10 (0.12)	-0.29*** (0.09)	-0.15 (0.09)	-0.13* (0.07)	0.43** (0.19)	-0.14 (0.12)
The household produced horticultural crops (1: yes, 0: no)	-0.10 (0.06)	0.05 (0.04)	-0.03 (0.03)	0.04 (0.03)	-0.07** (0.03)	0.01 (0.02)	-0.08 (0.06)	-0.05 (0.04)
Number of meals the household could afford during the past hungry season	0.25** (0.10)	0.03 (0.08)	0.17 (0.12)	0.08 (0.08)	0.21** (0.08)	0.06 (0.05)	0.42** (0.17)	0.25** (0.11)
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.51** (0.20)	0.49*** (0.16)	0.56*** (0.19)	0.33*** (0.12)	0.50*** (0.15)	0.40*** (0.10)	0.11 (0.22)	0.19 (0.15)
The household received information about agricultural prices (1: yes, 0: no)		0.12 (0.12)		0.09 (0.08)		0.07 (0.07)		0.28** (0.13)
The household received information about agricultural from extension agents (1: yes, 0: no)		0.03 (0.13)		0.09 (0.12)		0.05 (0.09)		0.54*** (0.17)
The household belongs to an agricultural association (1: yes, 0: no)		0.10 (0.17)		0.03 (0.18)		0.10 (0.12)		-0.03 (0.17)
Constant	1.33 (0.95)	1.24 (0.88)	-0.23 (1.81)	-2.35 (1.54)	2.48*** (0.84)	-0.00 (0.63)	-1.28 (2.24)	-1.97* (1.12)
athrho		0.13 (0.21)		0.19* (0.10)		0.10 (0.11)		-0.02 (0.31)
lnsigma		0.13*** (0.03)		0.19*** (0.04)		0.17*** (0.03)		0.26*** (0.05)
Observations		1,035		1,960		2,995		1,411

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6 – Descriptive statistics for the variables used in the analysis, disaggregated by poor and nonpoor

	Nonpoor	Poor	Full sample	
Variable	**Mean**	**Mean**	**Mean**	
Quantity of staple crop sold (value, log)	6.40	5.13	5.80	***
The household sold staple crops (1: yes, 0: no)	0.50	0.41	0.47	***
Maize price (median per district/province, log)	1.32	1.13	1.22	***
Tobacco price (median per district/province, log)	2.39	2.23	2.29	**
Owned land per adult equivalent (ha, log)	-0.49	-0.76	-0.62	***
Age (log)	3.72	3.73	3.73	
Education years (1-5)	0.48	0.47	0.48	
Education years (6+)	0.21	0.07	0.14	***
Female household head (1: yes, 0: no)	0.21	0.31	0.25	***
Number of equivalent household members (log)	0.93	0.90	0.91	**
Quantity of maize produced (median per province, Kg)	141.60	125.38	133.62	***
Quantity of tobacco produced (median per province, Kg)	209.07	210.37	214.17	
The household produced cash crops (1: yes, 0: no)	0.27	0.17	0.23	***
The household used fertiliser (1: yes, 0: no)	0.06	0.01	0.04	***
Daily agricultural wage (median per district, log)	1.50	1.42	1.45	***
Livestock owned (TLUs, log)	-1.32	-1.92	-1.62	***
The household lives in the north (1: yes, 0: no)	0.34	0.41	0.38	***
The household lives in the centre (1: yes, 0: no)	0.41	0.46	0.43	*
Someone in the household has a paid job (1: yes, 0: no)	0.42	0.26	0.34	***
The household has improved granaries (1: yes, 0: no)	0.16	0.10	0.13	***
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.29	0.09	0.18	***
The household owns a radio (1: yes, 0: no)	0.67	0.41	0.54	***
Proportion of healthy household members (household members without serious illnesses/total household members)	0.96	0.95	0.95	***
The community where the household lives experienced a flood (1: yes, 0: no)	0.02	0.01	0.02	*
The community where the household lives experienced a drought (1: yes, 0: no)	0.87	0.89	0.88	*
Proportion of months without available reserves of the most important staple crop for household's food safety	0.47	0.53	0.49	***
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	0.62	0.61	0.62	
The household produced horticultural crops (1: yes, 0: no)	0.47	0.32	0.40	***
Number of meals the household could afford during the past hungry season	2.09	1.86	1.98	***
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.73	0.89	0.82	***
The household received information about agricultural prices (1: yes, 0: no)	0.46	0.33	0.40	***
The household received information about agricultural from extension agents (1: yes, 0: no)	0.18	0.11	0.15	***
The household belongs to an agricultural association (1: yes, 0: no)	0.08	0.04	0.06	***

Note: Are considered as poor households those that are at the bottom 40% of the income distribution, while nonpoor households those at the top 40%.

Appendix A

Table A.1 – Marginal effects (*ycond* and *psel*) for Heckman estimation, national level

	Heckman			
	Quantity of staple crop sold (value, log)		**The household sold staple crops (1: yes, 0: no)**	
Variables	**ycond**		**psel**	
Maize price (median per district/province, log)	0.23***	(0.08)	0	(0.03)
Tobacco price (median per district/province, log)	0.04	(0.05)	0.01	(0.03)
Household equivalent income (Meticais, log)	0.49***	(0.04)	0.07***	(0.01)
Owned land per adult equivalent (ha, log)	0.24***	(0.06)	0.05***	(0.02)
Age of the household head (log)	-0.26**	(0.12)	-0.05	(0.04)
Education years (1-5)	-0.06	(0.08)	0.03	(0.02)
Education years (6+)	0.02	(0.13)	-0.05	(0.03)
Female household head (1: yes, 0: no)	-0.11	(0.10)	0	(0.03)
Number of equivalent household members (log)	0.53***	(0.11)	-0.03	(0.03)
Quantity of maize produced (median per province, Kg)	0.00**	(0.00)	-0.00***	(0.00)
Quantity of tobacco produced (median per province, Kg)	-0.00***	(0.00)	-0.00***	(0.00)
The household produced cash crops (1: yes, 0: no)	-0.27***	(0.08)	-0.05	(0.03)
The household used fertiliser (1: yes, 0: no)	-0.03	(0.15)	0.13**	(0.05)
Daily agricultural wage (median per district, log)	-0.2	(0.14)	-0.08**	(0.04)
Livestock owned (TLUs, log)	0.04*	(0.02)	-0.01	(0.01)
The household lives in the north (1: yes, 0: no)	0.60***	(0.17)	0.33***	(0.06)
The household lives in the centre (1: yes, 0: no)	0.15	(0.15)	0.27***	(0.05)
Someone in the household has a paid job (1: yes, 0: no)	-0.32***	(0.08)	-0.09***	(0.02)
The household has improved granaries (1: yes, 0: no)	0.40***	(0.08)	-0.04	(0.03)
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.25***	(0.09)	0.03	(0.03)
The household owns a radio (1: yes, 0: no)	0.08	(0.08)	0.05**	(0.02)
Proportion of healthy household members (household members without serious illnesses/total household members)	0.31	(0.33)	0.01	(0.10)
The community where the household lives experienced a flood (1: yes, 0: no)	-0.29	(0.24)	0.05	(0.07)
The community where the household lives experienced a drought (1: yes, 0: no)	-0.21*	(0.12)	-0.02	(0.04)
Proportion of months without available reserves of the most important staple crop for household's food safety	-0.36***	(0.12)	-0.24***	(0.03)
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	-0.1	(0.08)	-0.06***	(0.02)
The household produced horticultural crops (1: yes, 0: no)	-0.22***	(0.08)	0.02	(0.02)
Number of meals the household could afford during the past hungry season	0.23***	(0.08)	0.04**	(0.02)
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.38***	(0.13)	0.12***	(0.03)
The household received information about agricultural prices (1: yes, 0: no)	-0.02	(0.01)	0.05**	(0.02)
The household received information about agricultural from extension agents (1: yes, 0: no)	-0.01	(0.01)	0.05	(0.03)
The household belongs to an agricultural association (1: yes, 0: no)	0	(0.01)	0.01	(0.04)
Observations	4,406	4,406		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.2 – Marginal effects (*ycond*) for Heckman estimation, regional level

	Centre – North				**South**			
	Heckman				**Heckman**			
	Quantity of staple crop sold (value, log)		**Quantity of staple crop sold (value, log)**		**Quantity of staple crop sold (value, log)**		**Quantity of staple crop sold (value, log)**	
	ycond		***ycond***		***ycond***		***ycond***	
Maize price (median per district/province, log)	0.308***	(0.0735)	0.206**	(0.0812)	0.05	(0.449)	0.11	(0.234)
Tobacco price (median per district/province, log)	0.10	(0.0992)	0.06	(0.0502)			-0.25	(0.198)
Household equivalent income (Meticais, log)	0.584***	(0.0515)	0.543***	(0.0507)	0.347***	(0.0777)	0.304***	(0.0815)
Owned land per adult equivalent (ha, log)	0.226***	(0.0867)	0.288***	(0.0726)	0.13	(0.147)	0.11	(0.118)
Age of the household head (log)	-0.358***	(0.132)	-0.372***	(0.134)	0.656**	(0.291)	0.585**	(0.273)
Education years (1-5)	0.00	(0.0795)	-0.05	(0.0867)	-0.20	(0.213)	-0.21	(0.204)
Education years (6+)	-0.05	(0.128)	0.01	(0.140)	0.42	(0.365)	0.33	(0.342)
Female household head (1: yes, 0: no)	-0.07	(0.116)	-0.10	(0.118)	-0.04	(0.192)	-0.11	(0.199)
Number of equivalent household members (log)	0.642***	(0.134)	0.632***	(0.135)	0.19	(0.217)	0.22	(0.206)
Quantity of maize produced (median per province, Kg)	0.01	(0.00527)	0.00186***	(0.000663)			0.0277***	(0.00969)
Quantity of tobacco produced (median per province, Kg)	0.00	(0.00309)	-0.000942***	(0.000301)			0.01	(0.0132)
The household produced cash crops (1: yes, 0: no)	-0.224**	(0.0889)	-0.305***	(0.0863)	-0.13	(0.277)	-0.19	(0.256)
The household used fertiliser (1: yes, 0: no)	-0.06	(0.176)	-0.23	(0.168)	0.07	(0.392)	0.785***	(0.290)
Daily agricultural wage (median per district, log)	-1.62	(1.940)	-0.558**	(0.229)				
Livestock owned (TLUs, log)	0.0458*	(0.0239)	0.0523**	(0.0265)	0.00	(0.0652)	-0.03	(0.0611)
The household lives in the North (1: yes, 0: no)	0.62	(1.497)	0.524***	(0.140)				
Someone in the household has a paid job (1: yes, 0: no)	-0.395***	(0.0807)	-0.391***	(0.0864)	-0.19	(0.205)	-0.12	(0.211)
The household has improved granaries (1: yes, 0: no)	0.328***	(0.0939)	0.428***	(0.0863)	-0.27	(0.262)	-0.15	(0.215)
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.210**	(0.0922)	0.226**	(0.0964)	0.17	(0.245)	0.18	(0.252)
The household owns a radio (1: yes, 0: no)	0.11	(0.0817)	0.08	(0.0821)	-0.10	(0.182)	0.01	(0.202)
Proportion of healthy household members (household members without serious illnesses/total household members)	-0.22	(0.334)	0.07	(0.365)	1.18	(0.718)	1.21	(0.743)
The community where the household lives experienced a flood (1: yes, 0: no)	-0.25	(0.232)	-0.29	(0.238)	-0.27	(0.681)	-0.16	(0.544)
The community where the household lives experienced a drought (1: yes, 0: no)	-0.17	(0.114)	-0.20	(0.129)	0.45	(0.394)	0.23	(0.398)
Proportion of months without available reserves of the most important staple crop for household's food safety	-0.233*	(0.130)	-0.365***	(0.137)	-0.05	(0.254)	-0.15	(0.275)
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	-0.173**	(0.0831)	-0.15	(0.0890)	0.335*	(0.192)	0.437**	(0.195)
The household produced horticultural crops (1: yes, 0: no)	-0.246***	(0.0861)	-0.270***	(0.0862)	-0.02	(0.168)	-0.14	(0.219)
Number of meals the household could afford during the past hungry season	0.196***	(0.0755)	0.202**	(0.0836)	0.457***	(0.153)	0.424***	(0.150)
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.404***	(0.135)	0.462***	(0.150)	0.08	(0.205)	0.11	(0.208)
The household received information about agricultural prices (1: yes, 0: no)	-0.01	(0.0120)	-0.01	(0.00836)	0.236**	(0.112)	0.01	(0.117)
The household received information about agricultural from extension agents (1: yes, 0: no)	0.00	(0.00964)	0.00	(0.00764)	0.464***	(0.172)	0.02	(0.228)
The household belongs to an agricultural association (1: yes, 0: no)	-0.01	(0.0159)	-0.01	(0.0123)	-0.11	(0.157)	0.00	(0.0187)
Observations	1569		1569		1024		1024	
District dummies	Yes		No		Yes		No	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix B

In this Appendix, a robustness check on the relation between market participation and poverty is conducted. To this end, a wealth index is constructed using Principal Component Analysis (PCA) and polychoric correlations (Olsson, 1979; Kolenikov and Angeles, 2004; Kolenikov and Angeles, 2009).

The wealth index is based on available data on possession of durable goods, productive tools, and education. Principal component analysis (PCA) enables normalisation of the way in which assets are measured and extraction of the underlying orthogonal linear combinations that capture the maximum amount of information common to all such variables. However, it would be incorrect to apply PCA with dummy or categorical variables, unless special attention is paid to these variables. Since most of the variables we use for computing the wealth index are binary or categorical, we make use of polychoric correlations in order to correctly compute the correlation coefficients among different kinds of variables - continuous, binary and categorical (Kolenikov and Angeles, 2004; Kolenikov and Angeles, 2009)[22]. Fifteen items for which we had available information were selected from the TIA 2005 survey. They are displayed in Table B.1, and include possession of durable goods (bicycles, radios, tables, or lanterns); access to or ownership of productive farm tools (land per adult equivalent, ploughs, scythes, improved granaries, or livestock); housing characteristics (latrines, high-quality roofs, or high-quality walls); household characteristics related to work, education, and vulnerability (household members with a paid job, years of education of the head of household, or whether the household had experiences difficulties in providing enough food for all its members in the previous twelve months).

After computing the correlation matrix for the selected items using polychoric correlations, we ran the PCA. Since the PCA was run on the correlation matrix of the original data, we did not need to normalise the data (Kolenikov and Angeles, 2004).

In our subsequent analysis, only the first principal component was considered as representing our wealth index: this component explains about 27.9% of the total variance, and households with negative values for this component are considered to be poor (56% of total households). Figure B.1 shows the distribution (panel a) and the cumulative distribution (panel b) of the wealth index using kernel density (with

[22] Simulations showed that using PCA with polychoric correlations gives results that are superior in many ways to standard PCA analyses, which ignore the problem of non-normality of dummy or categorical variables (Kolenikov and Angeles, 2004).

Epanechnikov kernel). The distribution is positively skewed with a mean equal to 0 and median equal to -0.21.

Descriptive statistics for the variables used in previous estimations are presented in Table B.2, disaggregated for wealth between poor and non-poor. Asterisks are displayed in the usual fashion when the mean for the two groups is significantly different.

As already evidenced in the main text when income was considered (Table 6), most of the variables differ considerably between poorer and wealthier households. Table B.2 confirms that poorer households are generally less-educated, headed by women, live in the northern and central regions, and sell fewer staple crops at a lower price. Moreover, by running separate regressions for the bottom 40% and upper 40% of the wealth index distribution, we obtain the result that the price of maize has a positive and significant effect only for the latter group (*p-value* < 0.05). Again, most poorer households appear to engage in agriculture/husbandry as their main activity, but they receive considerably less information from extension agents or information on prices.

Here, in the same manner as in Section 6 for income and other assets, we present some qualitative analyses based on the wealth index just described. Plotted in Figure B.2 (panel a), the result of a local polynomial smoothing (Epanechnikov kernel, bandwidth 0.8) of the –actual - (log) value of staple crops sold on the wealth index. As expected, the result is an increasing curve, also in the case of wealth. However, in this case, the curve increases more steeply at low levels of wealth, while it remains more stable at higher levels of wealth. This might suggest that increasing asset endowments or wealth up to a certain threshold could significantly increase the quantity sold on the market, while the effect would taper off substantially for wealthier households. Another local polynomial smoothing plot for the relation between the wealth index and household income is also presented (Figure B.2, panel b). They are positively correlated, though the correlation coefficient is not especially high (0.36).

Figure B.1 – Distribution and cumulative distribution of the wealth index

a b

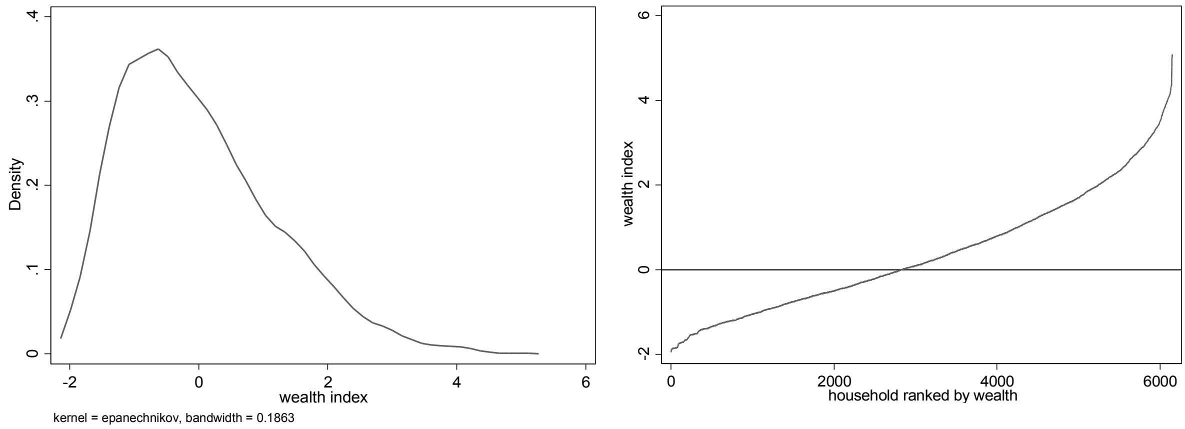


Figure B.2 – Local polynomial smoothing of the quantity of staple crops sold on wealth and correlation between household equivalent income and the wealth index

a b

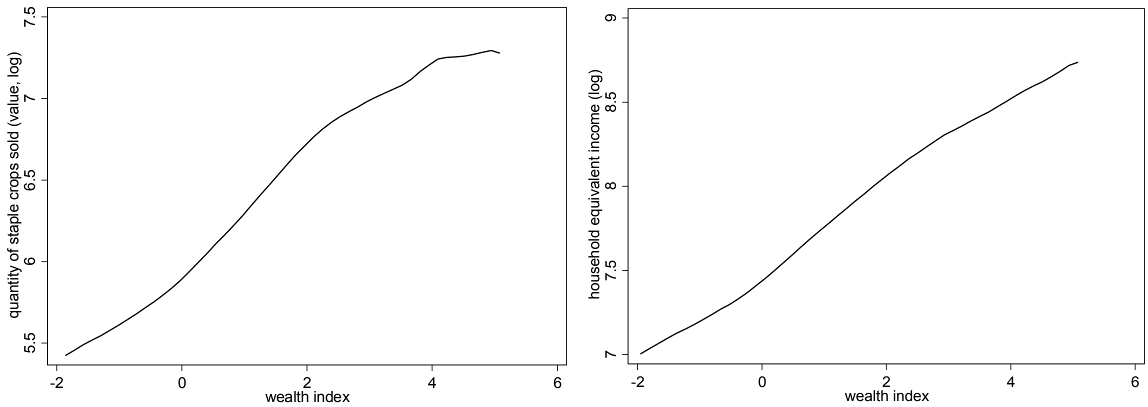


Table B.1

Variable	Obs	Mean	Std. Dev.	Min	Max
The household owns a bike (1: yes, 0: no)	6149	0.31	0.46	0	1
The household owns a plough (1: yes, 0: no)	6149	0.03	0.17	0	1
The household owns a scythe (1: yes, 0: no)	6149	0.35	0.48	0	1
The household owns a lantern (1: yes, 0: no)	6149	0.47	0.50	0	1
The household owns a radio (1: yes, 0: no)	6149	0.54	0.50	0	1
The household dwelling has a latrine (1: yes, 0: no)	6149	0.42	0.49	0	1
The household dwelling has a table (1: yes, 0: no)	6149	0.35	0.48	0	1
The household owns improved granaries (1: yes, 0: no)	6149	0.13	0.33	0	1
The household dwelling has a high-quality roof (1: yes, 0: no)	6149	0.16	0.36	0	1
The household dwelling has high-quality walls (1: yes, 0: no)	6149	0.06	0.23	0	1
Someone in the household has a paid job (1: yes, 0: no)	6149	0.34	0.47	0	1
Household head education years	6149	2.53	2.60	0	13
The household had difficulties in providing enough food to all members in the last 12 months (1: no, 0: yes)	6149	0.62	0.49	0	1
Owned land per adult equivalent (ha, log)	6149	-0.79	1.30	-6.68	4.13
Livestock owned (TLUs) (0: none, 1: less or equal than 0.5 TLUs, 2: less or equal than 2 TLUs, 3: more than 2 TLUs)	6149	0.95	0.83	0	3

Table B.2

Variable	Nonpoor Mean	Poor Mean	Full sample Mean	
Quantity of staple crop sold (value, log)	6.21	5.49	5.80	***
The household sold staple crops (1: yes, 0: no)	0.46	0.47	0.47	
Maize price (median per district/province, log)	1.28	1.17	1.22	***
Tobacco price (median per district/province, log)	2.37	2.23	2.29	**
Household equivalent income (Meticais, log)	7.94	7.13	7.49	***
Owned land per adult equivalent (ha, log)	-0.57	-0.65	-0.62	***
Age (log)	3.73	3.72	3.73	
Education years (1-5)	0.54	0.44	0.48	***
Education years (6+)	0.25	0.05	0.14	***
Female household head (1: yes, 0: no)	0.16	0.33	0.25	***
Number of equivalent household members (log)	1.04	0.80	0.91	***
Quantity of maize produced (median per province, Kg)	130.00	136.47	133.62	
Quantity of tobacco produced (median per province, Kg)	194.93	229.32	214.17	***
The household produced cash crops (1: yes, 0: no)	0.25	0.21	0.23	***
The household used fertiliser (1: yes, 0: no)	0.06	0.02	0.04	***
Daily agricultural wage (median per district, log)	1.46	1.45	1.45	
Livestock owned (TLUs, log)	-1.12	-2.17	-1.62	***
The household lives in the North (1: yes, 0: no)	0.33	0.42	0.38	***
The household lives in the Centre (1: yes, 0: no)	0.37	0.49	0.43	***
Someone in the household has a paid job (1: yes, 0: no)	0.42	0.28	0.34	***
The household has improved granaries (1: yes, 0: no)	0.18	0.09	0.13	***
The household used temporary or full-time agricultural workers (1: yes, 0: no)	0.30	0.09	0.18	***
The household owns a radio (1: yes, 0: no)	0.82	0.32	0.54	***
Proportion of healthy household members (household members without serious illnesses/total household members)	0.96	0.95	0.95	***
The community where the household lives experienced a flood (1: yes, 0: no)	0.02	0.02	0.02	
The community where the household lives experienced a drought (1: yes, 0: no)	0.88	0.88	0.88	
Proportion of months without available reserves of the most important staple crop for household's food safety	0.49	0.49	0.49	
The household bought maize or cassava during the past hungry season (1: yes, 0: no)	0.64	0.60	0.62	**
The household produced horticultural crops (1: yes, 0: no)	0.47	0.34	0.40	***
Number of meals the household could afford during the past hungry season	2.10	1.88	1.98	
Agriculture/husbandry is the main economic activity of the household (1: yes, 0: no)	0.73	0.89	0.82	***
The household received information about agricultural prices (1: yes, 0: no)	0.46	0.36	0.40	***
The household received information about agricultural from extension agents (1: yes, 0: no)	0.19	0.12	0.15	***
The household belongs to an agricultural association (1: yes, 0: no)	0.09	0.04	0.06	***

Multidimensional and Fuzzy Measures of Poverty and Inequality at National and Regional Level in Mozambique

Vincenzo Salvucci[a], Gianni Betti[a], Francesca Gagliardi[a]

[a] Dipartimento di Economia Politica e Statistica - Facoltà di Economia "Richard M. Goodwin"
Università degli Studi di Siena - Piazza S. Francesco, 7 - 53100 Siena, Italy

Abstract

This study provides a step-by-step account of how fuzzy measures of non-monetary deprivation and also monetary poverty may be constructed based on survey data such as those from the Mozambican Household Budget Survey 2008-09 (IOF08). For non-monetary deprivation, six dimensions are identified using explanatory and confirmatory factor analyses, and a weighting system is applied for the aggregation of individual items into the dimension they represent. Application to Mozambique is conducted using the Household Budget Survey 2008-09 (IOF08) data: estimates are provided at national level and also disaggregated at provincial and urban/rural level. Standard errors are provided using a recent methodology based on Jack-knife Repeated Replication.

Our results contrast with previous findings based solely on Head Count statistics and give a more complete mapping of poverty in Mozambique. Monetary and non-monetary deprivation seem to have very different distribution patterns, especially when analysed at sub-national level and by area of residence. Disaggregated Head Count statistics produce rankings of provinces and urban/rural areas that greatly differ from estimates based on non-monetary dimensions. In particular, the Northern and Central provinces suffer from non-monetary deprivation significantly more than the South, and the urban/rural deprivation gap widens in favour of urban areas when non-monetary dimensions are considered. Housing conditions and quality, and possession of less affordable durable goods emerge as the most unequally distributed non-monetary dimensions.

1. Introduction

This study provides a step-by-step account of how fuzzy measures of non-monetary deprivation and also monetary poverty may be constructed based on the Mozambican

Household Budget Survey 2008-09 (IOF08). To our knowledge, this is the first attempt to apply Fuzzy Set Theory to poverty measurement in Mozambique.

The dataset we use is the most recent budget survey available for Mozambique and is representative of the national, regional (North, Centre, South), provincial and urban/rural level.

In order to construct a Fuzzy Set index of poverty, monetary as well as non-monetary indicators are considered, and two different measures of deprivation are subsequently constructed: the Fuzzy Monetary (FM) and Fuzzy Supplementary (FS). For purposes of comparability with the official poverty analyses for Mozambique (INE, 2010; MPD-DNEAP, 2010), we use per capita (real) daily consumption as our monetary poverty indicator.

For non-monetary deprivation, we identify six dimensions, including housing conditions; more widespread and affordable durable goods; less common, more expensive durable goods; housing quality; income-related deprivation; health and education. The dimensions are identified using explanatory and confirmatory factor analyses, and a weighting system is applied for the aggregation of individual items into the dimension they represent.

In the study we present monetary and non-monetary poverty estimates at national, provincial and urban/rural level. Standard errors are provided using a recent methodology based on Jack-knife Repeated Replication (Verma and Betti, 2011). This methodology allows us to present reliable estimates, especially at sub-national level, where the sample size is smaller.

In particular, we compute more precise poverty estimates at provincial level and urban/rural level for each province, thus providing a finer poverty mapping for Mozambique.

We believe that including non-monetary dimensions in the analysis of poverty in Mozambique is important and informative for the most part of Mozambicans lives close to the poverty line, so that small changes in their income levels are likely to produce sensible modification in the Head Count statistics. For example, using consumption per adult equivalent instead of consumption per capita for the computation of the Head Count Ratio yields very different results. Poverty estimates based solely on Head Count statistics are thus not very robust. Moreover, the official poverty analyses based on Mozambican Budget Survey data generally produced strange or non-robust results at provincial level, with strong fluctuations in the Head Count Ratio and re-ranking of poor and rich provinces (Van den Boom, 2011; pp. 7-8).

The work proceeds as follows: in Section 2 we present previous studies and official statistics of poverty in Mozambique; Section 3 illustrates the concept of multidimensional poverty, as well as the Fuzzy Set technique and its application to poverty estimation. In Section 4 we introduce the dataset that is used throughout the study, while in Section 5 we set out the empirical analysis and the resulting poverty estimates. Section 6 concludes.

2. Poverty in Mozambique

Mozambique is among the poorest countries in the world, with a per capita income level of approximately $428, ranking 197 out of 210 countries (World Bank, 2010). After the end of the civil war in 1992, Mozambique underwent a process of sustained growth and poverty reduction that led the country to be considered as a success story by the World Bank and international donors (World Bank, 2008).

Nevertheless, poverty levels remain very high and poor living conditions are widespread throughout the country.

The process of poverty reduction has been deeply monitored and analysed by three official national assessments (MPF, 1998; MPF, 2004; MPD-DNEAP, 2010) and several other studies by both Mozambican and international analysts (Hanlon, 2007; Castel-Branco, 2010; Ossemane, 2010; Van den Boom, 2011).

What emerges from the three main household surveys conducted in the 1996-2008 period and from other field-specific surveys, is that Mozambican citizens substantially improved their situation with respect to some non-monetary dimensions: access to education and health services, household asset ownership, and quality of housing. On the other hand, monetary poverty remained fairly stable between 2002 and 2008: the Head Count Ratio slightly increased from a value of 54.1% in 2002-03 to 54.7% in 2008-09. However, it is important to note that this stabilisation followed a sharp fall from its previous (69%) levels in 1996-97.

Recent figures on poverty reduction can be found in the Third National Poverty Assessment (MPD-DNEAP, 2010), which is based on the results of the Mozambican Household Budget Survey 2008-09 (IOF08) and outlines an analysis of monetary poverty and non-monetary poverty.

As previously explained, monetary and non-monetary poverty are different phenomena where non-monetary indicators are recognised as dimensions that capture long-term poverty trends more than monetary ones. The Mozambican Government and international donors invested considerably in reducing non-monetary poverty. In particular, education and health are considered key intervention areas, and progressively more people have been granted access to schools and health facilities in urban as well as rural areas (Chao and Kostermans, 2002; Government of Mozambique, 2005; Republic of Mozambique, 2006).

Nonetheless, monetary poverty did not decrease between 2002 and 2008. The Third National Poverty Assessment lists different causes that contribute to maintain such a high level of monetary poverty, improvements in non-monetary dimensions notwithstanding. Most of them are related to the agricultural sector, since agriculture remains the main economic activity and source of income for most Mozambicans. In particular, MPD-DNEAP (2010) points out that the harvest of 2008 was affected by unexpected weather shocks, which reduced the amount of disposable food, especially for Central provinces. Moreover, households faced higher transportation and imported food costs due to the rise in international food and fuel prices. Such a high dependence from climate shocks and international prices contributed to the persistence of low agricultural productivity between 2002 and 2008, and to maintaining large part of the Mozambican population in a state of great vulnerability (MPD-DNEAP, 2010; pp. xii-xiii).

When the analysis is conducted at sub-national - provincial and urban/rural - levels, the pattern of poverty reduction becomes less clear. With specific regard to monetary poverty, between 2002 and 2008 the Southern provinces and some of the rural areas in the North experienced a sharp fall in their Head Count Ratio, while Central regions witnessed an increase. More precisely, monetary poverty decreased in 5 provinces (Niassa, from 52% to 32%; Cabo Delgado, from 63% to 37%; Tete, from 60% to 42%; Inhambane, from 81% to 58%; Maputo City, from 54% to 36%), remained stable in three provinces (Maputo Province, 68%; Nampula, 54%; Gaza, 61%), and increased in the remaining 3 provinces (Zambezia, from 45% to 71%; Manica, from 44% to 51%; Sofala, from 36% to 58%). Nationwide, rural poverty increased from 55.3% in 2002-03 to 56.9% in 2008-09, whereas urban poverty decreased from 51.5% to 49.6% in the same period.

For non-monetary dimensions, the Assessment compares the results for 2008-09 with those from 2002-03. Each of the three dimensions considered (housing conditions, ownership of durable goods, and access to public goods and services) is separately compared with the same dimension six years before but without computing a general composite welfare indicator. This approach we argue is less informative than the one used in the present study. In what follows we consider all the available information about durable goods, housing conditions, and health and education, by computing dimension-specific deprivation indexes plus a general indicator for non-monetary poverty. Moreover, we contend this official analysis to constitute a possible benchmark.

The results of the Assessment indicate that on average housing conditions improved between 2002 and 2008, though differences at sub-national level remain high nowadays. Ownership of durable goods also improved: the percentage of households owning a radio, a TV, a fridge, a mobile, a telephone, a car, and a bike or motorbike increased by 5.7 points. Turning to the access to public goods and services, it emerges that access to education peaked such that in 2008-09 more than 76% of all children aged 6-13 were attending school which reflects a big jump if compared to a figure of 66.8% in 2002-03. Moreover, geographic inequality in access to education decreased over time while access to health facilities improved. Though the average proportion of households at less than 45 minutes from the nearest health facility raised from 54.5% to 65.2%, it spiked in rural areas from 31.5% to 69.7% in the North, and from 35.0% to 47.6% in the Centre, while in the South it only increased from 48.8% to 53.6%.

At the same time, other non-monetary dimensions of deprivation did not improve in a substantial way: access to safe water and chronic malnutrition, for example, remained more or less stable[1].

1 Chronic malnutrition (stunting) is still suffered by 46.4% of under-five children, which is among the highest percentages in the world (WHO, 2011).

3. Multidimensional poverty and fuzzy set theory

In order to understand poverty and social exclusion, it is necessary to consider deprivation simultaneously in different terms (i.e. as low income as well as different non-monetary aspects of deprivation). The need to adopt a multidimensional approach has been noted, among others, by Kolm (1977); Atkinson and Bourguignon (1982); Maasoumi (1986); Tsui (1995); Sen (1999). Moreover, the multidimensional nature of poverty is a widely recognised fact, not only by the international scientific community, but also by many official statistical agencies (e.g. Eurostat, Istat) as well as by international institutions (United Nations, World Bank).

In the present work we go beyond the conventional study of poverty based simply on the poor/non-poor dichotomy defined in relation to a chosen poverty line. Instead, poverty and multidimensional deprivation are treated as matters of degree based on the individual's position in the distribution of income and other aspects of living condition. State of deprivation is thus seen in the form of 'fuzzy sets' to which all members of the population belong yet to varying degrees. This fact brings with it more complete and realistic view of the phenomenon but also an increased complexity at both the conceptual and the analytical levels.

A number of authors have applied the concepts of fuzzy sets to the analysis of poverty and living conditions (Chiappero Martinetti, 1994; Vero and Werquin, 1997, inter alia). Our application is based on the specific methodology developed by Cerioli and Zani (1990), Cheli (1995), Cheli and Lemmi (1995), Betti and Verma (1999), Betti et al. (2006).

Under the so-called traditional approach, poverty is characterized by a simple dichotomization of the population into poor and non poor defined in relation to a chosen poverty line, z. This approach presents two main limitations: firstly, it is unidimensional (i.e. it refers to only one proxy of poverty, namely low income or consumption expenditure), and secondly it reduces the population to a simple dichotomy. However, poverty is a much more complex phenomenon that is not formed solely of its monetary dimension but must also take account non-monetary indicators of living condition. Moreover it is not an attribute that characterises an individual as being either present or absent, but is rather a difficult to define predicate that manifests itself in different shades and degrees.

The fuzzy approach considers poverty as a matter of degree rather than an attribute that is simply present or absent for individuals in the population. In this case, two additional aspects have to be introduced:

i. The choice of membership functions (m.f.), i.e. quantitative specifications of individuals' or households' degrees of poverty and deprivation;
ii. The choice of rules for the manipulation of the resulting fuzzy sets.

The traditional approach can be seen as a special case of the fuzzy approach, where the membership function may be seen as $\mu_i^H = 1$ if $y_i < z$, $\mu_i^H = 0$ if $y_i \geq z$, where y_i is the income of individual i and z is the poverty line.

An early attempt to incorporate the concept of poverty as a matter of degree at methodological level was made by Cerioli and Zani (1990) who drew inspiration from the theory of *Fuzzy Sets* initiated by Zadeh (1965). Subsequently, Cheli and Lemmi (1995) proposed the so called Totally Fuzzy and Relative (TFR) approach in which the m.f. is defined as the distribution function $F(y_i)$ of income, normalised (linearly transformed) so as to equal 1 for the poorest and 0 for the richest person in the population.

3.1 Income poverty: the Fuzzy Monetary (FM) measure

In the present study we make use of a fuzzy monetary indicator as found in Betti et al. (2009). The proposed FM indicator is defined as a combination of the $(1-F_{(M),i})$ indicator, namely the proportion of individuals richer than individual i (Cheli and Lemmi, 1995), and of the $(1-L_{(M),i})$ indicator, namely the share of the total income received by all individuals richer than individual i (Betti and Verma, 1999). Formally:

$$\mu_i = FM_i = \left(1-F_{(M),i}\right)^{\alpha-1}\left(1-L_{(M),i}\right) = \left(\frac{\sum_{\gamma=i+1}^{n} w_\gamma | y_\gamma > y_i}{\sum_{\gamma=2}^{n} w_\gamma | y_\gamma > y_1}\right)^{\alpha-1}\left(\frac{\sum_{\gamma=i+1}^{n} w_\gamma y_\gamma | y_\gamma > y_i}{\sum_{\gamma=2}^{n} w_\gamma y_\gamma | y_\gamma > y_1}\right) \quad (1)$$

where y_γ is the income, $F_{(M),i}$ is the income distribution function, w_γ is the sample weight of individual of rank γ ($\gamma = 1,\ldots, n$) in the ascending income distribution, $L_{(M),i}$ represents the value of the Lorenz curve of income for individual i. The parameter α is estimated so that the overall FM indicator (which is calculated simply as the weighted mean of the individual FM_i), is equal to the Head Count Ratio computed for the official poverty line.

3.2 Non-monetary poverty: the Fuzzy Supplementary (FS) measure

In addition to the level of monetary income, the standard of living of households and individuals can be described by a host of indicators, such as housing conditions, possession of durable goods, health conditions, education, perception of hardship. To quantify and put together diverse indicators of deprivation several steps are necessary. In particular, decisions are required for assigning numerical values to the ordered categories, weighting the score to construct composite indicators, choosing their appropriate distributional form and scaling the resulting measures in a meaningful way.

First, from the large set which may be available, a selection has to be made of indicators which are substantively meaningful and worthwhile for the analysis of interest. Secondly, it is useful to group different indicators into statistical components (or dimensions) in order to reduce dimensionality. Whelan et al. (2001) suggest, as a first step in an analysis of life-style deprivation, to systematically examine the range of deprivation items to see whether the items cluster into distinct groups. Factor analysis can be used to identify such clusters of interrelated variables. To quantify and put together diverse indicators several steps are necessary.

1. Identification of items;
2. Transformation of the items into the [0, 1] interval;
3. Exploratory and confirmatory factor analysis;
4. Calculation of weights within each dimension (each group);
5. Calculation of scores for each dimension;
6. Calculation of an overall score and the parameter α;
7. Construction of the fuzzy deprivation measure in each dimension (and overall).

Aggregation over a group of items in a particular dimension h (h = 1, 2, …, m) is given by a weighted mean taken over j items: $s_{hi} = \frac{\sum w_{hj} \cdot s_{hj,i}}{\sum w_{hj}}$ where w_{hj} is the weight of the j-th deprivation variable in the h-th dimension. An overall score for the i-th individual is calculated as the unweighed mean:

$$s_i = \frac{\sum_{h=1}^{m} s_{hi}}{m} \quad (2)$$

Then, we calculate the FS indicator for the i-th individual over all dimensions as:

$$FS_i = (1-F_{(S),i})^{\alpha-1}(1-L_{(S),i}) \quad (3)$$

As for the FM indicator, the estimates of α is determined so as to make the overall non-monetary deprivation rate (which is calculated simply as the weighted mean of the individual FS_i) numerically identical to the Head Count Ratio computed for the official poverty line. The parameter α estimated is then used to calculate the FS indicator for each dimension of deprivation separately. The FS indicator for the h-th deprivation dimension and for the i-th individual is defined as combination of the $(1-F_{(S),hi})$ indicator and the $(1-L_{(S),hi})$ indicator.

$$\mu_{hi} = FS_{hi} = \left(1-F_{(S),hi}\right)^{\alpha-1}\left(1-L_{(S),hi}\right) = \left[\frac{\sum_{\gamma=i+1}^{n} w_{h\gamma} | s_{h\gamma} > s_{hi}}{\sum_{\gamma=2}^{n} w_{h\gamma} | s_{h\gamma} > s_{h1}}\right]^{\alpha-1} \left[\frac{\sum_{\gamma=i+1}^{n} w_{h\gamma} s_{h\gamma} | s_{h\gamma} > s_{hi}}{\sum_{\gamma=2}^{n} w_{h\gamma} s_{h\gamma} | s_{h\gamma} > s_{h1}}\right], \quad (4)$$

$h = 1, 2, \ldots, m;\ i = 1, 2, \ldots, n;\ \mu_{hn} = 0$

The $(1-F_{(S),hi})$ indicator for the i-th individual is the proportion of individuals who are less deprived in the h-th dimension than the individual studied. $F_{(S),hi}$ is the value of the score distribution function evaluated for individual i in dimension h and $w_{h\gamma}$ is the sample weight of the i-th individual of rank γ in the ascending score distribution in the h-th dimension.

The $(1-L_{(S),hi})$ indicator is the share of the total lack of deprivation score assigned to all individuals less deprived than individual i. $L_{(S),hi}$ is the value of the Lorenz curve of score in the h-th dimension for the i-th individual.

As for the Fuzzy Monetary and the Fuzzy Supplementary indicators, the overall index corresponding to each dimension, FS_h, is calculated simply as the weighted mean of the individual FS_{hi}. Here it is interesting to note that the overall ranking of the FS indicator cannot directly be obtained from the rankings in each dimension, albeit the ranking obtained with FS_i being consistent with the ranking obtained from FS_{hi} [2].

2 A possible alternative definition of the overall Fuzzy Supplementary indicator could be the simple average of the corresponding indicators. An advantage would be that the overall indicators would fulfill

4. Data

The dataset used in the study is the Mozambican Household Budget Survey 2008-09 (IOF08) (*Inquerito aos Agregados Familiares sobre Orçamento Familiar 2008-09*), a nationally representative household survey conducted by the National Institute of Statistics (INE). The IOF08 was conducted from August 2008 to September 2009. The survey has a stratified structure with three steps of selection: i) selection of the primary sampling units (PSUs), ii) selection of the enumeration areas[3] within the PSUs, and iii) selection of the households within the enumeration areas.

Nine households were selected in each rural area and twelve in each urban area. Twenty-one strata were constructed, one for each urban/rural sample of the 11 provinces of Mozambique (the province of Maputo City does not have a rural area). The IOF08 has a sample size of 51,177 individuals in 10,832 households, divided into 5,223 urban households and 5,609 rural households. It is representative at the national, regional (North, Centre, South), provincial and urban/rural level. Sampling weights are provided in the survey dataset.

The survey includes information on general characteristics of the individuals and of the households, on daily, monthly and durable goods final consumption expenditures, own consumption, transfers and gifts. A related survey at community level is also available, though only for rural areas, containing information on community characteristics. Supplementary information for the IOF08 can be found in (INE, 2010; MPD-DNEAP, 2010).

Concerning socioeconomic status, we use data on (real) per capita daily consumption, available from the IOF08. Such variable is used by the Government for official analyses of poverty which makes our results immediately comparable to existing ones. Such indicator takes into account daily and monthly expenditures, durable goods use and rent, self-consumption, and in kind earnings. All different sources of consumption are averaged out on a daily basis, providing a good measure of economic status for Mozambican households. This measure of income also considers the inflation that

consistency properties with respect to decomposition (Chakravarty et al., 1998; among others). A drawback would be that the weighted mean of the individual would not be equal to the Fuzzy Monetary and the Head Count Ratio indicators.

3 An enumeration area (EA) represents the area assigned to each enumerator for distributing questionnaires to households and it is the smallest building block of the geographical frame for the Mozambican Household Budget Surveys.

occurred during the implementation of the surveys, the different values of the *Metical* - the Mozambican currency - in different periods of the year, and spatial differences in price levels among different provinces and areas (rural/urban). Thirteen relatively homogeneous regions and poverty lines are identified for Mozambique, so as appreciate the differences in price levels in the various provinces and rural/urban areas[4].

In order to compute a measure for non-monetary poverty we use information on ownership of durable goods, housing quality, health status and education level. In particular, we consider ownership of bikes, motorbikes, radios, watches, beds, TVs, computers, printers, sewing machines, fans or air conditioners, telephones or mobiles, new or used cars, irons, other tools, and fridges or freezers.

With regards to housing quality, we use information on: the house where the household lives being a proper house or a thatched hut; walls, roof or floor being made of high-quality materials; information on access to safe water, on sanitation quality, on the most used source of energy for cooking or lighting purposes; and on the number of rooms per (squared) household member.

We also take the highest education level attained in the household and whether someone in the household can read or write to be a realistic measure of non-monetary poverty. Having a household member affected by chronic illnesses and having a chronically malnourished (stunted) child, both health-related indicators, are also added. Finally, other household characteristics such as having a bank account, having someone in the house with a formal or informal job, together with the average number of daily meals the household members can afford are included.

4.1 Problems with the data

The IOF08 is a very rich and detailed dataset. It has been carefully designed and implemented, to the effect of providing reliable information and statistical results. Nonetheless, a few problems with the data were encountered while conducting analysis on multidimensional poverty. In particular, it was found that sampling weights were not calibrated at the household level following a non-response or other problems occurred in the surveying process. Moreover, such weights ranged from 54.6 to 93,452.2, and as a

4 This is the same methodology used in official analyses of poverty.

result of this, a few households with very high weights significantly influenced statistical results.

In terms of household real consumption - the variable used to assess socioeconomic status - we conformed to official analyses that divide it by the number of household members, and on the basis of such a variable we estimated poverty rates. However, this overlooks issues of intra-household allocation of resources and economies of scale, which might considerably matter when dealing with poverty estimates in a country whose average household size is approximately made up of six members. Indeed, the Head Count Ratio computed dividing household consumption per adult equivalent produces very different poverty estimates where the percentage of poor was 36.8% vs. 54.69%.

5. Empirical analysis and results

In this section we describe the steps involved in the measurement of multidimensional poverty in Mozambique at national, provincial and urban/rural level, as outlined in previous sections. This is followed by an analysis of the results for Mozambique's monetary poverty and the different dimensions of non-monetary deprivation.

As introduced in Section 4, the Fuzzy Monetary measures, FM, are based on a household's real consumption divided by its size. Real consumption is obtained by taking into account regional differences in price levels, inflation and seasonal fluctuations. In order to obtain FM, we need to take into account both the proportion of households richer than each particular household and the cumulative share of consumption such richer households receive. Finally, the resulting distribution is transformed such that its mean is equal to the Head Count Ratio: this ensures comparability between the two measures and the two approaches, namely the traditional and the multidimensional one.

As for the Fuzzy Supplementary measures, we use information about thirty-two basic items, as described in Section 4. The deprivation dimensions are initially determined using an exploratory factor analysis: this procedure permits to describe the variability among observed variables - our basic items - in terms of a lower number of unobserved, uncorrelated variables, which are called factors. In the exploratory factor analysis the observed variables are expressed as a linear combination of the underlying factors, without any a priori assumption about the factor structure.

The results of the exploratory factor analysis are then calibrated according to the literature and to the experience acquired during the fieldwork in Mozambique. For example, owning a mobile was moved to dimension FS2 (more widespread durable goods), while number of rooms per household member was included in dimension FS1 (housing conditions), despite both being more correlated with other variables or factors. The results of the exploratory factor analysis and the subsequent reorganisation and calibration is found in Appendix A.1.

Finally, a confirmatory factor analysis was performed by imposing a priori assumptions on the underlying factor structure. This allowed us to test whether the proposed calibration of initial items into a lower number of dimensions made statistical sense (Table 2).

After these preliminary steps, thirty-two basic indicators were grouped into six dimensions, roughly corresponding to: i) housing conditions; ii) more widespread and affordable durable goods; iii) less common, more expensive durable goods; iv) housing quality; v) income-related deprivation; vi) health and education. The complete list of the selected indicators and the resulting dimensions are reported in Table 1, while the results of the confirmatory factor analysis are summarised in Table 2. They are quite satisfactory as:

- The Goodness of Fit Index (GFI) is 0.83. It is based on the ratio of the sum of squared discrepancies to the observed variances; it ranges from 0 to 1 with higher values indicating a good fit.
- The Adjusted Goodness of Fit Index (AGFI) is 0.81. It is the GFI adjusted for degrees of freedom of the model, that is the number of the fixed parameters. It can be interpreted in the same manner.
- The Parsimonious GFI is 0.77. It adjusts GFI for the number of estimated parameters in the model and the number of data points.
- The Root Mean Square Residual (RMR) is 0.08. The fit is considered really good if RMR is equal or below 0.06.
- The Root Mean Squared Error of Approximation (RMSEA) is 0.07. It is based on

the analysis of residuals, with small values indicating a good fit. Values below 0.1, 0.05 and 0.01 indicate a good, very good and outstanding fit respectively.

For what concerns the aggregation of different indicators in each single dimension, a weighting procedure was carried out, as described in Section 3. Depending on the distribution of each indicator in the population and its correlation with other indicators in the same dimension, we constructed item-specific composite weights with equal value for all households in the population. The item-specific weights, W_j, are composed of two parts: Wa_j, which is an inverse function of the percentage of people deprived in item j, and Wb_j, an inverse function of the correlation between item j and all the other items in the same dimension. For each dimension we have that $W_j = Wa_j \cdot Wb_j$.

Intuitively, the first component of the weights, Wa_j, takes into account that if a high percentage of people possess j, then the few who do not possess j are very deprived; the second component, Wb_j, tries to achieve parsimony assigning a lower weight to items that are highly correlated in the same dimension (e.g. high-quality walls and high-quality roof in the 'housing conditions' dimension).

The result is the identification of six different fuzzy supplementary measures, one for each dimension: FS1, FS2, FS3, FS4, FS5, FS6. Subsequently, we aggregate the different non-monetary dimensions into a single composite Fuzzy Supplementary poverty indicator, FS. This is done by assigning equal weights to each supplementary dimension, based on the assumption that all dimensions are equally important in determining supplementary deprivation. The resulting FS distribution is also scaled so that its mean is equal to the Head Count Ratio, as we did for the monetary poverty indicator, FM. The rescaling ensures that the traditional and the fuzzy indicators are comparable.

The results for the Fuzzy Monetary (FM), the composite Fuzzy Supplementary (FS) measure, and for the individual Fuzzy Supplementary dimensions (FS1-FS6) at national level are outlined in Subsection 5.1, while those relative to the provincial and urban/rural level are presented in Subsection 5.2 and Subsection 5.3, respectively. A detailed description of the technique used for the computation of the standard errors and relative advantages compared to other methods is found in Appendix B.

5.1 Poverty estimates at national level

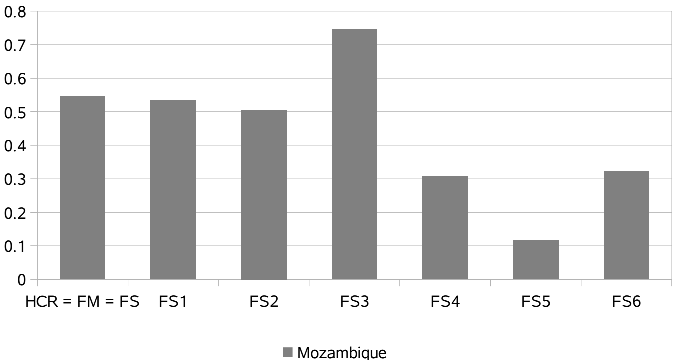
As previously outlined, the overall Fuzzy Monetary (FM) and Fuzzy Supplementary (FS) dimensions are constructed such that their mean is equal to the official Head Count Ratio so they do not convey additional information to our analysis at national level (Head Count Ratio = FM = FS = 54.69%). Hence, in this subsection we only focus on the values of the supplementary dimensions FS1-FS6.

From Table 3, it can be seen that the factor with highest level of deprivation is FS3 which corresponds to less common, expensive durable goods. Most Mozambicans do not possess any of the items included in this dimension and a level of deprivation of about 0.75 is thus reasonable. Conversely, the deprivation value for less expensive durable goods (FS2) is lower, showing that some durable goods -especially mobile phones and bikes - are becoming more common in the country.

The level of deprivation for housing conditions (FS1) is also very high (0.53), and reflects the fact that many households lack basic facilities in their dwellings. Even so, the proportion of households lacking decent household quality (FS4) is significantly lower (0.31). Income-related deprivation (FS5) appears to be relatively low: this result is probably influenced by the inclusion of a dummy for whether someone in the household had a job (formal or informal) or not. Since most of the households interviewed (about 98%) had a member with a formal or informal job, the entire dimension was pushed towards low levels of deprivation (0.12). When this variable is eliminated from the FS5 dimension, the average deprivation raises significantly (0.64). This was then taken into account in the following analyses, other sensitivity checks were undertaken, and the results presented in Appendix A.2.

Finally, the result for health and education (FS6) shows that education and health conditions in Mozambique are improving. However, one needs to be warned that the relatively low average value of deprivation for this dimension (0.32) is likely to be affected by the low level of deprivation characteristic of chronic illnesses and ability to read and write. Indeed, the level of child malnutrition in Mozambique is still among the highest in the world (WHO, 2011).

Figure 1 – Deprivation by dimension, national level



Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education.

5.2 Poverty estimates at provincial level

When fuzzy set poverty analysis is carried out at sub-national level then it becomes evident how the inclusion of multiple dimensions substantially increases the amount of available information.

Mozambique is divided into eleven provinces[5]. These territories are quite heterogeneous with regard to economic development, culture, ethnic and linguistic composition. Consequently, huge differences in poverty rates exist among different zones and provinces in Mozambique. Even though some insights emerged from the official Head Count reports, the multidimensional analysis of poverty we undertook using Fuzzy Set Theory allows us to highlight important characteristics that would otherwise go unnoticed in a traditional poverty assessment.

In particular, by looking at the Fuzzy Monetary (FM) and Fuzzy Supplementary (FS) statistics presented in Table 4 it appears striking that some of the provinces with low rates of monetary poverty are also much more deprived in other dimensions, and the converse is also true. The Northern provinces (Niassa, Cabo Delgado, Nampula) and the Central province of Tete, all have much higher Fuzzy Supplementary (FS) averages with respect to their Fuzzy Monetary (FM) ones. The other Central provinces (Zambezia,

5 The eleven provinces of Mozambique are grouped into three bigger zones: the North, which includes the provinces of Niassa, Cabo Delgado and Nampula; the Centre, with the provinces of Zambezia, Tete, Manica and Sofala; the South, containing the provinces of Gaza, Inhambane, Maputo Province and Maputo City.

Manica and Sofala) have similar statistics in both the FM and FS dimension, while the Southern provinces show FS averages that are lower than their respective FM averages. The estimated averages for the Head Count Ratio, the Fuzzy Monetary and Fuzzy Supplementary dimensions together with their standard errors are presented in the first rows of Table 4 and in Figure 2.

The analysis of Fuzzy Supplementary dimensions indicates that the South is generally more developed than the Centre and the North, with Maputo City being much less deprived than all other provinces. These characteristics remained hidden using the standard poverty Head Count analysis. This is probably due to various causes: first, consumption is highly dependent on temporary and/or seasonal fluctuations - e.g. a bad harvest in 2008 -, while other dimensions as those included in the computation of the Fuzzy Supplementary statistics are more robust to such changes. Indeed, buying an asset, a durable good or investing in education requires an evaluation of a household's economic status that is only partially related to the level of income/consumption in a given year. Moreover, a large part of the Mozambican population has consumption levels that are close to the poverty line, hence even small fluctuations can alter the poverty Head Count statistics in a substantial way. This is one of the main drawbacks of using a dichotomous index like the Head Count Ratio for the analysis of a complex phenomenon such as poverty. In fact, poverty Head Count analyses based on Mozambican Budget Surveys generally yielded strange or non-robust results, with strong fluctuations in the Head Count Ratio and re-ranking of poor and rich provinces (Van den Boom, 2011; pp. 7-8).

A deeper investigation into supplementary factors yields additional results (Table 4). As for housing conditions (FS1), we can identify three distinct groups of provinces on the basis of their FS1 averages: the Central provinces (Zambezia, Tete, Manica and Sofala) and the province of Nampula are the most deprived in this dimension with an average of about 0.60 for Nampula, Manica and Sofala but roughly 0.70 for Zambezia and Tete. In the second group, with an average deprivation of about 0.40, we find two Northern provinces (Niassa and Cabo Delgado) and two Southern provinces (Gaza and Inhambane). Finally, the least deprived provinces are again Maputo Province and Maputo City, the latter with an average level of deprivation of 0.03.

In the FS2 dimension we put together some durable goods that are more widespread than others, like mobile phones, bikes and motorbikes, radios, watches and TVs. Indeed, most provinces show similar average levels of deprivation in this dimension, ranging within the 0.44-0.55 interval where Nampula is the most and Niassa and Manica the least deprived with 0.60, 0.36, and 0.33 scores respectively.

Conversely, the FS3 dimension consists of those durable goods that are less affordable and thus less common among Mozambicans such as cars, fridges or freezers, irons, computers, printers, other tools, and sewing machines. As evidenced in the analysis at national level, this is the factor for which average levels of deprivation are highest. This is particularly true in the North and in the Centre, where five provinces (Niassa, Cabo Delgado, Nampula, Zambezia and Tete) have average values that exceed 0.80, in contrast to other Central provinces (Manica and Sofala) with values around 0.75 which perform a little better. Once again, the Southern provinces of Gaza, Maputo Province and Maputo City have much lower deprivation levels with scores 0.52, 0.43, and 0.28 respectively, confirming the finding that Southern provinces are less deprived than Northern and Central provinces in various dimensions.

As for access to safe water, energy sources for cooking, in-house lighting and the likes - included in the FS4 dimension -, we find that the average level of deprivation is relatively low. For Northern and Central provinces it ranges between 0.28 for Sofala and 0.41 for Niassa, while all Southern provinces perform comparatively better.

As presented in Subsection 5.1, the FS5 dimension (income-related deprivation) is the one for which average levels of deprivation are lowest. In this case, there are no noticeable differences between provinces. However, when the variable “formal or informal job” is taken out, then it emerges that there is a group of provinces including Manica, Sofala, Maputo, and Maputo City, with average deprivation values between 0.40 and 0.50 whilst all other provinces perform comparatively worse with values of around 0.65-0.75.

Finally, the last supplementary dimension (FS6) takes into account education measured as education level and ability to read and write and health measured as child malnutrition and chronic illnesses. In this case, Maputo Province and Maputo City record an average level of 0.12-0.15, while the estimated values for other provinces range between 0.28 for Sofala and 0.38 for Zambezia, which amounts to more than twice the level of deprivation of the two most Southern provinces.

What this subsection makes clear is that the analysis of dimensions other than consumption substantially improves the mapping of provincial differences regarding poverty. In particular, the higher level of development of the Southern provinces distinctly surfaced in more than one dimension (FS, and particularly FS1, FS3, FS4, FS6). At the same time, understanding which factors are most influential upon deprivation, yields a deeper insight about which characteristics are more unequally distributed throughout the country.

The estimated averages and standard errors for the supplementary dimensions are found in Table 4, while a graphical analysis of the results is shown in Figure 2, 3, 4 and 5. In Figure 2 we present the Head Count Ratio, Fuzzy Monetary and Fuzzy Supplementary averages for all provinces, which highlights the differences that exist between monetary and overall non-monetary deprivation. In Figure 3, instead, all the different supplementary dimensions are shown, divided by region and province. In Figure 4 both the monetary and individual non-monetary dimensions are shown for each region and province on a net graph. This kind of graph provides additional information about the overall condition of each province compared to other provinces in the same region. Figure 5 is particularly informative since it allows a comparison of all provinces in all dimensions, and reveals the gap between the Centre-North and the South for supplementary dimensions of deprivation as clearly evident.

In Table 5 we present the ranking of each province in all dimensions of deprivation. While being very similar for the Head Count Ratio and the Fuzzy Monetary (FM) measure, it greatly differs when it comes to supplementary dimensions.

Figure 2 – Head Count Ratio (HCR), Fuzzy Monetary (FM) and Fuzzy Supplementary (FS), by province

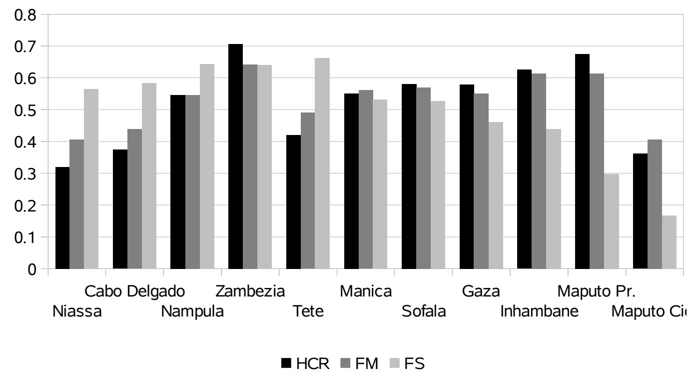
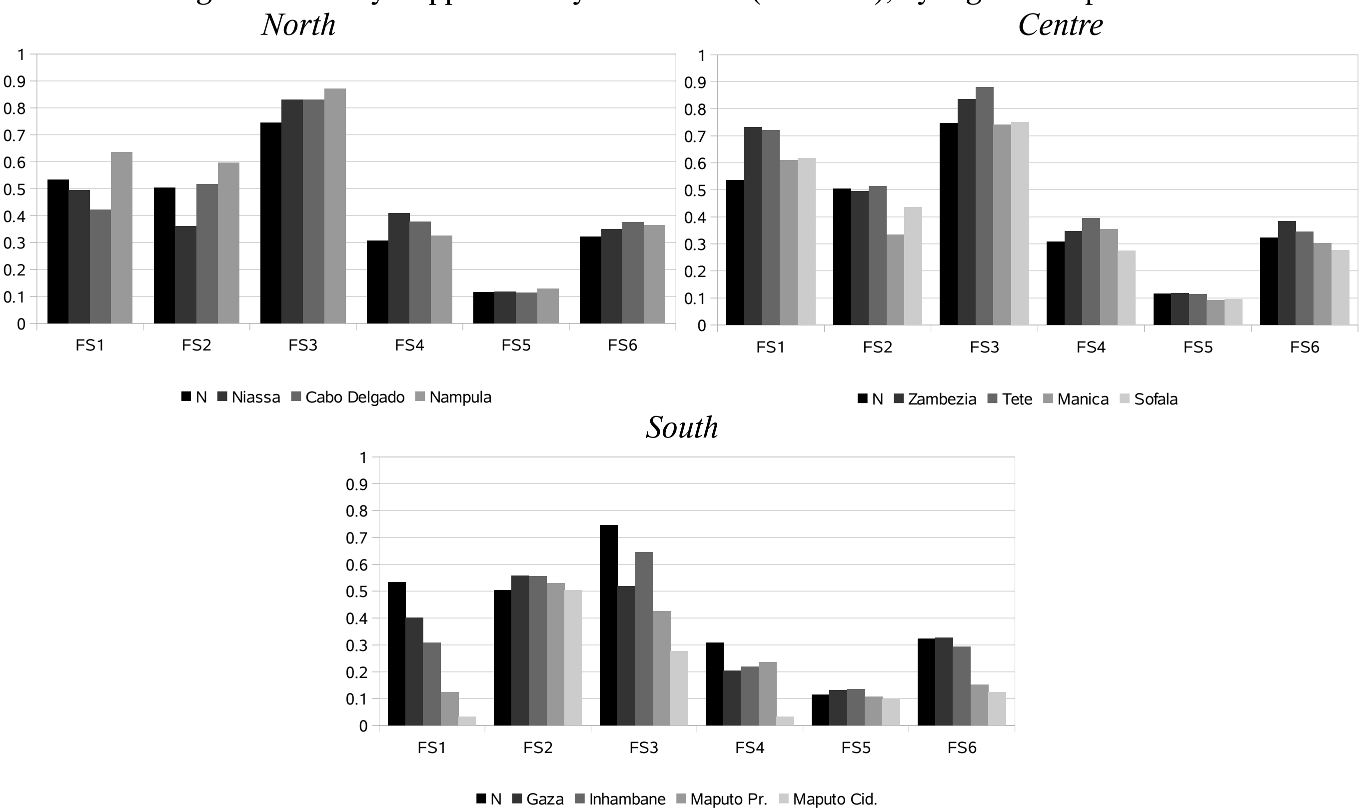
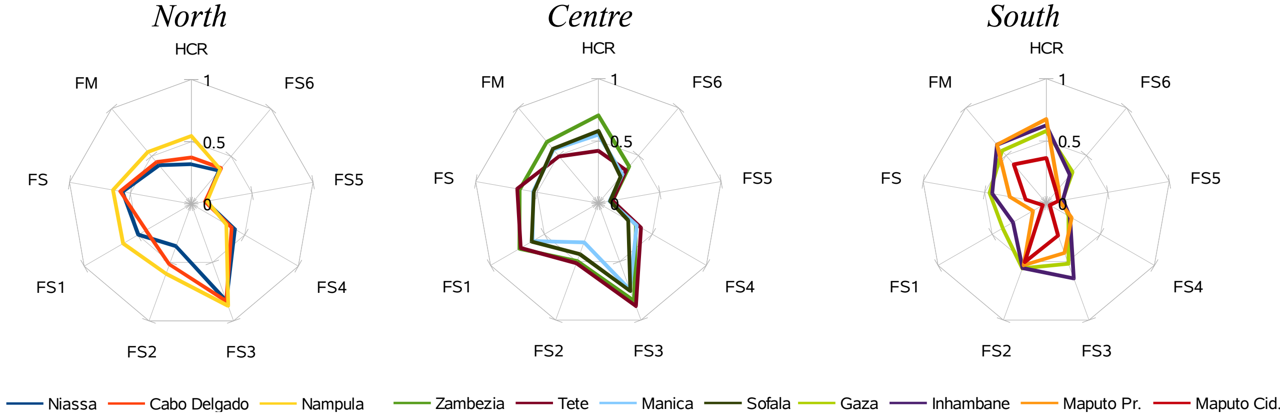


Figure 3 – Fuzzy Supplementary dimensions (FS1-FS6), by region and province



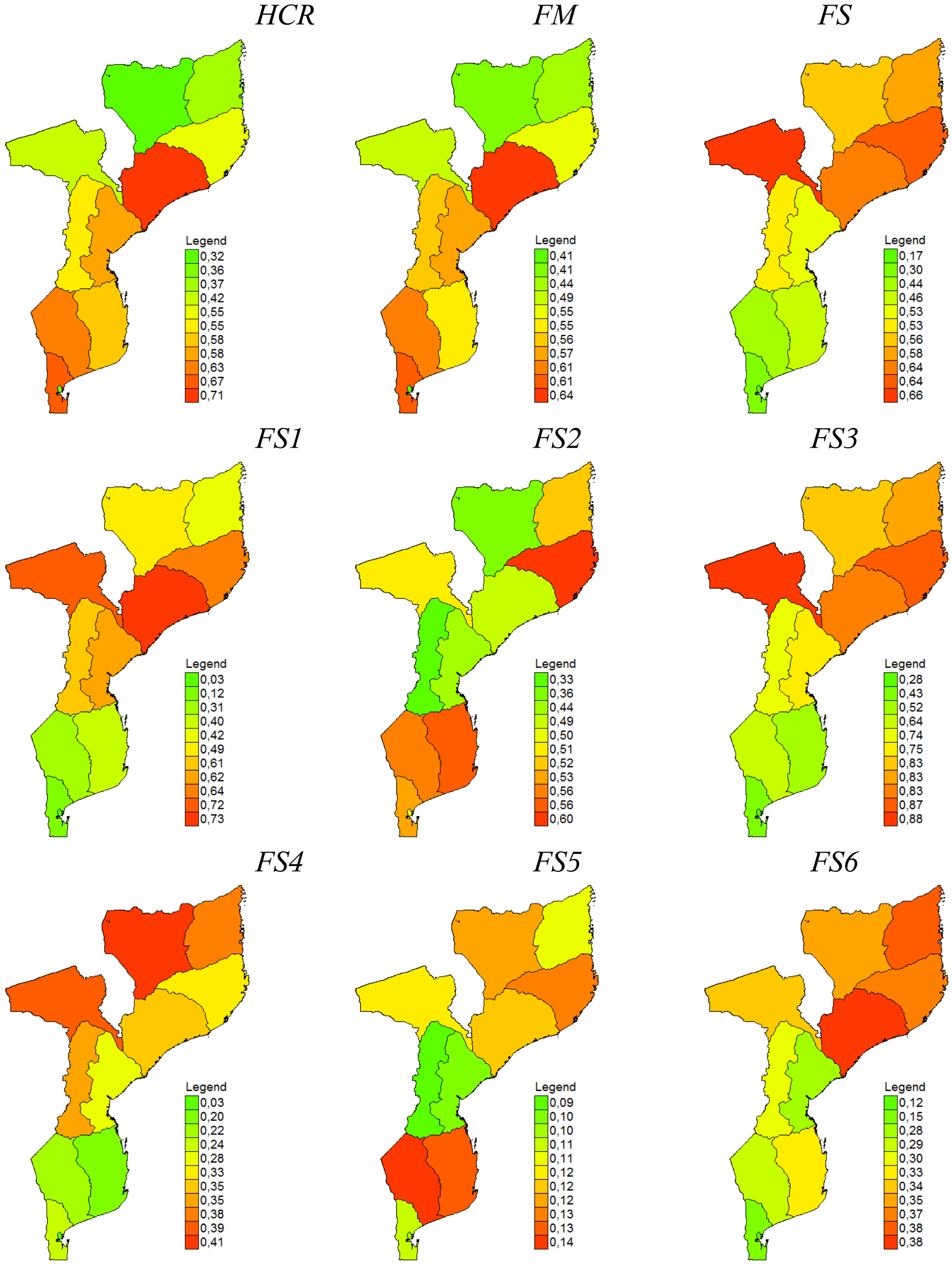
Note: FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education.

Figure 4 – Deprivation by dimension and province



Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education.

Figure 5 – Maps of deprivation, by dimension



Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education.

5.3 Multidimensional poverty estimates by province and area of residence (urban/rural)

In what follows, we present multidimensional deprivation as estimated by province and by area of residence (urban/rural). The huge differences in poverty estimates that exist between urban and rural areas at both national and sub-national level in Mozambique were already discussed in Section 2. Nonetheless, unexpected results emerge from the analysis of supplementary dimensions of deprivation (Table 6): when these are introduced, the urban/rural deprivation gap widens substantially, contrasting with the official analyses based on consumption that estimate a differential of about seven percentage points: at national level the official Head Count Ratio of rural and urban areas for 2008-09 is about 56.9% and 49.6%, respectively (MPD-DNEAP, 2010). Such gap is very small when compared with urban/rural poverty gaps in other Sub-Saharan countries (Van den Boom, 2011; World Bank, 2011): for example, while 50% of rural people in Kenya have income below the poverty line, only 32% of the urban population is considered poor. An even wider gap is found in Uganda, where the rural Head Count Ratio is 34%, as opposed to 14% in urban areas. Lastly, in Ghana the percentage of poor people in rural areas is more than three times that of poor people in urban areas, 39% and 11% respectively.

However, when supplementary dimensions of deprivation are considered, a different picture also emerges for Mozambique. The aggregated Fuzzy Supplementary (FS) deprivation level for urban areas is 0.34, whereas the one for rural areas exceeds 0.63. Such difference is due to the urban/rural gap found in the underlying supplementary dimensions. In particular, housing conditions (FS1), possession of less common, more expensive durable goods (FS3), housing quality (FS4) and - to a lesser extent - health and education (FS6) all show very different deprivation levels for urban and rural areas (Figure 6).

For housing conditions (FS1) the urban deprivation level is 0.26, the rural one as high as 0.65. For more expensive durable goods (FS3) they are equal to 0.52 and 0.84 respectively. The values for the housing quality dimension (FS4) are 0.13 for urban areas and 0.38 for rural areas while those for the health and education dimension (FS6) are 0.21 for urban and 0.37 for rural. Much smaller differences exist in the more widespread durable goods (FS2) and income-related (FS5) deprivation dimensions[6]. The wide deprivation gap between urban and rural areas typical of most supplementary dimensions at national level is also reflected at the provincial one. Point estimates and standard errors are found in Table 6 divided by province and area of residence.

The central regions of Manica and Sofala exhibit the greatest difference between supplementary deprivation values in urban and rural areas. In the supplementary dimensions FS1 (housing conditions), FS3 (more expensive, less affordable durable goods) and FS4 (housing quality) such difference is conspicuous, ranging 30 to 60 percentage points. While the urban areas of these two provinces are among the least deprived areas of Mozambique in all dimensions, the opposite holds for their rural counterparts. The urban/rural deprivation gap for the dimensions FS1, FS3 and FS4 is substantial also for other provinces such as Niassa, Cabo Delgado, and especially Nampula, Zambezia, and Tete. Moreover, the Southern provinces of Gaza and Maputo Province also show significant differences between rural and urban areas.

Urban and rural deprivation levels are instead comparable for more widespread durable goods (FS2) and income-related (FS5) supplementary dimensions. Some of the rural areas score even better than their relative urban areas in FS2 (Tete, Manica). As pointed out in previous paragraphs, excluding the variable "formal or informal job" from the FS5 dimension changes the results for this dimension substantially. When this variable is excluded the difference between urban and rural areas increases largely for Niassa, Nampula, Tete, Sofala, Gaza and Inhambane. Both figures, with and without the variable "formal or informal job", are presented in Figure 6.

For what concerns FS6 (health- and education-related indicators), rural areas are systematically more deprived than urban areas. This is plausible, as healthcare facilities and schools are more widespread in urban areas. The average gap between areas of residence amounts to more than ten percentage points, notwithstanding the commitment of the Mozambican government to increase the availability of health and education facilities in rural areas (Chao and Kostermans, 2002; Government of Mozambique, 2005; Republic of Mozambique, 2006).

---

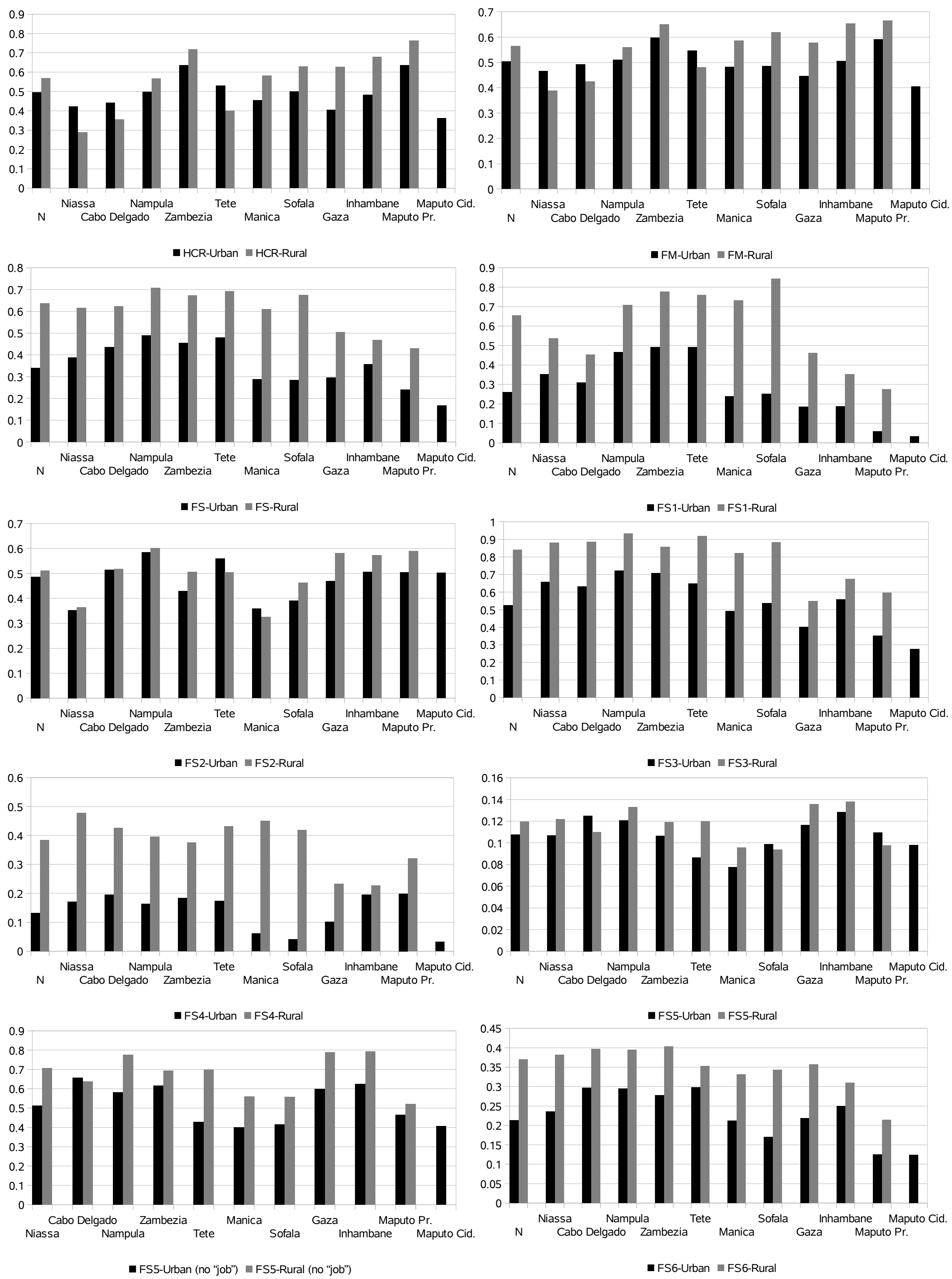
6 Again, the urban/rural difference increases for dimension FS5 when the variable "formal or informal job" is not considered: in this case the average urban deprivation becomes 0.51, while the rural becomes one 0.70.

As previously shown, both monetary deprivation dimensions - Head Count and Fuzzy Monetary - analysed at national level determine radically different results to non-monetary dimensions. This holds true also for the analysis at provincial and urban/rural level. Head Count Ratio and Fuzzy Monetary estimates reveal that the poorest region in Mozambique is the rural area of Maputo Province, while the rural areas of Niassa, Cabo Delgado and Tete are richer than their urban counterparts and present the same low deprivation levels of Maputo City, the capital. The urban/rural deprivation gap of Manica and Sofala in these monetary dimensions is not as wide as when it is measured by the supplementary dimensions, while the urban/rural Head Count Ratio gap of Gaza and Inhambane is found significantly wider. Introducing supplementary dimensions to the analysis of poverty in Mozambique substantially increases the amount and quality of available information, providing figures that often contrast with the ones derived solely from monetary poverty estimates.

In Figure 6 we present the average deprivation levels for each dimension and for all provinces distinguished by area of residence. The province of Maputo City does not have a rural area.

In Table 7 we show the ranking of each urban/rural area for all provinces in all dimensions of deprivation. As it was found for provincial analysis, the Head Count Ratio and the Fuzzy Monetary (FM) measure generate very similar rankings, while those produced by supplementary dimensions differ both within and in comparison to monetary measures. On average, it appears that the ranking of some rural area worsens sensibly as in the case of Cabo Delgado, Tete, and Niassa, yet it improves for other provinces like Zambezia or Maputo Province.

Figure 6 – Deprivation by dimension, provincial and urban/rural level



Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education. The province of Maputo City does not have a rural area.

6. Conclusions

In this study we have shown how it is possible to construct poverty measures relative to monetary and non-monetary dimensions using Fuzzy Set Theory. We applied this technique to the Mozambican Household Budget Survey 2008-09 (IOF08) dataset, the most recent budget survey available for Mozambique.

Our main contribution to the analysis and measurement of poverty in Mozambique is twofold. On the one hand, we operationalise a concept of poverty wider than monetary poverty, therefore involving supplementary dimensions. At the same time, we obtain reliable estimates of poverty rates at sub-national and urban/rural level, by using the Jackknife Repeated Replications method to compute standard errors.

To our knowledge, this is the first study that applies Fuzzy Set Theory to the measurement of poverty in Mozambique. As a result, the figures provided in the study substantially increase the amount and quality of available information about Mozambican households' deprivation. Our estimates - especially those obtained for non-monetary dimensions - complement the ones derived solely from the Head Count Ratio. They also provide new evidence with respect to provincial and urban/rural deprivation levels.

With regards to monetary poverty, the Fuzzy Monetary estimates essentially confirm the official results obtained using the Head Count Ratio. In particular, the ranking of poorer and richer provinces remains unchanged, also when the analysis is carried out at the urban/rural level. This is due to both measures, the Head Count and Fuzzy Monetary, being based on consumption data.

Instead, innovative results come from the inclusion of six supplementary dimensions of deprivation in the analysis of poverty: housing conditions; more widespread and affordable durable goods; less common, more expensive durable goods; housing quality; income-related deprivation; health and education. When these dimensions are considered, some of the provinces showing relatively low Head Count Ratios are found to be among the most deprived with respect to supplementary dimensions of deprivation, and conversely. In particular, the Northern provinces and the Central province of Tete, all show much higher Fuzzy Supplementary (FS) averages with respect to their Fuzzy Monetary (FM) averages. The remaining Central provinces have similar statistics in both the FM and FS dimension, while the Southern provinces show lower FS averages than their respective FM averages.

The higher level of development of the Southern provinces distinctly becomes relevant to more than one supplementary dimension: housing conditions (FS1); less common, more expensive durable goods (FS3); housing quality (FS4); and, to a lesser extent, health and education (FS6).

Furthermore, in our analysis we point out that deprivation values found in urban and rural areas are very different. When we consider non-monetary dimensions of deprivation we find that the urban/rural gap is much wider than when it is measured using Head Count Ratio or Fuzzy Monetary statistics. The aggregated Fuzzy Supplementary deprivation level for urban areas is estimated to be 0.34, whereas the one for rural areas exceeds 0.63. Moreover, while the ranking of some rural area such as Cabo Delgado, Niassa, and Tete sensibly worsens, it improves for the provinces of Zambezia and Maputo Province).

One partial explanation of this large difference between the results for monetary and non-monetary poverty is that some of the items included in the supplementary analysis are non-essential items, like fridge, car or PC. For example, the highest average level of deprivation is found for FS3 (less common, expensive durable goods), as most Mozambicans do not possess any of the items included in this dimension, especially in the North and in the Centre. In fact it might be objected that the inclusion of these items in the analysis of poverty is not entirely justified. However, the said difference between monetary and non-monetary poverty is large also for those supplementary dimensions like housing conditions, housing quality, or health and education, which certainly denote a situation of deprivation.

Moreover, our findings are broadly in line with those reported by analyses of poverty for other Sub-Saharan countries: for example Bradshaw and Steyn (2001) and Ngwane et al. (2001) study multidimensional poverty with various methods and find that the rankings of South African regions are different when non monetary poverty measures are used . The same is true for Ghana (Appiah-Kubi et al., 2007) and for a group of other Sub-Saharan countries (Batana, 2008). Perhaps the main difference is that we find larger differences between monetary and non monetary poverty values among the different regions and provinces of Mozambique.

Our results are particularly relevant since Mozambique is among the poorest countries in the world, given its per capita income level of approximately $428 (World Bank, 2010), and several donor countries and international agencies involved in poverty reduction plans. Accurate information and measurement of poverty at the local level is thus required and may be used to redirect funds.

References

Anand, S., Sen, A., 1997. Concepts of human development and poverty: a multidimensional perspective. UNDP New York, .

Atkinson, A., 2002. Social indicators: The EU and social inclusion. Oxford University Press, USA.

Atkinson, A., 2003. Multidimensional deprivation: contrasting social welfare and counting approaches. Journal of Economic Inequality 1(1), 51-65.

Atkinson, A., Bourguignon, F., 1982. The comparison of multi-dimensioned distributions of economic status. The Review of Economic Studies 49(2), 183.

Ballini, F., Betti, G., Çalik, G., Da?kiran, Z., Karaka?, M., 2009. The Siena Micro-Simulation Model (SM2) for the tax-benefit system in Turkey.

Betti, G., Cheli, B., Lemmi, A., Verma, V., 2006. Multidimensional and longitudinal poverty: an integrated fuzzy approach. In: Lemmi, A., Betti, G. (Ed.), Fuzzy set approach to multidimensional poverty measurement, Springer.

Betti, G., Ferretti, C., Gagliardi, F., Lemmi, A., Verma, V., 2009. Proposal for new multidimensional and fuzzy measures of poverty and inequality at national and regional level. Working Paper n. 83, Dipartimento di Metodi Quantitativi, Università di Siena.

Betti, G., Verma, V., 1999. Measuring the degree of poverty in a dynamic and comparative context: a multi-dimensional approach using fuzzy set theory. Proceedings, ICCS-VI, Vol. 11, pp. 289-301, Lahore, Pakistan, August 27-31, 1999.

Bourguignon, F., Chakravarty, S., 2003. The measurement of multidimensional poverty. Journal of Economic inequality 1(1), 25-49.

Castel-Branco, M., 2010. Intervention at the launch of the book “Economia extractiva e desafios de industrializacao em Mocambique". Instituto de Estudos Sociais e Economicos, 21/10/10, Maputo.

Cerioli, A., Zani, S., 1990. A fuzzy approach to the measurement of poverty. In: Dagum, C., Zenga, M. (Ed.), Income and wealth Distribution, Inequality and Poverty, Springer Verlag Berlin, 272-284.

Chakravarty, S., Mukherjee, D., Ranade, R., 1998. On the family of subgroup and factor decomposable measures of multidimensional poverty. Research on Economic Inequality 8, 175-194.

Chao, S., Kostermans, K., 2002. Improving health for the poor in Mozambique. The fight continues. Health, Nutrition and Population Discussion Paper. World Bank, Washington DC.

Cheli, B., 1995. Totally fuzzy and relative measures of poverty in dynamic context. Metron 53, 183-205.

Cheli, B., Betti, G., 1999. Fuzzy analysis of poverty dynamics on an Italian pseudo panel, 1985-1994. Metron 57, 83-104.

Cheli, B., Lemmi, A., 1995. A" Totally" fuzzy and relative approach to the multidimensional analysis of poverty. Economic Notes 24, 115-134.

Chiappero Martinetti, E., 1994. A new approach to evaluation of well-being and poverty by fuzzy set theory. Giornale degli Economisti e Annali di Economia 53, 367-388.

Chiappero Martinetti, E., 2006. Capability approach and fuzzy set theory: description, aggregation and inference issues. In: Lemmi, A., Betti, G. (Ed.), Fuzzy Set Approach to Multidimensional Poverty Measurement, Springer.

Duclos, J., Sahn, D., Younger, S., 2006. Robust Multidimensional Poverty Comparisons. The Economic Journal 116(514), 943-968.

Durbin, J., 1959. A note on the application of Quenouille's method of bias reduction to the estimation of ratios. Biometrika 46(3/4), 477-480.

Efron, B., Stein, C., 1981. The jackknife estimate of variance. Annals of Statistics 9, 586-596.

Eurostat, 2002. Income, Poverty and Social Exclusion: 2nd Report. Luxembourg: Office for Official Publications of the European Communities.

Eurostat, 2009. Description of SILC user database variables: cross-sectional and longitudinal. Version 2007. Luxembourg: Office for Official Publications of the European Communities.

Filippone, A., Cheli, B., D'Agostino, A., 2001. Addressing the interpretation and the aggregation problems in totally fuzzy and relative poverty measures. ISER Working Paper number number 2001-22 University of Essex.

Government of Mozambique, 2005. Education sector strategy plan II 2005-2009.

Hanlon, J., 2007. Is poverty decreasing in Mozambique?. IESE Conference Paper n.14.

INE (2004) Inquérito Nacional aos Agregados Familiares Sobre Orçamento Familiar 2002/3, http://www.ine.gov.mz/inqueritos_dir/iaf/

INE (2010) Inquérito Sobre Orçamento Familiar, 2008/09 – Quadros Básicos, http://www.ine.gov.mz/inqueritos_dir/iaf/

Kish, L., Frankel, M., 1974. Inference from complex samples. Journal of the Royal Statistical Society B(36), 1-37.

Klasen, S., 2000. Measuring poverty and deprivation in South Africa. Review of Income and Wealth 46(1), 33-58.

Kolm, S., 1977. Multidimensional egalitarianisms. The Quarterly Journal of Economics 91, 1-13.

Maasoumi, E., 1986. The measurement and decomposition of multi-dimensional inequality. Econometrica 54, 771-779.

MPF, 2004. Poverty and Well-Being in Mozambique: The Second National Assessment (2002-2003). National Directorate of Planning and Budget, Ministry of Planning and Finance, Economic Research Bureau, Ministry of Planning and Finance, International Food Policy Research Institute, Purdue University.

MPD-DNEAP, 2010. Poverty and wellbeing in Mozambique: third national poverty assessment. Maputo: Ministério da Planificação e Desenvolvimento - Direcção Nacional de Estudos e Análise de Políticas.
MPF, 1998. Poverty and Well-Being in Mozambique: The First National Assessment (1996-1997). Ministry of Planning and Finance, Universidade Eduardo Mondlane, International Food Policy Research Institute.
Ossemane, R., 2010. O dilema do crescimento empobrecedor. 26/10/10, Tete, Mozambique.
Republic of Mozambique, 2006. Action plan for the reduction of absolute poverty 2006-2009.
Sen, A., 1999. Development as freedom. Oxford University Press, Oxford.
Shao, J., Tu, D., 1995. The Jackknife and Bootstrap. Springer-Verlag, Inc.
Smeeding, T., Saunders, P., Coder, J., Jenkins, S., Fritzell, J., Hagenaars, A., Hauser, R., Wolfson, M., 1993. Poverty, inequality and family living standards impacts across seven nations: The effect of non-cash subsidies for health, education and housing. Review of Income and Wealth 39, 229-256.
Townsend, P., 1979. Poverty in the United Kingdom. Allen Lane, Harmondsworth.
Tsui, K., 1995. Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach. Journal of Economic Theory 67(1), 251-265.
TurkStat, 2007. Income and Living Conditions Survey, 2007. Turkish Statistical Institute, Ankara.
Van den Boom, B., 2011. Analysis of poverty in Mozambique. Available at: http://undp.org.mz/en/Publications/Other-Publications/Analysis-of-poverty-in-Mozambique
Verma, V. (Ed.), 1993. Sampling Errors in Household Surveys: A Technical Study. New York: United Nations Department for Economic and Social Information and Policy Analysis, Statistical Division, INT-92-P80-15E, .
Verma, V., Betti, G., 2011. Taylor linearization sampling errors and design effects for poverty measures and other complex statistics. Journal of Applied Statistics 38(8), 1549-1576.
Vero, J., Werquin, P., 1997. Reexamining the Measurement of Poverty: How Do Young People in the Stage of Being Integrated in the Labor Force Manage. Economie et Statistique 8(10), 143-156.
Whelan, C., Layte, R., Maitre, B., Nolan, B., 2001. Income, deprivation, and economic strain. An analysis of the European Community Household Panel. European Sociological Review 17(4), 357.
WHO, 2011. Nutritional status. Available at: http://www.childinfo.org/undernutrition_nutritional_status.php
World Bank, 2008. Beating the odds: sustaining inclusion in Mozambique's growing economy. Washington D.C.: The World Bank.
World Bank, 2010. World Development Report 2010: Development and Climate Change. Washington D.C.: The World Bank.
World Bank, 2011. World Development Indicators 2010. Washington D.C.: The World Bank.
Zadeh, L., 1965. Fuzzy sets. Information and control 8(3), 338-353.

Table 1 – Dimensions and indicators of non-monetary deprivation

1	Housing conditions	2	More widespread durable goods	3	Less common, expensive durable goods	4	Housing quality	5	Income-related deprivation	6	Health and education
d1	Bed	d8	Telephone or mobile phone	d14	Tools	d21	Energy source for cooking	d26	A bank account	d29	Ability to read and write
d2	Proper house	d9	TV	d15	Electric or coal iron	d22	Energy source for in-house lighting	d27	A formal or informal job	d30	Education level
d3	High-quality walls	d10	Bike	d16	Fridge or freezer	d23	Hotplate or gas ring	d28	More than two meals per day	d31	Chronic illness
d4	High-quality roof	d11	Radio	d17	New or second-hand car	d24	Has access to safe water			d32	Child malnutrition
d5	High-quality floor	d12	Watch	d18	Computer	d25	Fan or air conditioner				
d6	Has WC or latrine	d13	Motorbike	d19	Printer						
d7	Number of rooms/ (household members)^2			d20	Sewing machines						

Table 2 – Confirmatory factor analysis results

Goodness of fit (GFI)[a]	0.834
Adjusted GFI[b]	0.8086
Parsimonious GFI[c]	0.77
Root Mean Square Residual[d]	0.0807
RMSEA[e]	0.0748

Notes:
[a] It is based on the ratio of the sum of squared discrepancies to the observed variances; it ranges from 0 to 1 with higher values indicating a good fit. [b] It is the GFI adjusted for degrees of freedom of the model, that is the number of the fixed parameters. It can be interpreted in the same manner. [c] It adjusts GFI for the number of estimated parameters in the model and the number of data points. [d] The fit is considered really good if RMR is equal or below 0.06. [e] The Root Mean Squared Error of Approximation (RMSEA) is based on the analysis of residuals, with small values indicating a good fit.

Table 3 – Deprivation by dimension, national level

	Mozambique (n = 10831)								
	HCR	**FM**	**FS**	**FS1**	**FS2**	**FS3**	**FS4**	**FS5**	**FS6**
Mean	0.5469	0.5469	0.5469	0.5345	0.5043	0.7457	0.3079	0.1158	0.3224
SE	0.0119	0.0078	0.0056	0.0065	0.0070	0.0076	0.0103	0.0020	0.0063

Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education. The province of Maputo City does not have a rural area.

Table 4 – Deprivation by dimension, provincial level

	Niassa	**Cabo Delgado**	**Nampula**	**Zambezia**	**Tete**	**Manica**	**Sofala**	**Gaza**	**Inhambane**	**Maputo Pr.**	**Maputo Cid.**
Head Count Ratio	0.3194	0.3740	0.5468	0.7054	0.4203	0.5509	0.5803	0.5795	0.6254	0.6746	0.3615
SE	0.0322	0.0353	0.0291	0.0339	0.0444	0.0427	0.0484	0.0475	0.0320	0.0245	0.0239
mean FM	0.4068	0.4386	0.5459	0.6423	0.4903	0.5605	0.5690	0.5504	0.6132	0.6134	0.4054
SE	0.0237	0.0217	0.0186	0.0218	0.0273	0.0249	0.0377	0.0329	0.0233	0.0165	0.0182
mean FS	0.5649	0.5829	0.6431	0.6406	0.6615	0.5313	0.5265	0.4603	0.4385	0.2970	0.1672
SE	0.0244	0.0171	0.0143	0.0123	0.0192	0.0208	0.0255	0.0225	0.0249	0.0159	0.0066
mean FS1	0.4941	0.4229	0.6353	0.7319	0.7200	0.6101	0.6179	0.4021	0.3076	0.1233	0.0323
SE	0.0284	0.0190	0.0148	0.0150	0.0273	0.0284	0.0231	0.0290	0.0298	0.0159	0.0029
mean FS2	0.3615	0.5175	0.5969	0.4948	0.5132	0.3345	0.4360	0.5578	0.5554	0.5305	0.5031
SE	0.0213	0.0211	0.0171	0.0141	0.0285	0.0190	0.0408	0.0236	0.0222	0.0231	0.0137
mean FS3	0.8316	0.8319	0.8710	0.8349	0.8796	0.7410	0.7511	0.5181	0.6439	0.4253	0.2768
SE	0.0200	0.0209	0.0139	0.0280	0.0179	0.0245	0.0236	0.0291	0.0288	0.0240	0.0159
mean FS4	0.4084	0.3763	0.3269	0.3471	0.3945	0.3545	0.2756	0.2048	0.2187	0.2355	0.0327
SE	0.0396	0.0342	0.0285	0.0245	0.0358	0.0395	0.0367	0.0278	0.0438	0.0368	0.0042
mean FS5	0.1186	0.1131	0.1293	0.1172	0.1150	0.0910	0.0957	0.1316	0.1355	0.1060	0.0979
SE	0.0065	0.0057	0.0042	0.0064	0.0067	0.0057	0.0075	0.0043	0.0044	0.0073	0.0069
mean FS6	0.3489	0.3755	0.3652	0.3843	0.3445	0.3017	0.2775	0.3272	0.2936	0.1517	0.1234
SE	0.0146	0.0168	0.0125	0.0228	0.0147	0.0157	0.0242	0.0305	0.0162	0.0156	0.0087
n	814	780	1575	1523	768	804	851	803	814	900	1199

Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education. The province of Maputo City does not have a rural area.

Table 5 – Relative ranking according to the different dimensions of deprivation (1: lowest deprivation; 10: highest deprivation)

	HCR	**FM**	**FS**	**FS1**	**FS2**	**FS3**	**FS4**	**FS5**	**FS6**
Niassa	1	2	7	6	2	7	11	8	8
Cabo Delgado	3	3	8	5	7	8	9	5	10
Nampula	5	5	10	9	11	10	6	9	9
Zambezia	11	11	9	11	4	9	7	7	11
Tete	4	4	11	10	6	11	10	6	7
Manica	6	7	6	7	1	5	8	1	5
Sofala	8	8	5	8	3	6	5	2	3
Inhambane	7	6	4	4	10	3	2	10	6
Gaza	9	9	3	3	9	4	3	11	4
Maputo Pr.	10	10	2	2	8	2	4	4	2
Maputo Cid.	2	1	1	1	5	1	1	3	1

Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education. The province of Maputo City does not have a rural area.

Table 6 – Deprivation by dimension, provincial and urban/rural level

	Mozambique		Niassa		Cabo Delgado		Nampula		Zambezia		Tete	
	urban	**rural**	**urban**	**rural**	**urban**	**rural**	**urban**	**rural**	**urban**	**rural**	**urban**	**rural**
Head Count Ratio	0.4962	0.5691	0.4224	0.2892	0.4429	0.3552	0.499	0.5667	0.6362	0.718	0.5302	0.4015
SE	0.0164	0.0170	0.0512	0.0388	0.0866	0.0382	0.0548	0.0343	0.0680	0.0381	0.0798	0.0502
mean FM	0.5042	0.5656	0.4664	0.3893	0.4922	0.4239	0.5105	0.5607	0.5985	0.6503	0.5462	0.4808
SE	0.0114	0.0108	0.0221	0.0300	0.0588	0.0226	0.0400	0.0204	0.0485	0.0243	0.0478	0.0309
mean FS	0.339	0.6378	0.3887	0.6165	0.4358	0.623	0.4901	0.7072	0.4557	0.6743	0.479	0.6926
SE	0.0103	0.0067	0.0313	0.0301	0.0387	0.0191	0.0312	0.0155	0.0457	0.0119	0.0709	0.0190
mean FS1	0.2596	0.6545	0.3521	0.5357	0.3099	0.4538	0.4646	0.7068	0.4923	0.7756	0.4927	0.7588
SE	0.0113	0.0094	0.0269	0.0359	0.0232	0.0233	0.0325	0.0160	0.0543	0.0147	0.0958	0.0274
mean FS2	0.487	0.5119	0.3531	0.3639	0.5145	0.5183	0.5847	0.602	0.4285	0.5068	0.5599	0.5053
SE	0.0117	0.0087	0.0232	0.0267	0.0527	0.0227	0.0437	0.0159	0.0278	0.0159	0.0497	0.0323
mean FS3	0.5248	0.8421	0.6586	0.8822	0.6328	0.8862	0.7224	0.9333	0.7072	0.8581	0.6481	0.9191
SE	0.0133	0.0095	0.0334	0.0240	0.0675	0.0192	0.0385	0.0113	0.0484	0.0319	0.0779	0.0161
mean FS4	0.1317	0.3848	0.1711	0.4779	0.1951	0.4258	0.1637	0.3953	0.184	0.3768	0.1743	0.432
SE	0.0119	0.0136	0.0475	0.0493	0.0434	0.0418	0.0454	0.0357	0.0401	0.0280	0.0500	0.0410
mean FS5	0.1075	0.1195	0.1069	0.1221	0.1251	0.1099	0.1205	0.133	0.1064	0.1191	0.0866	0.1199
SE	0.0033	0.0024	0.0103	0.0079	0.0127	0.0064	0.0092	0.0046	0.0127	0.0072	0.0136	0.0075
mean FS6	0.2135	0.3699	0.2355	0.3821	0.2972	0.3969	0.2951	0.3945	0.2769	0.4039	0.2978	0.3525
SE	0.2135	0.3699	0.0221	0.0178	0.0232	0.0204	0.0161	0.0164	0.0349	0.0262	0.0371	0.0160
n	5222	5609	384	430	240	540	570	1005	336	1187	192	576

	Manica		Sofala		Gaza		Inhambane		Maputo Pr.		Maputo Cid.
	urban	**rural**	**urban**	**rural**	**urban**	**rural**	**urban**	**rural**	**urban**	**rural**	**urban**
Head Count Ratio	0.4537	0.5828	0.5005	0.6293	0.4049	0.6267	0.4836	0.6784	0.6366	0.7633	0.3615
SE	0.0505	0.0543	0.0721	0.0643	0.0429	0.0592	0.0612	0.0376	0.0276	0.0502	0.0239
mean FM	0.4818	0.5864	0.4865	0.6197	0.4459	0.5786	0.5059	0.6533	0.5914	0.6649	0.4054
SE	0.0322	0.0314	0.0474	0.0535	0.0268	0.0412	0.0427	0.0277	0.0195	0.0308	0.0182
mean FS	0.2888	0.6111	0.2849	0.6751	0.2969	0.5045	0.3566	0.469	0.2401	0.43	0.1672
SE	0.0255	0.0264	0.0318	0.0362	0.0269	0.0276	0.0545	0.0275	0.0134	0.0427	0.0066
mean FS1	0.2387	0.7322	0.2507	0.8436	0.185	0.4609	0.1875	0.3525	0.0584	0.2748	0.0323
SE	0.0314	0.0363	0.0404	0.0279	0.0276	0.0361	0.0534	0.0357	0.0058	0.0514	0.0029
mean FS2	0.3591	0.3263	0.3906	0.4639	0.4695	0.5817	0.5061	0.5738	0.5049	0.5902	0.5031
SE	0.0320	0.0230	0.0334	0.0626	0.0325	0.0287	0.0271	0.0288	0.0244	0.0520	0.0137
mean FS3	0.4919	0.823	0.5364	0.8831	0.4025	0.5494	0.56	0.6752	0.3523	0.5958	0.2768
SE	0.0350	0.0305	0.0488	0.0235	0.0363	0.0357	0.0466	0.0355	0.0219	0.0616	0.0159
mean FS4	0.0615	0.4509	0.0414	0.4196	0.1017	0.2327	0.1958	0.2272	0.1987	0.3213	0.0327
SE	0.0234	0.0519	0.0109	0.0588	0.0220	0.0348	0.0908	0.0496	0.0268	0.1058	0.0042
mean FS5	0.0775	0.0955	0.0989	0.0937	0.1164	0.1357	0.1285	0.1381	0.1096	0.0976	0.0979
SE	0.0069	0.0072	0.0147	0.0080	0.0076	0.0050	0.0089	0.0050	0.0082	0.0151	0.0069
mean FS6	0.2123	0.3311	0.1703	0.3435	0.2185	0.3567	0.2499	0.3099	0.1248	0.2145	0.1234
SE	0.0228	0.0195	0.0216	0.0367	0.0263	0.0382	0.0260	0.0200	0.0151	0.0380	0.0087
n	336	468	527	324	336	467	382	432	720	180	1199

Table 7 – Relative ranking according to the different dimensions of deprivation (1: lowest deprivation; 10: highest deprivation)

	HCR	**FM**	**FS**	**FS1**	**FS2**	**FS3**	**FS4**	**FS5**	**FS6**
Niassa urban	6	5	7	9	2	11	6	9	7
Niassa rural	1	1	16	16	4	17	21	16	18
Cabo Delgado urban	7	9	9	8	14	9	9	17	11
Cabo Delgado rural	2	3	17	11	15	19	18	11	20
Nampula urban	10	11	13	13	19	14	5	15	10
Nampula rural	13	13	21	17	21	21	16	19	19
Zambezia urban	17	17	10	14	6	13	8	8	9
Zambezia rural	20	19	18	20	13	16	15	13	21
Tete urban	12	12	12	15	16	10	7	2	12
Tete rural	4	6	20	19	11	20	19	14	16
Manica urban	8	7	4	5	3	4	3	1	4
Manica rural	14	15	15	18	1	15	20	4	14
Sofala urban	11	8	3	6	5	5	2	7	3
Sofala rural	16	18	19	21	7	18	17	3	15
Gaza urban	5	4	5	3	8	3	4	12	6
Gaza rural	15	14	14	12	18	6	13	20	17
Inhambane urban	9	10	6	4	12	7	10	18	8
Inhambane rural	19	20	11	10	17	12	12	21	13
Maputo Pr. urban	18	16	2	2	10	2	11	10	2
Maputo Pr. rural	21	21	5	7	20	8	14	5	5
Maputo Cid. urban	3	2	1	1	9	1	1	6	1

Note: HCR=Head Count Ratio; FM=Fuzzy Monetary; FS=Fuzzy Supplementary; FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education. The province of Maputo City does not have a rural area.

Appendix A.1 - Exploratory Factor Analysis results and subsequent calibration

Exploratory Factor Analysis results	Calibration results
Factor 1 *(became dimension 1)*	**1. Housing conditions**
d1. Bed	d1. Bed
d2. Proper house	d2. Proper house
d3. High-quality walls	d3. High-quality walls
d4. High-quality roof	d4. High-quality roof
d5. High-quality floor	d5. High-quality floor
d6. Has WC or latrine	d6. Has WC or latrine
d8. Telephone or mobile phone *(moved to dimension 2)*	d7. Number of rooms/(household members)^2
Factor 2 *(became dimension 4)*	**2. More widespread durable goods**
d21. Energy source for cooking	d8. Telephone or mobile phone
d22. Energy source for in-house lighting	d9. TV
Factor 3	d10. Bike
d9. TV *(moved to dimension 2)*	d11. Radio
d14. Tools	d12. Watch
d15. Electric or coal iron	d13. Motorbike
d16. Fridge or freezer	**3. Less common, expensive durable goods**
d23. Hotplate or gas ring *(moved to dimension 4)*	d14. Tools
d25. Fan or air conditioner *(moved to dimension 4)*	d15. Electric or coal iron
d26. A bank account *(moved to dimension 5)*	d16. Fridge or freezer
Factor 4 *(became dimension 2)*	d17. New or second-hand car
d10. Bike	d18. Computer
d11. Radio	d19. Printer
d12. Watch	d20. Sewing machines
Factor 5 *(became dimension 6)*	**4. Housing quality**
d7. Number of rooms/(household members)^2 *(moved to dimension 1)*	d21. Energy source for cooking
d29. Ability to read and write	d22. Energy source for in-house lighting
d30. Education level	d23. Hotplate or gas ring
Factor 6 *(became dimension 3)*	d24. Has access to safe water
d17. New or second-hand car	d25. Fan or air conditioner
d18. Computer	**5. Income-related deprivation**
d19. Printer	d26. A bank account
d20. Sewing machines	d27. A formal or informal job
Factor 7	d28. More than two meals per day
d27. A formal or informal job *(moved to dimension 5)*	**6. Health and education**
Factor 8	d29. Ability to read and write
d13. Motorbike *(moved to dimension 2)*	d30. Education level
d28. More than two meals per day *(moved to dimension 5)*	d31. Chronic illness
Factor 9	d32. Child malnutrition
d24. Has access to safe water *(moved to dimension 4)*	
Factor 10 *(became dimension 6)*	
d31. Chronic illness	
d32. Child malnutrition	

Appendix A.2 – Sensitivity Analysis

Here we present the results of a sensitivity analysis that is performed to assess the robustness of the selection of the items for the subsequent Fuzzy Set poverty analysis. This is done by implementing a leave-one-out process in which one item at the time is excluded from the analysis. After the exclusion, an Exploratory Factor Analysis is run and we check for the number of changes that occur with respect to the baseline analysis. In particular, we check whether the number of dimensions (factors) varies, and how many items pass from a factor to another (Columns 1 and 2 of Table A.2.1).

Afterwards, the same calibration process shown in Appendix A.1 is applied, and we report the summary statistics that describe the goodness of fit of such calibration: Goodness of Fit (GFI), GFI adjusted for degrees of freedom (AGFI), Root Mean Square Residual (RMR), Parsimonious GFI, and Root Mean Squared Error of Approximation (RMSEA).

Finally, in Table A.2.2 the average values of deprivation for each supplementary dimension are presented.

From the results of the sensitivity analysis it can be seen that for most of the items the procedure is quite robust. The number of dimensions remains fairly stable in that for 19 items it does not change at all compared with the baseline, while for the remaining 13 items it decreases at most by one dimension, namely from 8 to 7. As for the number of items that change their position from one dimension to another in each run, we have that for 13 items no change is observed. For other 16 items this number is comprised between 1 and 3 changes, while only the exclusion of "proper house", "high-quality roof" and "child malnutrition" generates more sensible modifications of respectively 11, 8, and 4 changes.

All the GFI statistics derived from the calibration procedure and the Confirmatory Factor Analyses yield satisfactory results, in line with the baseline analysis presented in the study.

The average values of deprivation for supplementary dimensions are also shown to be robust to the exclusion of most items (Table A.2.2). For FS1 (housing conditions), only the exclusion of improved sanitation related variables determines a significant increase in its average, while for FS2 (more widespread durable goods) the biggest change comes from the exclusion of “owning a radio”. When we do not consider the possession of an electric or coal iron, then the FS3 dimension (less common, expensive durable goods) increases its average deprivation value by more than 10 percentage points. For the dimension FS4 (housing quality) the items that cause greater modifications are “access to safe water” and “energy source for in-house lighting”. While the former causes deprivation to increase by 13 percentage points, the latter decreases it from 0.31 to 0.21. The most relevant changes, however, occur in the dimension FS5 (income-related deprivation). Here, the exclusion of certain items - the ones related to health and education - brings about a decrease in deprivation of about nine percentage point. Furthermore, when the variable “formal or informal job” is excluded from the analysis, the deprivation value for this dimension increases from 0.12 to 0.64. Even though it leaves unaltered the values of all the other dimensions, this represents a very big change, and for this reason we analysed it in greater detail throughout the study. The dimension FS6, instead, appears to be more robust to the exclusion of items. The most significant changes for all dimensions and summary statistics are bold-marked in Tables A.2.1 and A.2.2.

Table A.2.1 – Sensitivity Analysis, summary statistics

Excluded item	**Number of dimensions**	**Number of items changing dimension**	**GFI**	**AGFI**	**RMR**	**Parsimonious GFI**	**RMSEA**
-	8	**-**	0.834	0.8086	0.0807	0.77	0.0748
Has access to safe water	**7**	2	0.831	0.8042	0.0828	0.7649	0.077
New or second-hand car	8	-	0.8365	0.8105	0.0806	0.7699	0.0759
Has WC or latrine	8	-	0.837	0.8112	0.0816	0.7704	0.0758
Bike	8	2	0.842	0.8168	0.08	0.775	0.0744
Proper house	**7**	**11**	0.8389	0.8138	**0.0987**	0.774	0.0799
Tools	8	-	0.8326	0.806	0.0802	0.7663	0.0764
Telephone or mobile phone	8	-	0.8428	0.8179	0.0805	0.7758	0.0747
Computer	**7**	3	0.8488	0.8248	0.078	0.7813	0.0712
A bank account	8	-	0.8399	0.8144	0.0786	0.773	0.075
Energy source for cooking	8	-	0.8385	0.8128	0.0804	0.7718	0.0752
Sewing machines	**7**	3	0.8308	0.8039	0.0829	0.7647	0.0771
Electric or coal iron	8	-	0.8438	0.819	0.0785	0.7767	0.0746
Hotplate or gas ring	8	-	0.844	0.8192	0.0777	0.7769	0.074
Fridge or freezer	8	-	0.8416	0.8165	0.0764	0.7747	0.0736
Energy source for in-house lighting	**7**	3	0.8366	0.8107	0.0811	0.7701	0.0758
A formal or informal job	**7**	2	0.8314	0.8046	0.0826	0.7652	0.0769
Ability to read and write	**7**	3	0.8466	0.8223	**0.064**	0.7793	0.0719
Bed	8	1	0.84	0.8146	0.0802	0.7731	0.075
Chronic illness	**7**	-	0.8298	0.8028	0.0832	0.7638	0.0773
Motorbike	8	-	0.8316	0.8048	0.0827	0.7654	0.0769
High-quality walls	**7**	3	0.8423	0.8173	0.0819	0.7753	0.0746
Watch	8	1	0.8359	0.8098	0.0816	0.7694	0.0761
More than two meals per day	8	-	0.8335	0.8071	0.0824	0.7672	0.0766
High-quality floor	8	2	0.8457	0.8212	0.0818	0.7784	0.074
Radio	8	1	0.841	0.8158	0.0805	0.7741	0.0749
Printer	**7**	3	0.8459	0.8214	0.0791	0.7786	0.0721
Number of rooms/(household members)^2	**7**	3	0.8362	0.8102	0.0814	0.7697	0.0756
Education level	**7**	3	0.8511	0.8274	**0.0639**	0.7833	0.071
High-quality roof	**7**	**8**	0.8557	0.8328	0.0736	0.7876	0.0701
TV	8	-	0.8431	0.8181	0.0787	0.776	0.0744
Fan or air conditioner	8	1	0.842	0.8168	0.078	0.775	0.0743
Child malnutrition	8	**4**	0.8328	0.8063	0.0821	0.7666	0.0766

Table A.2.2 – Sensitivity Analysis, deprivation values for supplementary dimensions

Excluded item	FS1	FS2	FS3	FS4	FS5	FS6
-	0.5345	0.5043	0.7457	0.3079	0.1158	0.3224
Has access to safe water	0.5302	0.5002	0.7419	**0.4396**	0.1119	0.318
New or second-hand car	0.5343	0.5041	0.7445	0.3077	0.1156	0.3221
Has WC or latrine	**0.6039**	0.5042	0.7456	0.3077	0.1157	0.3222
Bike	0.533	0.5277	0.7451	0.3065	0.1145	0.3208
Proper house	0.5339	0.5043	0.7457	0.3079	0.1158	0.3224
Tools	0.5343	0.5041	0.7356	0.3077	0.1156	0.3221
Telephone or mobile phone	0.5349	0.5142	0.7458	0.3082	0.1162	0.3228
Computer	0.5343	0.5042	0.7455	0.3077	0.1157	0.3222
A bank account	0.5342	0.5041	0.7434	0.3067	0.1027	0.3221
Energy source for cooking	0.5343	0.5041	0.7435	0.2929	0.1157	0.3222
Sewing machines	0.534	0.5039	0.7433	0.3075	0.1154	0.3219
Electric or coal iron	0.536	0.5058	**0.8861**	0.3092	0.1172	0.3239
Hotplate or gas ring	0.534	0.5039	0.7433	0.3095	0.1154	0.3219
Fridge or freezer	0.5343	0.5041	0.7356	0.3077	0.1157	0.3222
Energy source for in-house lighting	0.5347	0.5046	0.7436	**0.2125**	0.1161	0.3226
A formal or informal job	0.5226	0.493	0.7391	0.2966	**0.6386**	0.3102
Ability to read and write	0.5344	0.5043	0.7435	0.3069	**0.0309**	0.3295
Bed	0.5414	0.5044	0.7457	0.308	0.1159	0.3225
Chronic illness	0.5315	0.5014	0.7424	0.3043	**0.0298**	**0.3496**
Motorbike	0.5344	0.5033	0.7456	0.3078	0.1158	0.3223
High-quality walls	0.5296	0.5042	0.7456	0.3078	0.1157	0.3222
Watch	0.5345	0.4994	0.7457	0.3079	0.1159	0.3224
More than two meals per day	0.5344	0.5043	0.7435	0.3069	**0.031**	0.3223
High-quality floor	0.5294	0.5043	0.7457	0.3078	0.1158	0.3223
Radio	0.5343	**0.5377**	0.7456	0.3077	0.1157	0.3222
Printer	0.5343	0.5042	0.7457	0.3077	0.1157	0.3222
Number of rooms/(household members)^2	0.5488	0.5023	0.7449	0.306	0.1139	0.3202
Education level	0.5342	0.5041	0.7434	0.3067	**0.0308**	**0.2946**
High-quality roof	0.5319	0.5043	0.7457	0.3079	0.1158	0.3224
TV	0.5344	0.5006	0.7456	0.3078	0.1157	0.3222
Fan or air conditioner	0.5341	0.5039	0.7434	0.3071	0.1155	0.3219
Child malnutrition	0.534	0.5039	0.7433	0.3065	**0.0307**	0.3247

Note: FS1=housing conditions; FS2=more widespread and affordable durable goods; FS3=less common, more expensive durable goods; FS4=housing quality; FS5=income-related deprivation; FS6=health and education. The province of Maputo City does not have a rural area.

Appendix B – Jackknife Repeated Replication (JRR) for variance estimation

The Jackknife Repeated Replication (JRR) is one of a class of methods for estimating sampling errors from comparisons among sample replications which are generated through repeated resampling of the same parent sample.

We prefer to use the JRR instead of other methods like bootstrapping as it is less computer-intensive, in that the JRR only provides estimates of the variance of the point estimator, while bootstrapping estimates its whole distribution. Moreover, the bootstrap may be seen as a random approximation of the general version of the JRR we chose, but while it gives different results when repeated on the same data, the jackknife gives exactly the same result each time (Shao and Tu, 1995).

In the JRR procedure each replication needs to be a representative sample in itself and reflect the full complexity of the parent sample. However, in so far as the replications are not independent, special procedures are required in constructing them so as to avoid bias in the resulting variance estimates. We prefer the JRR to similar methods such as the Balanced Repeated Replication because the JRR is generally simpler and more flexible.

Originally introduced as a technique of bias reduction, the Jackknife method has by now been widely tested and used for variance estimation (Durbin, 1959). Efron and Stein (1981) provide a discussion of the Jackknife methodology. For a landmark empirical study of such applications, see Kish and Frankel (1974). For a general description of JRR and other practical variance estimation methods in large-scale surveys, see Verma (1993). For a comparative analysis between JRR and Taylor linearisation methods see Verma and Betti (2011).

The JRR variance estimates take into account the effect on the variance of aspects of the estimation process which are allowed to vary from one replication to another. In principle this can include complex effects such as those of imputation and weighting. But it has to be noted that often in practice it is not possible to repeat such operations from scratch at each replication.

The basic model of the JRR for application in the context described above may be summarised as follows. Consider a design in which two or more primary units have been selected independently from each stratum in the population. Within each primary sampling unit (PSU), subsampling of any complexity may be involved, including weighting of the ultimate units.

In the ‘standard’ version, each JRR replication can be formed by eliminating one sample PSU from a particular stratum at a time, and increasing the weight of the remaining sample PSU's in that stratum appropriately so as to obtain an alternative but equally valid estimate to that obtained from the full sample.

The above procedure involves creating as many replications as the number of primary units in the sample. The computational work involved is sometimes reduced by reducing the number of replications required. For instance, the PSUs may be grouped within strata, and JRR replications formed by eliminating a whole group of PSUs at a time. This is possible only when the stratum contains several units. Alternatively, or in addition, the groupings of units may cut across strata. It is also possible to define the replications in the standard way (‘delete one-PSU at a time Jackknife’), but only construct and use one subsample at a time.

In the kind of multistage samples encountered in most national household surveys, the standard JRR method without such grouping of units can be applied. However, one common situation in which grouping of units is unavoidable is when the sample or a part of it is a direct sample of ultimate units or of small clusters, so that the number of replications under ‘standard’ JRR is too large to be practical. Normally, the appropriate procedure to reduce this number would be to form new computational units by forming random groupings of the units within strata. The presence of small and variable-sized PSUs may also require some grouping in the practical application of the procedure.

Briefly, the standard JRR involves the following.

Let u be a full-sample estimate of any complexity, and $u_{(hi)}$ be the estimate produced using the same procedure after eliminating primary unit i in stratum h and increasing the weight of the remaining (a_h-1) units in the stratum by an appropriate factor g_h (see below). Let $u_{(h)}$ be the simple average of the $u_{(hi)}$ over the a_h values of i in h. The variance of u is then estimated as:

$$var(u)=\sum\left[(1-f_h)\cdot\frac{a_h-1}{a_h}\cdot\sum\left(u_{(hi)}-u_{(h)}\right)^2\right]. \qquad (5)$$

A major advantage of a procedure like the JRR is that, under quite general conditions for the application of the procedure, the same and relatively simple variance estimation formula (5) holds for u of any complexity.

A possible variation which may be mentioned is to replace $u_{(h)}$, the simple average of the $u_{(hi)}$ over the a_h replications created from h, by the *full-sample* estimate u:

$$var(u)=\sum\left[(1-f_h)\cdot\frac{a_h-1}{a_h}\cdot\sum\left(u_{(hi)}-u\right)^2\right]. \qquad (5')$$

This version tends to provide a 'conservative' estimate of variance, but normally the difference to (5) is small. We have used form (5) in the empirical analysis for Mozambique.

Concerning the re-weighting of units retained in a stratum after dropping one unit, normally the factor g_h is taken as (6.a), but for reasons noted below, we propose the form in (6.w):

$$g_h=\frac{a_h}{a_h-1}\ , \qquad (6.a)$$

$$g_h=\frac{w_h}{w_h-w_{hi}}\ , \qquad (6.w)$$

where $w_h=\sum w_{hi}$, $w_{hi}=\sum w_{hij}$, the sum of sample weights of ultimate units j in primary selection i. Note that (6.a) gives the variance of a simple aggregate, while (6.w) gives the corresponding (lower) variance of a mean, or of total as a ratio estimate.

Form (6.w) is used throughout in our empirical analysis here. This form retains the total weight of the included sample cases unchanged across the replications created -the same total as that for the full sample-. With the sample weights scaled such that their sum is equal (or proportional) to some external more reliable population total, population aggregates from the sample can be estimated more efficiently, often with the same precision as proportions or means.

Health Provider Choice and Implicit Rationing in Health Care: Evidence from Mozambique

Vincenzo Salvucci[a]

[a] Dipartimento di Economia Politica e Statistica - Facoltà di Economia "Richard M. Goodwin", Università degli Studi di Siena - Piazza S. Francesco, 7 - 53100 Siena, Italy

salvucci3@unisi.it

Abstract

An analysis of health provider choice in Mozambique together with a model of implicit rationing in health care is presented. We make use of the Mozambican Household Survey on Living Conditions (IAF) 2002-03 and 2008-09. While the former is well-known and has been used in other studies, the IAF 2008-09 is a newly released dataset. Results that are substantially different from similar analyses on older Mozambican data are obtained. In particular, we find that the role of income is more important in 2002-03 and in 2008-09 than it was in 1996-97. In consequence, we further expand the analysis by studying how socioeconomic disparities leads to implicit rationing in health care. When using the IAF 2008-09 only, moderate evidence that some sort of health care rationing is in place is found, i.e. someone does not seek care because he/she foresees he/she will not be granted quality health care access or treatment.

JEL classification: I11
Keywords: health care demand, rationing, Mozambique

1. Introduction

The study of health related behaviour in developing countries is gaining further interest in economics because of its implications in economic activity and outcomes. In this article an analysis of health provider choice in Mozambique together with a model of implicit rationing in health care is presented. Concerning the analysis of health provider choices in Mozambique it is found that the role of income is more important in 2002-03 and in 2008-09 than it was in 1996-97 (Lindelow, 2005). In consequence we further expand the analysis by studying how socioeconomic disparities lead to implicit rationing in health care.

When fees for health care services are nominal or nihil but the system is resource poor - as is generally the case in Sub-Saharan Africa, and in particular in

Mozambique - the allocation of these services is often unequal and regressive (Stierle et al., 1999; Kapiriri and Martin, 2007), mostly favouring the urban areas and the elites, who find queue-avoiding mechanisms (Pinho, 2008; Rosen et al., 2010).

Hence, an analysis of health care rationing is first carried out. Then, a probability of accessibility to medical services in Mozambique is obtained, with special emphasis on non-price determinants of health care allocation and use. Previous modelling only considers the probability to seek treatment from a certain provider given personal, household, and community characteristics. In order to fully identify the rationing mechanisms at work, the probability of *receiving* a good treatment, conditional on seeking care, is also modelled here. As the literature on queue jumping and rationing illustrates, for those who cannot jump the queue, getting to the chosen health care provider does not ensure receiving the desired treatment. This is modelled by means of a probability of receiving satisfying treatment that depends primarily on economic status, education, and distance to the health provider. Using the IAF 2008-09 only, moderate evidence that some sort of health care rationing is in place is found, i.e. someone does not seek care because he/she foresees he/she will not be granted quality health care access or treatment.

The model used in the first part of the article is a standard random utility model with dichotomous choice between care and self-care (Gertler and Van Der Gaag, 1990) as the dependent variable, which is estimated by means of a probit model. For the polytomous choice of the health care provider, a more flexible model is used (Dow, 1996; Lindelow, 2005), and estimated by means of a multinomial logit model.

Two cross-sectional surveys, the Mozambican Household Survey on Living Conditions (*Inquerito aos Agregados Familiares sobre Orçamento Familiar*, IAF) 2002-03 and 2008-09 are employed. While the former is well-known and has been used in other studies, the IAF 2008-09 is a newly released dataset. These surveys are not only richer and more detailed compared to those implemented in 1996-97, but also yield more robust results and deeper insights into households' health related behaviour.

Section 2 presents the theoretical model used for the analysis of both the binary choice care/no care and of the multinomial choice of the provider. In Section 3 the datasets used is presented and Section 4 introduces the econometric specification and discusses the results. In Section 5 the problem of implicit rationing is introduced while a possible strategy to devise such rationing is discussed. Section 6 concludes.

2. Theoretical model

Following relevant works on provider choice and health related behaviour in developing countries (Mwabu, 1984; Akin et al., 1986; Dor and Van Der Gaag, 1993; Ellis et al., 1994; Akin et al., 1995; Glick et al., 2000; Sahn et al., 2003; Ssewanyana et al., 2004; Lindelow, 2005), the decision to seek care is analysed by means of a random utility model. This model is used to study the dichotomous decision to seek care or not, the dichotomous decision to seek formal care or not, and the polytomous decision about the health provider where to seek care. The random utility model seems particularly suitable for a developing country framework where more sophisticated models meet practical difficulties in their application (Gertler et al., 1987; Gertler and Van Der Gaag, 1990).

2.1 The choice to seek care or not and the choice to seek care from a certain provider

For the dichotomous case about seeking care or not, we take the expected utility of seeking or not seeking care -i.e. choosing self-care-, respectively equal to U_c and U_{sc}. These depend on expected health status when seeking or not seeking care, H_c and H_{sc}, and on consumption of goods other than health, C_c and C_{sc}.

$$U_i = U(H_i, C_i), \quad (1)$$

where $i = c, sc$. C_c and C_{sc} are modelled as residual consumption after the decision to seek care or not. H_c and H_{sc} will depend on individual, household and community characteristics, while C_c and C_{sc} can be derived from the budget constraint as:

$$C_c = Y - P_c^*, \quad (2)$$

$$C_{sc} = Y, \quad (3)$$

P_c^* is the cost of seeking care, and in turn P_c^* is composed of a direct payment, P_c, plus the opportunity cost of time (w) multiplied by the time spent to reach the health provider (T_c). Hence, $P_c^* = P_c + w \cdot T_c$ and $P_{sc}^* = 0$. Each individual will maximise U_i:

$$U_i^* = max(U_s, U_{sc}), \quad (4)$$

In the polytomous choice, the number of possible outcomes is increased by extending it to self-care, health post, health centre, hospital, other formal providers and traditional medical practitioner (TMP). In this case the expected utility of seeking care from a certain provider, U_j, is considered, where j = (*self-care, health post, health centre, hospital, other formal providers, TMP*).

U_j depends on expected health status when seeking care from provider j, H_j, and on consumption of goods other than health, C_j. C_j is again modelled as residual consumption after the decision to seek care from provider j. H_j will depend on individual, household and community characteristics, while C_j can be derived from the budget constraint as:

$$C_j = Y - P_j^*, \tag{5}$$

P_j^* is the cost of seeking care from provider j, and in turn P_j^* is composed of a direct payment, P_j, plus the opportunity cost of time (w) multiplied by the time spent to reach the health provider j (T_j). Hence $P_j^* = P_j + w \cdot T_j$ and $P_{self\text{-}care}^* = 0$.

Each individual will maximise U_j:

$$U_j^* = max(U_{self\text{-}care}, U_{health\ post}, U_{health\ centre}, U_{hospital}, U_{other\ formal\ providers}, U_{TMP}) \tag{6}$$

In both the dichotomous and the polytomous cases, the maximisation procedure yields a system of demand function in the form of probabilities. These systems can be estimated -respectively- by means of a probit model and of a polytomous choice model such as the multinomial logit model.

3. Data

The datasets used are the Household Survey on Living Conditions 2002-03 and 2008-09 (*Inquerito aos Agregados Familiares sobre Orçamento Familiar*, IAF), which are nationally representative household surveys conducted by the National Institute of Statistics (INE). The IAF 2002-03 was conducted during 2002 and 2003, from July to June, while the IAF 2008-09 was conducted from August 2008 to September 2009. The surveys have a stratified structure with three steps of selection: i) selection of the primary sampling units (PSUs); ii) selection of the enumeration areas within the PSUs; iii) selection of the households within the enumeration areas. Twelve households were selected in each rural area and nine in each urban area. 21 strata were constructed, one for each urban/rural sample of the 11 provinces (the province of Maputo city does not have a rural area). The final sample size of the IAF 2002-03 amounts to 44,083 individuals in 8,727 households: 4,020 urban households and 4,707 rural households. The IAF 2008-09 instead has a sample size of 51,177 individuals in 10,832 households, divided into 5,223 urban households and 5,609 rural households. Both surveys are representative at the national, regional (North, Centre, South), provincial and urban/rural level. Sampling weights are provided in the survey dataset.

The surveys include information on general characteristics of the individuals and of the households, home consumption, transfers and gifts, as well as on daily, monthly, and durable goods final consumption expenditures. Community surveys are also available, but only for rural areas, containing information on community characteristics. Supplementary information for the IAF 2002-03 can be found in MPF (2004) and in INE (2004), while for the IAF 2008-09 it can be found in INE (2010).

With regards to health related issues, the IAF 2008-09 has a more detailed health section, but both surveys contain information on permanent illnesses, on past episodes of sickness, on the choice of health care providers, and relative satisfaction with the treatment. Moreover, in the community questionnaires there is a comprehensive information on distances to different health care providers.

4. Econometric model, variables and estimation results

Starting from the model outlined in Section 1, we use different econometric specifications when studying the dichotomous case of care/self-care and the polytomous case of health care provider choice. In the dichotomous case, a simple linear equation is used, following among others Mwabu (1984); Akin et al. (1986); Akin et al. (1995). From Equation (1) we thus specify:

$$U_i = a_0 H_i + a_1 C_i + e_i \,, \qquad (7)$$

Where H_i depends on a vector of controls, Z, while C_i depends on income, Y, and time distance, T_i. As shown in Lindelow (2005, pp. 439-440), Equation (7) can be reparameterised and estimated by means of a probit model. In this way the characteristics influencing the choice to seek care or not are analysed.

For the polytomous choice problem, instead, the empirical model is extended with the inclusion of cross-price terms in each provider's equation[1]. In this case, a multinomial logit model is used in order to assess the probability to seek care from each provider[2]. The multinomial logit model is a valid alternative provided that the independence of irrelevant alternatives (IIA) assumption holds, otherwise other specifications should be used. In what follows the IIA assumption is tested by means of a Hausman test to make sure that estimates are reliable. All econometric

[1] Dow (1996) provides justifications for this and for other potential extensions.

[2] Again, details on the reparametrisation and on the possibility to estimate such a model by means of a multinomial logit model can be found in Lindelow (2005, pp. 450-453).

estimations are corrected for the stratified survey design and weighted according to the sampling weights provided in order to ensure representativeness.

4.1 Variables used and results

In the econometric estimation of the models outlined in Section 1 and Section 3 the dependent variable is first the dichotomous variable care/self-care and then a polytomous variable which has as possible outcomes the different health care provider alternatives[3].

For independent variables several individual, family, and community characteristics are considered. As price variables, only the time-price is considered since the official fees are either very low or unevenly paid by the population throughout the country (Lindelow et al., 2004). The opportunity cost of time that is employed is the median -computed for each community- of per capita hourly consumption: this is not as a good measure of opportunity cost as is wage, but for many developing countries, and in particular for Mozambique, data on wages are either missing or non-existent. This is mainly due to the fact that few people work in the formal sector. In contrast, hourly consumption, though suffering from possible endogeneity problems, is available for the whole sample.

Individual income is measured by (real) per capita daily consumption (household consumption divided by household size). This is the measure of income that is also used in official analyses of poverty carried out by the Mozambican government. A composite measure of consumption that takes into account daily and monthly expenditures, durable goods use and rent, self-consumption, in kind earnings is used. All these different sources of consumption are averaged out on a daily basis, providing a good measure of economic status for Mozambican households. This measure of income also takes into account the inflation occurred during the implementation of the surveys, the different values of the Metical - the Mozambican currency - over different periods of the year, and spatial differences in price levels among different provinces and areas (rural/urban)[4].

The information on illness episodes, decision to seek care and where, problems experienced during the medical consultation and time needed to reach the closest

---

[3] We do not take into account the whole range of possible options since some of them are chosen by very few ill individuals. So we put into one group all the other formal providers others than health post, health centre, hospital and traditional medical practitioner.

[4] 13 regional poverty lines have been computed for Mozambique, in order to take into account the differences in price levels in the various provinces and rural/urban areas (Simler et al., 2004).

health facility is also used as a key variable. The other variables in the regressions include standard controls: individual characteristics such as age, gender, education, ability to read and write; family characteristics like household size and household members per room; community characteristics as accessibility through public transport and conditions of the main road, and which health care provider is present and how far it is from the community[5]. Descriptive statistics for the variables used are presented in Table 1.

4.2 Seeking care or not – Probit results

In Table 2 the marginal effects of the probit results for the dichotomous decision to seek care or not are introduced. One can notice that some sets of variables such as socioeconomic status - expressed by the logarithm of per capita daily consumption (Model 1), and quintiles of per capita daily consumption (Model 2) -, are strongly associated with seeking care. Though being stronger for 2002-03, this association is present in 2008-09, too[6]. Computed marginal effects show an increase of about 9% and 7% in the probability of seeking care, respectively for 2002-03 and 2008-09.

Such result contrasts with previous findings by Lindelow (2005) on Mozambican data for 1996-97, but in line with the literature on health care seeking (Dor et al., 1987; Sahn et al., 2003; Ssewanyana et al., 2004). This may be due to differences in the sample used - urban and rural households in the present article versus only rural households in Lindelow (2005) - or it might reflect a change in the behaviour of Mozanbican society between 1996-97 and 2002-03[7].

Higher level of education is also associated with a greater probability of seeking care: this is reflected by the coefficient on the ability to read and write, which appears to be more important than formal educational training in 2002-03, and by years of formal education in 2008-09. Moreover, distance to the closest health care provider is also associated with the decision to seek care. This effect is in line with existing literature (Lindelow, 2005), coefficients are highly significant both in 2002-03 and in 2008-09, and coefficient directions are as expected.

[5] The distance variables are used both alone and in interaction with the shadow price of time to obtain the prices of the different providers. These are expressed respectively by dummies in 2002-03 and by a continuous variable in 2008-09. Some of the community-level variables are only available for rural communities. Values for urban areas are imputed on the basis of information available through other sources.

[6] Both the Wald and LR test reject dropping the expenditure-related variables.

[7] In the sense that maybe the better-off have developed a greater attention to health care in these years.

4.3 Seeking formal care or not – Probit-with-selection results

A slightly different issue related to the choice of seeking care is the choice of seeking care from a formal provider, i.e. excluding self-care and traditional medical practitioners. In order to study this problem, a probit-with-selection model, checking for the probability to seek formal care provided that one seeks care, is first run[8]. In both cases, estimation results (not shown) are very similar to the ones presented in Subsection 4.2. Some differences are found for 2002-03, where the education and female dummies have a statistically significant positive coefficient. That more educated individuals are more likely than other categories to seek care from formal health care providers is a plausible and well-established result. On the other hand, the fact that ill female individuals seek more formal care is an interesting finding that deserves to be investigated more in depth[9]. One account of this finding depends certainly on the fact that women are more likely to seek formal sector care for pre-natal and birthing care which typically represents a large proportion of all health visits in poor countries.

4.4 Choosing where to seek care – Multinomial logit results

The following analysis deals with the choice of health care provider, as outlined in Section 1. Marginal effects computed in Table 3 and Table 4 show that the interpretation of the results here is less clear than before.

In 2002-03 socioeconomic status has a positive and significant effect for health centres and hospitals. Moreover, ability to read and write maintains the expected positive sign for health centre, hospital, and other formal providers. In addition, having 6 or more years of education plausibly decreases the probability of self-care or to seek care from a traditional medical practitioner[10]. Concerning distance, a negative effect is observed for the health post and less significantly for the health centre alternatives, while a significant positive effect is displayed for self-care or traditional care.

In 2008-09 socioeconomic status has the expected positive sign only for the health centre and other formal provider alternatives[11]; while education and ability to read

[8] While for 2002-03 there is evidence of a selection bias, the same is unattested for 2008-09, for which reason in this second case a simple probit model was run.

[9] A possible explanation comes from standard utility model frameworks such as Acton's, where individuals with higher opportunity costs of time (generally adult males) demand less medical care.

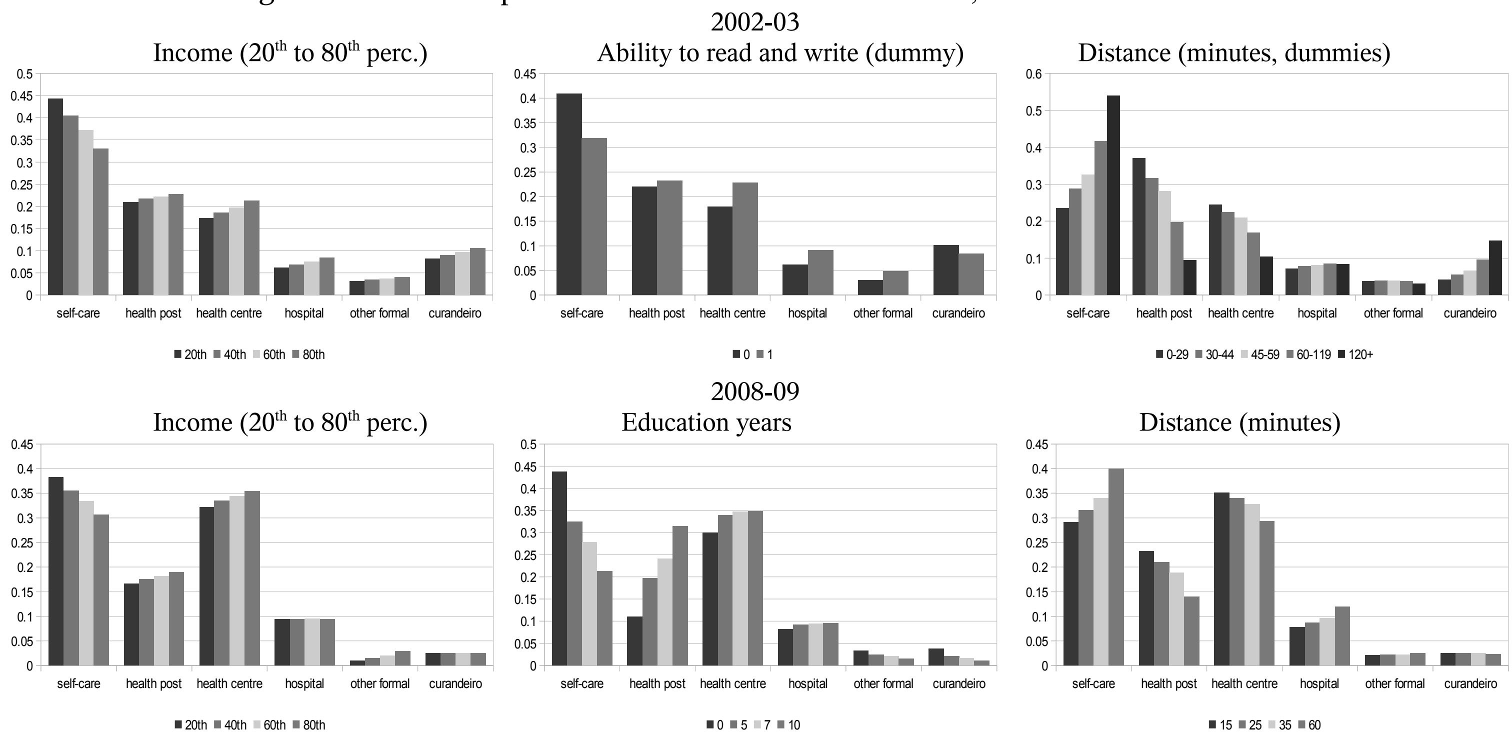
[10] As introduced in Subsection 4.3, women seek care from traditional medical practitioners less.

[11] Though the evidence is mixed when consumption is expressed as quintile dummies.

and write positively influence the choice of seeking care from a health centre and other formal providers respectively. Finally, living in a rural area or in an inaccessible community is negatively associated with seeking care from hospital, other formal providers, and a health centre.

Another observation concerns a change in health seeking behaviour between 2002-03 and 2008-09. Looking at the percentage of ill individuals seeking care from a traditional medical practitioner in 2002-03 shows that this figure is four times as large as the same figure in 2008-09 (14.5% of ill individuals in 2002-03 and 3.75% of ill individuals in 2008-09). Despite only the last consultation being registered in the survey, this suggests that ill individuals are becoming more prone to seek care from formal providers. This is probably influenced by the increase in the number of rural health providers which occurred between the two dates. A graphical analysis of marginal effects (for relevant variables) is found in Figure 1.

Figure 1 – Predicted probabilities for the effect of income, education and distance



Note: Each column represents the probability to choose the outcome indicated in the x-axis at different points.
Source: author's calculations.

4.5 Robustness checks

This subsection describes a few robustness checks performed on our base model. A first robustness check consists of running the same multinomial logit regressions but using a constructed measure of wage as the shadow price of time. In particular, information about wages in the household questionnaire is used to compute the median wage for each province. Given that a complete information about wages was only present in the IAF 2008-09, this robustness check is only undertaken for this year. What emerges from this check (not shown) is that the coefficient of the income variable is still significant for the health post, health centre, and other formal providers alternatives. While the coefficients for education and distance are only significant for the health post alternative, ability to read and write is found to influence the choice of health centre and other formal providers.

In another robustness check, an alternative specification of income, namely consumption per adult equivalent, is used instead of consumption per capita. The former takes into account possible economies of scale within the household. For 2008-09, results (not shown) say that income remains significant for the choice of seeking care from health post, health centre and other formal providers. Again, the coefficients for education and distance are only significant for the health post alternative whereas ability to read and write seems to influence the choice of health centre, and other formal providers.

5. Implicit rationing in health care access

In Mozambique, the vast majority of health care providers are public and fees are set very low with many exemptions for children under-five and other categories (Lindelow et al., 2004). However, especially in rural areas, it is not easy to provide health care service that is good and accessible to all. Lack of medicine, lack of trained staff, long waiting times, requests of informal payments have been frequently reported as existing problems (Newman et al., 1998; INE, 2004; Posse and Baltussen, 2009). In such circumstances it is likely that some ill individuals do not receive a full or successful treatment.

Given the above, there are at least two types of possible rationing mechanisms in place. The first, and more obvious, is the one due to objective obstacles to health care access, for example, occurring on a first-come first-served basis. The other is more subtle and is due to the fact that personal characteristics of the ill may also affect the way in which health care is provided. In particular, social status and distance (cultural and physical distance) contribute to influence health access and care.

This section is thus devoted to the analysis of such mechanism, which we call "implicit rationing", i.e. someone do not seek care because he/she foresees he/she will not be granted quality health care access or treatment.

This is not new to qualitative anthropological and economic studies on health care access (Ahmed et al., 2000; Khan et al., 2000; Griffiths and Stephenson, 2001; Baltussen et al., 2002; Ager and Pepper, 2005; Goins et al., 2005; Baltussen and Ye, 2006; Chibwana et al., 2009) and was confirmed by repeated discussions had with patients, nurses, missionaries, and NGO professionals in Mozambique. Ill individuals try to seek care from health care providers, but sometimes they do not receive the treatment they expect. As a consequence of that they no longer seek care from such providers. Nonetheless, we experienced that expectations about access and quality of treatment were not the same among different individuals. In particular, the poorer and those living far away from the chosen health provider, and those with lower social status, all had much worse expectations about their possibility to be granted quality health care access and treatment. The richer or those of higher social status seemed more able to find queue avoiding mechanisms and to receive good treatment. If one knows that for his personal characteristics he is more likely to have problems in access to health care providers, he will probably decide to seek care only when it is strictly necessary. Many individuals we interacted with clearly took the risk of "bad" treatment into account. Hence, they thought twice before walking many hours to the health care provider or paying for public transport, and wasting a workday. This was also confirmed by respondents who had experienced sickness episodes but did not seek care at all or did not seek formal care. In what follows, only two quotes are presented, which summarise most of the answers got in informal interviews at different health providers in the districts of Erati and Monapo.

"I went to the hospital to get drugs for leprosy, they told me to come back on the next day. When I went on the next day, they told me: 'Come next week'. I am not going anymore. Every time I go, I have to walk many hours and waste a working day" (a peasant woman).

"When they go to the hospital with me, then the nurse finds the medicines; when they go on their own, he says there are no medicines" (a missionary).

In what follows, a model that introduces a probability to receive good medical treatment and estimates this probability for the sample of ill individuals is presented. Relying on the estimated probability of receiving good treatment, we try to understand whether some implicit rationing exists. Consider the following expected utility introduced in Section 1, conditional on receiving health care from provider *j*:

$$U_j = U(H_j, C_j), \quad (8)$$

where H_j is again the expected health status after receiving treatment from *j*, and C_j is the consumption of goods other than health, after paying for the treatment from provider *j*.

The model is then extended to the case when receiving good treatment is not certain, but it happens with probability p_i. p_i is modelled as the individual probability of receiving a good treatment and depends on education, income, demographic characteristics, area of residence, distance to the facility, and other controls already presented in Section 4. The distance to the facility is interpreted as a proxy for network effects, as we expect that if one lives far away from a health facility the probability of knowing someone working in the facility is lower, and in turn it is lower the probability of receiving a good treatment. On the other hand, a positive effect coming from higher education and income is expected.

If the expected health status, H_j, depends on p_i, then we have that:

$$H_j = p_i \cdot H_j^+ + (1-p_i) \cdot H_0, \quad (9)$$

where H_j^+ is the expected health status in case of good treatment, while H_0 is the expected health status in case of bad treatment (in order to make things simple we assume H_0 is equal to the expected health status when one chooses self-care).

We assume that each provider has a certain quality which is based on objective characteristics of the provider[12]. All else being equal, the fact that having a good treatment is not certain increases the probability of choosing self-care. For example, if an ill individual is not certain that seeking care from a health provider would make a difference with self-care (because he/she expects that for his/her characteristics he/she will not be granted a quality access or treatment), then this increases the probability of choosing self-care.

In the health section of the surveys' questionnaire respondents were asked to answer whether they experienced problems in access or treatment and which problem they

[12] In this feature our model differs from the one in Dor et al. (1987), which instead models providers' quality as depending on individual characteristics.

had. Choices were lack of hygiene, long waiting time, lack of qualified staff, too expensive, lack of medicines, unsuccessful treatment, corruption, other problems[13]. Those who were ill but did not seek care were instead asked to provide a reason why they did so. The available options were: not necessary, too expensive, too far away, lack of means of transportation[14], and other reasons. In our analysis we try to understand what is behind the "other" reasons and whether this answer hides an implicit rationing in health care.

For example, it results that most of those who were ill and did not seek care, and who answered that they did so for "other reasons", were living in rural areas (82% as opposed to a national average of 68%). Moreover, these individuals have a mean per capita consumption that is lower than mean per capita consumption for the entire population (17.5 Mt against 21.5 Mt) yet also lower than mean per capita consumption for the entire population of ill individuals (22.5 Mt). This suggests this group constitutes a subgroup of particularly disadvantaged individuals who perceive that given their characteristics they will experience problems in health care access or treatment.

As for the estimation strategy, a three-step procedure is followed:

- Step 1: a probit-with-selection (*heckprobit*) model is used to assess the probability of experiencing problems in access or treatment, provided that one seeks care. The list of explanatory variables includes all the previously introduced covariates. Furthermore, for 2008-09 a variable that registers the number of primary health units - health centres and health posts - per person in the district is also available. This variable was constructed from the annex to the National Inventory of Health Infrastructures, Services and Resources 2007, provided by the Mozambican Ministry of Health and National Institute of Health[15]. In principle, all the variables that affect the probability of seeking care also influence the probability of experiencing problems in access or treatment. However, as evidenced in Wooldridge (2002) among others, Heckman-type models are better identified with exclusion restrictions similar to instrumental variables. Among the independent variables, it was chosen to

[13] We exclude the "too expensive" problem, because of possible sources of endogeneity.
[14] This option was only available in the IAF 2008-09.
[15] Available at the address:
http://www.misau.gov.mz/pt/misau/dpc_direccao_de_nacional_planificacao_e_cooperacao/departamento_de_informacao_para_a_saude_e_monitoria_e_avaliacao/documentos_chave_do_sistema_de_informacao_para_a_saude/sam_inventario_nacional_de_infra_estruturas_de_saude_2007

exclude from the second tier of the model the number of days in which the individual was ill, as this is likely to affect the probability to seek care relatively more than the probability of experiencing problems in access or treatment. However, since this variables is continuous only for 2008-09 - and continuous variables work better as instruments in the estimation procedure -, implicit rationing in health care was only analysed for this year. Indeed, the number of days the individual was ill seems to work well in the model. It is highly significant in the selection equation and it results that there is evidence for a selection issue (Table 5).

- Step 2: after the estimation, the conditional probability of experiencing problems in access or treatment - Pr(problems=1|seek care=1) - is predicted for the whole sample of ill individuals, also for those who did not seek care for various reasons. This provides a measure that we interpret as potential rationing.

- Step 3: the following question is answered: is there potential rationing in the choice of seeking care? That is to say, are those who were ill but did not seek care for "other" reasons[16] the same who had a medium-high probability of experiencing problems in access or treatment? Evidence on this issue is that the percentage of ill individuals who did not seek care for "other" reasons is not negligible (24.4% of the ill who did not seek care in 2002-03, 15.2% in 2008-09), and indeed is generally associated with a medium-high probability of experiencing problems in access or treatment.

5.1 Estimation of the conditional probability and rationing

In Table 5 the results for the probit-with-selection estimation, with marginal effects, are shown, while in Table 7 the imputed rationing probability for ill individuals who did not seek care for "other" reasons is presented. The results in Table 5 confirm the expectation about the sign of most regressors. In particular, individuals from rural areas and those living far away from a health care provider are more likely to experience problems in access or treatment. Moreover, socioeconomic status significantly affects the probability of experiencing problems in access or treatment. A unit increase in (log) consumption decreases the probability to experience problems by 4%. The fact that education is not significant in these regressions also contradicts the possible hypothesis that experiencing problems in access or treatment

[16] We exclude those who declared they did not seek care because it was not necessary, because it was too expensive or because the health care provider was too distant, only keeping those who declared they did not seek care for "other" reasons. Respondents could give more than one answer.

might depend on education-related characteristics, like understanding the prescribed treatment well and implement it correctly[17].

The (log) number of primary health units divided by the population of the district also displays the expected sign in the problems and selection equation, though it is only significant at 10% level.

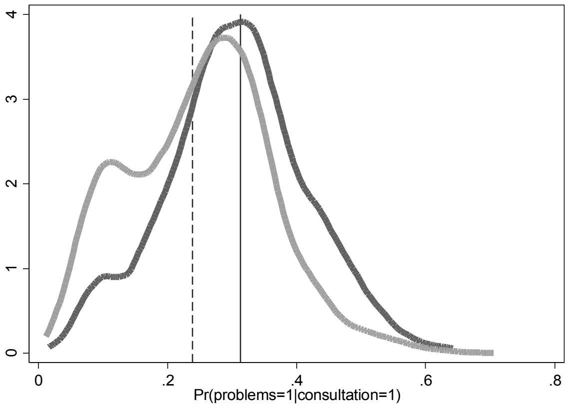
As previously introduced, the number of days the individual was ill is positive and significant in the selection equation, and it results that there is evidence for a selection issue. The parameter ρ is negative and significantly different from zero. In consequence, the lambda term ($\rho \cdot \sigma$) has a negative sign which suggests that the error terms in the two equations are negatively correlated. Hence, (unobserved) factors that make care-seeking (selection) more likely tend to be associated with a lower probability of experiencing problems in health care access or treatment. In particular, those who report having experienced problems are relatively less likely to seek care.

After computing the conditional probability of experiencing problems in access or treatment given care seeking, it was checked whether such probability correctly predicts the behaviour of those ill individuals who actually sought care and did or did not experience problems. Table 6 shows that the computed probability correctly predicts the behaviour of 58% of the sample in 2008-09. In the columns it is displayed the percentage of individuals experiencing problems or not, while in the rows there is the percentage of individuals having a conditional probability lower or higher than the distribution mean.

In Figure 2, instead, the distribution of the conditional probability of experiencing problems for those who were ill and sought care and did not actually experience problems (light grey), or experienced problems (dark grey), is plotted. It is possible to notice that the dark grey distribution lies on the right of the light grey distribution, indicating that the constructed probability discriminates, at least in part, between the two groups. Moreover, the means of the two distribution result to be significantly different.

[17] In Dor et al. (1987) there is the hypothesis that the quality of treatment is dependent on one's education.

Figure 2 – Conditional probability of experiencing problems, 2008-09



Note: conditional probability of experiencing problems for those who were ill and sought care and did not actually experience problems (light grey) or experienced problems (dark grey). The dashed vertical line indicates the mean of the light grey distribution, while the solid vertical line indicates the mean of the dark grey one: they are different at 1% significance level.

From Table 7 it is also possible to see that a cumulative percentage of 73% of ill individuals not seeking care for "other" reasons have an imputed probability of experiencing problems in access or treatment greater than 20%, and that 31% of such individuals have an imputed probability greater than 30%. Moreover, it results that 48% of ill individuals not seeking care for "other" reasons have an imputed probability greater than the distribution mean (26.8%).

These finding are taken to support the hypothesis that some sort of implicit rationing exists. Someone is induced not to seek care at all because he/she fears that after reaching the health care provider his/her economic and travel efforts could be wasted because of problems in access or treatment. Moreover, it seems that the probability of experiencing problems in access or treatment is not equally distributed in the population, but that the poorer and those that are more distant from health providers are more penalised.

6. Conclusions

In this article we examined the issues of health care seeking and implicit rationing in Mozambique. In the first part of the article individual, household and community characteristics that most affect the dichotomous choice of seeking care or not, the choice of seeking care from a formal health care provider provided that one seeks care, and the polytomous choice of where to seek care among different providers were analysed. Nationally representative cross-sectional data from the Mozambican Household Survey on Living Conditions 2002-03 and 2008-09 (*Inquerito aos*

Agregados Familiares sobre Orçamento Familiar, IAF) were used. The role of socioeconomic indicators together with education and distance variables as important determinants for the choice to seek care was highlighted. Moreover gender and education variables also seemed to affect an individual's decision to seek care from formal providers. With regards to the choice of different healthcare providers, the evidence on the explicative power of different variables was mixed: socioeconomic indicators were important for the choice of seeking care from hospitals (only in 2002-03), health centres and other formal providers (only in 2008-09), while education and distance also seemed to affect the decision to seek formal care instead of self-care and traditional medical practitioner's care. That the percentage of ill individuals seeking care from informal providers such as the traditional medical practitioner declined between 2002-03 and 2008-09 was also emphasised.

In the second part of the article, an analysis of whether some sort of implicit health care rationing is in place in Mozambique was carried out. To this aim, an imputed conditional probability of experiencing problems in access or treatment was constructed, on the basis of all the previously introduced explicative variables, but excluding the number of days the individual was ill and using it as an instrument in the probit-with-selection model. Next, a study of whether those who were ill but did not seek care for "other" reasons had an imputed probability of experiencing problems high enough to dissuade them from seeking care was computed. It was found that the percentage of those showing an imputed probability of experiencing problems higher than 20-30% out of the sample of people not seeking care for "other" reasons was not negligible. We interpret this to mean that some form of implicit healthcare rationing exists, against the poorer and those who are more distant from health providers.

Acknoweldgements

I am grateful to Kenneth Simler, Channing Arndt and Valentina Gil for providing help with data collection and analysis, and to Francesca Bettio, Maureen Mackintosh, Rebecca Hanlin and Cecilia Navarra for useful comments on preliminary drafts.

References

Acton, J. P. (1975) 'Nonmonetary Factors in the Demand for Medical Services: Some Empirical Evidence', *Journal of Political Economy*, 87(3): 595-614.

Ager, A. and K. Pepper (2005) 'Patterns of health service utilization and perceptions of needs and services in rural Orissa', *Health Policy and Planning*, 20(3): 176-184.

Ahmed, S. M., A. M. Adams, M. Chowdhury and A. Bhuiya (2000) 'Gender, socioeconomic development and health-seeking behaviour in Bangladesh', *Social Science and Medicine*, 51(3): 361-371.

Akin, J., C. Griffin, D. K. Guilkey and B. M. Popkin (1986) 'The Demand for Primary Health Care Services in the Bicol Region of the Philippines', *Economic Development and Cultural Change*, 34(4): 755-782.

Akin, J., D. Guilkey and E. HazelDenton (1995) 'Quality of services and demand for health care in Nigeria: a multinomial probit estimation', *Social Science & Medicine*, 40(11): 1527-1537.

Baltussen, R. and Y. Ye (2006) 'Quality of Care of Modern Health Services as Perceived by Users and Non-Users in Burkina Faso', *International Journal for Quality in Health Care*, 18(1): 30-34.

Baltussen, R., Y. Ye, S. Haddad and R. Sauerborn (2002) 'Perceived Quality of Care of Primary Health Care Services in Burkina Faso', *Health Policy and Planning*, 17(1): 42-48.

Chibwana, A. I., D. P. Mathanga, J. Chinkhumba and C. H. Campbell Jr (2009) 'Socio-cultural predictors of health-seeking behaviour for febrile under-five children in Mwanza-Neno district, Malawi', *Malaria Journal*, 8(219): 1-8.

Dor, A., P. Gertler and J. Van Der Gaag (1987) 'Non-price rationing and the choice of medical care providers in rural Côte d'Ivoire.', *Journal of Health Economics*, 6(4): 291-304.

Dor, A. and J. Van Der Gaag (1993) 'Quality rationing and the demand of adults for medical care in rural Cote d'Ivoire' in: Mills and Lee (ed.), *Health Economics Research in Developing Countries*. Oxford: Oxford University Press, pp.291-304.

Dow, W. (1996) *Unconditional demand for health care in Côte d'Ivoire: does selection on health status matter?*. Washington, D.C.: World Bank Publications.

Ellis, R., D. McInnes and E. Stephenson (1994) 'Inpatient and outpatient health care demand in Cairo, Egypt', *Health Economics*, 3(3): 183-200.

Gertler, P., L. Locay and W. Sanderson (1987) 'Are user fees regressive? The welfare implications of health care financing proposals in Peru', *Journal of Econometrics*, 36(1-2): 67-88.

Gertler, P. and J. Van Der Gaag (1990) *The willingness to pay for medical care: evidence from two developing countries*. Baltimore, MD and London: Johns Hopkins University Press.

Glick, P., J. Razafindravonona and I. Randretsa (2000) Education and Health Services in Madagascar: Utilization Patterns and Demand Determinants, Cornell Food and Nutrition Policy Program Working Paper No. 107.

Goins, R., K. Williams, M. Carter, S. Spencer and T. Solovieva (2005) 'Perceived barriers to health care access among rural older adults: a qualitative study', *The Journal of Rural Health*, 21(3): 206-213.

Griffiths, P. and R. Stephenson (2001) 'Understanding users'perspectives of barriers to maternal health care use in Maharashtra, India', *Journal of biosocial science*, 33(3): 339-359.

INE (2004) Inquérito Nacional aos Agregados Familiares Sobre Orçamento Familiar 2002/3, http://www.ine.gov.mz/inqueritos_dir/iaf/

INE (2010) Inquérito Sobre Orçamento Familiar, 2008/09 – Quadros Básicos, http://www.ine.gov.mz/inqueritos_dir/iaf/

Iyer, A., G. Sen and A. George (2007) 'The dynamics of gender and class in access to health care: Evidence from rural Karnataka, India', *International Journal of Health Services*, 37(3): 537-554.

Kaija, D. and P. Okwi, 2006. Quality and Demand for Health Care in Rural Uganda: Evidence from 2002/03 Household Survey.

Kapiriri, L. and D. Martin (2007) 'Bedside rationing by health practitioners: a case study in a Ugandan hospital', *Medical Decision Making*, 27(1): 44-52.

Khan, A., J. Walley, J. Newell and N. Imdad (2000) 'Tuberculosis in Pakistan: socio-cultural constraints and opportunities in treatment', *Social Science & Medicine*, 50(2): 247-254.

Lindelow, M. (2004) 'The Utilization of Curative Health Care in Mozambique: Does Income Matter?', *The Centre for the Study of African Economies Working Paper Series*, 11: 1-36.

Lindelow, M. (2005) 'The Utilisation of Curative Healthcare in Mozambique: Does Income Matter?', *Journal of African Economies*, 14(3): 435-482.

Lindelow, M., P. Ward, N. Zorzi and World Bank Africa Regional Office Human Development Dept (2004) *Primary health care in Mozambique: Service delivery in a complex hierarchy*. Human Development Sector, Africa Region, The World Bank.

McFadden, D. (1973) 'Conditional Logit Analysis of Qualitative Choice Behaviour' in: Zarembka (ed.), *Frontiers of Econometrics*. New York: New York Academic Press, pp.105-42.

Ministry of Health and National Institute of Health, 2007. National Inventory of Health Infrastructures, Services and Resources, Mozambique 2007 (Inventario Nacional de Infra-estruturas de Saude, Serviços e Recursos, Moçambique 2007), http://www.misau.gov.mz/pt/misau/dpc_direccao_de_nacional_planificacao_e_cooperacao/departamento_de_informacao_para_a_saude_e_monitoria_e_avaliacao/documentos_chave_do_sistema_de_informacao_para_a_saude/sam_inventario_nacional_de_infra_estruturas_de_saude_2007

MPF (2004) 'Poverty and Well-Being in Mozambique: The Second National Assessment (2002-2003)', National Directorate of Planning and Budget, Ministry of Planning and Finance, Economic Research Bureau, Ministry of Planning and Finance, International Food Policy Research Institute, Purdue University.

Mwabu, G. (1984) *Conditional logit analysis of household choice of medical treatments in rural villages in Kenya*. Nairobi: Kenyatta University College.

Newman, R., S. Gloyd, J. Nyangezi, F. Machobo and J. Muiser (1998) 'Satisfaction with outpatient health care services in Manica Province, Mozambique', *Health policy and planning*, 13(2): 174.

Pinho, M. (2008) 'Racionamento dos cuidados de saúde: problemática inerente', *Cadernos de Saúde Pública*, 24(3): 690-695.

Posse, M. and R. Baltussen (2009) ‘Barriers to access to antiretroviral treatment in Mozambique, as perceived by patients and health workers in urban and rural settings’, *AIDS patient care and STDs*, 23(10): 867-875.

Rosen, S., I. Sanne, A. Collier and J. Simon (2010) ‘Rationing Antiretroviral Therapy for HIV/AIDS in Africa: Efficiency, Equity, and Reality’, *Health and Development Discussion Paper*, 4.

Sahn, D., S. Younger and G. Genicot (2003) ‘The demand for health care services in rural Tanzania’, *Oxford Bulletin of Economics and Statistics*, 65(2): 241-260.

Simler, K., S. Mukherjee, G. Dava and G. Datt (2004) Rebuilding after war: Micro-level determinants of poverty reduction in Mozambique, International Food Policy Research Institute Research Report 132.

Ssewanyana, S., J. Nabyonga, I. Kasirye and D. Lawson (2004) ‘Demand for health care services in Uganda: Implications for poverty reduction’, *EPRC Research Paper*.

Stierle, F., M. Kaddar, A. Tchicaya and B. Schmidt-Ehry (1999) ‘Indigence and access to health care in sub-Saharan Africa’, *The International journal of health planning and management*, 14(2): 81-105.

Train, K. (1986) *Qualitative choice analysis: theory, econometrics, and an application to automobile demand*. Cambridge, MA: The MIT Press.

Table 1 – Descriptive statistics

	2002-03					2008-09				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Price - health post	43638	62965.610	2871.211	0	421649.7	49580	121.340	5.592	0	832.195
Price - health centre	43627	70245.230	2985.573	0	421649.7	49838	124.084	5.763	0	832.195
Price - hospital	43677	73394.110	2749.679	0	421649.7	49785	148.655	7.622	0	832.195
Price - other formal provider	43527	67169.470	2980.873	0	421649.7	49765	149.674	7.762	0	832.195
Price - traditional medical practitioner	43825	8791.176	1363.425	0	267159.3	49978	20.372	2.010	0	832.195
Per capita daily expenditures (log)	44083	9.106	0.026	5.995	14.099	51177	2.871	0.029	-0.502	7.912
Primary education - 1st cycle	44083	0.136	0.008	0	1					
Primary education - 2nd cycle and more	44083	0.109	0.010	0	1					
Education years						50309	4.313	0.085	0	12
Education years squared						50309	29.173	0.901	0	144
Female	44079	0.596	0.009	0	1	51177	0.577	0.007	0	1
Age	43986	32.804	0.411	0	95	51140	23.198	0.450	0	105
Age squared	43986	1414.507	31.602	0	9025	51140	988.881	28.944	0	11025
Ability to read and write	35207	0.339	0.012	0	1	51130	0.550	0.011	0	1
Household size	44083	5.382	0.087	1	31	51177	6.888	0.132	1	34
Household members per room	44066	2.025	0.046	0.143	11	51165	2.252	0.057	0.111	21
Inaccessible community	44083	0.173	0.021	0	1	51177	0.169	0.019	0	1
Rural	44083	0.719	0.014	0	1	51177	0.687	0.017	0	1
Time to closest health facility (0-29)	44014	0.361	0.019	0	1					
Time to closest health facility (30-44)	44014	0.150	0.013	0	1					
Time to closest health facility (45-59)	44014	0.091	0.010	0	1					
Time to closest health facility (60-119)	44014	0.115	0.011	0	1					
Time to closest health facility (120+)	44014	0.314	0.024	0	1					
Time to closest health facility						51177	35.908	1.232	1	100
Time to closest health facility squared						51177	1927.485	112.794	1	10000
0-2 days ill	6181	0.237	0.008	0	1					
3-5 days ill	6181	0.336	0.012	0	1					
6-9 days ill	6181	0.232	0.009	0	1					
More than 10 days ill	6181	0.199	0.010	0	1					
Days ill						6517	5.366	0.114	0	98
Niassa	44083	0.035	0.005	0	1	51177	0.082	0.014	0	1
Cabo Delgado	44083	0.113	0.013	0	1	51177	0.118	0.013	0	1
Nampula	44083	0.241	0.017	0	1	51177	0.202	0.017	0	1
Zambezia	44083	0.199	0.014	0	1	51177	0.133	0.011	0	1
Tete	44083	0.081	0.008	0	1	51177	0.077	0.007	0	1
Sofala	44083	0.045	0.005	0	1	51177	0.028	0.003	0	1
Manica	44083	0.091	0.010	0	1	51177	0.085	0.012	0	1
Inhambane	44083	0.056	0.006	0	1	51177	0.142	0.020	0	1
Gaza	44083	0.051	0.004	0	1	51177	0.042	0.004	0	1
Maputo Province	44083	0.044	0.004	0	1	51177	0.039	0.005	0	1
Maputo City	44083	0.057	0.004	0	1	51177	0.052	0.006	0	1
Trimester of interview - 1	44083	0.353	0.027	0	1	51177	0.295	0.026	0	1
Trimester of interview - 2	44083	0.246	0.025	0	1	51177	0.256	0.025	0	1
Trimester of interview - 3	44083	0.224	0.023	0	1	51177	0.212	0.021	0	1
Trimester of interview - 4	44083	0.180	0.015	0	1	51177	0.237	0.021	0	1

Table 2 – Probit: Seek care or not – Marginal effects

	2002-03				2008-09			
Variable	Model 1		Model 2		Model 1		Model 2	
Per capita daily expenditures (log)	0.0892***	(0.0195)			0.0717***	(0.0200)		
Per capita daily expenditures - Q2			0.114***	(0.0379)			0.0182	(0.0270)
Per capita daily expenditures - Q3			0.160***	(0.0351)			0.0343	(0.0292)
Per capita daily expenditures - Q4			0.149***	(0.0369)			0.046	(0.0302)
Per capita daily expenditures - Q5			0.199***	(0.0356)			0.149***	(0.0349)
Primary education - 1st cycle	-0.00535	(0.0330)	-0.00218	(0.0330)				
Primary education - 2nd cycle and more	0.0241	(0.0396)	0.029	(0.0384)				
Education years					0.0254**	(0.0112)	0.0258**	(0.0112)
Education years squared					-0.00143	(0.000956)	-0.00148	(0.000961)
Female	-0.00586	(0.0196)	-0.00528	(0.0196)	0.0163	(0.0183)	0.0158	(0.0181)
Age	0.00487**	(0.00222)	0.00469**	(0.00222)	-0.0014	(0.00116)	-0.00148	(0.00113)
Age squared	-8.35e-05***	(2.76e-05)	-8.04e-05***	(2.76e-05)	5.79e-07	(1.64e-05)	1.27e-06	(1.60e-05)
Ability to read and write	0.0919***	(0.0253)	0.0932***	(0.0250)	0.038	(0.0252)	0.0381	(0.0250)
Household size	0.00151	(0.00501)	0.00133	(0.00503)	0.00227	(0.00358)	0.00253	(0.00356)
Household members per room	0.0289**	(0.0123)	0.0305**	(0.0122)	-0.00973	(0.00854)	-0.00986	(0.00847)
Inaccessible community	0.015	(0.0414)	0.013	(0.0417)	-0.139***	(0.0369)	-0.140***	(0.0368)
Rural	-0.0403	(0.0290)	-0.0427	(0.0290)	-0.118***	(0.0268)	-0.115***	(0.0265)
Time to closest health facility (30-44)	-0.0477	(0.0313)	-0.05	(0.0311)				
Time to closest health facility (45-59)	-0.207***	(0.0452)	-0.210***	(0.0449)				
Time to closest health facility (60-119)	-0.181***	(0.0420)	-0.187***	(0.0415)				
Time to closest health facility (120+)	-0.278***	(0.0385)	-0.282***	(0.0382)				
Time to closest health facility					-0.00407**	(0.00191)	-0.00404**	(0.00189)
Time to closest health facility squared					3.79e-05**	(1.87e-05)	3.77e-05**	(1.86e-05)
3-5 days ill	0.126***	(0.0278)	0.126***	(0.0276)				
6-9 days ill	0.250***	(0.0329)	0.250***	(0.0325)				
More than 10 days ill	0.291***	(0.0361)	0.291***	(0.0356)				
Days ill					0.0160***	(0.00274)	0.0156***	(0.00269)
Observations	42254		42254		50713		50713	
Subpopulation	4469		4469		6393		6393	

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Province and trimester dummies omitted

Table 3 – Multinomial logit: Where to seek care – Marginal effects

	2002-03											
Variable	Self-care		Health post		Health centre		Hospital		Other formal		TMP	
Price - health post	3.32e-07	(4.61e-07)	-2.20e-06***	(4.29e-07)	1.52e-06***	(4.49e-07)	1.44e-07	(1.13e-07)	1.40e-07	(1.65e-07)	6.26e-08	(1.43e-07)
Price - health centre	-9.79e-08	(8.06e-07)	2.89e-06***	(7.11e-07)	-4.23e-06***	(8.21e-07)	4.24e-07**	(1.82e-07)	2.78e-07	(2.63e-07)	7.36e-07***	(2.50e-07)
Price - hospital	-3.55e-07	(6.39e-07)	4.38e-08	(6.73e-07)	1.76e-06**	(7.92e-07)	-6.18e-07***	(2.04e-07)	-4.39e-07**	(1.92e-07)	-3.92e-07	(2.59e-07)
Price - other formal provider	9.53e-07	(8.51e-07)	-1.01e-06*	(5.21e-07)	1.53e-07	(8.46e-07)	7.07e-08	(1.98e-07)	2.09e-07	(2.70e-07)	-3.72e-07**	(1.88e-07)
Price - traditional medical practitioner	5.84e-07	(6.68e-07)	-1.10e-06	(1.06e-06)	5.34e-07	(7.84e-07)	-1.53e-07	(2.38e-07)	-5.44e-09	(1.80e-07)	1.43e-07	(1.30e-07)
Per capita daily expenditures (log)	-0.117***	(0.0248)	0.0269	(0.0227)	0.0536***	(0.0174)	0.0150***	(0.00550)	0.00914	(0.00586)	0.0119	(0.00907)
Primary education - 1st cycle	0.0367	(0.0386)	-0.014	(0.0356)	0.00279	(0.0280)	0.00127	(0.00785)	-0.0112	(0.00723)	-0.0155	(0.0127)
Primary education - 2nd cycle and more	0.0186	(0.0518)	-0.0283	(0.0394)	0.0449	(0.0412)	0.0284*	(0.0146)	-0.0142*	(0.00776)	-0.0493***	(0.0130)
Female	0.00852	(0.0213)	0.00893	(0.0210)	0.00549	(0.0181)	0.00772	(0.00596)	-0.00632	(0.00586)	-0.0243***	(0.00901)
Age	-0.00454*	(0.00258)	0.00139	(0.00237)	0.00141	(0.00196)	0.00116*	(0.000641)	-0.000269	(0.000714)	0.000854	(0.000823)
Age squared	8.65e-05***	(3.24e-05)	-4.16e-05	(3.16e-05)	-2.90e-05	(2.39e-05)	-1.23e-05	(7.94e-06)	6.04e-06	(8.47e-06)	-9.65e-06	(1.04e-05)
Ability to read and write	-0.119***	(0.0297)	0.0224	(0.0315)	0.0718***	(0.0238)	0.0219**	(0.00874)	0.0173**	(0.00872)	-0.0149	(0.0114)
Household size	-0.00957*	(0.00500)	0.0032	(0.00401)	0.00832**	(0.00349)	-0.000449	(0.00126)	-0.000724	(0.00128)	-0.00078	(0.00195)
Household members per room	-0.0179	(0.0140)	0.000829	(0.0121)	-0.00607	(0.0115)	0.00970**	(0.00407)	0.00246	(0.00386)	0.0109**	(0.00448)
Inaccessible community	0.0024	(0.0463)	-0.00819	(0.0511)	-0.0223	(0.0437)	-0.0345***	(0.0130)	-0.00231	(0.0187)	0.0649**	(0.0276)
Rural	0.0294	(0.0391)	0.101**	(0.0449)	-0.0227	(0.0404)	-0.0781**	(0.0365)	-0.0228	(0.0166)	-0.00706	(0.0176)
Time to closest health facility (30-44)	0.0476	(0.0404)	-0.133***	(0.0434)	0.0506*	(0.0296)	0.00989	(0.00753)	0.00558	(0.0117)	0.019	(0.0172)
Time to closest health facility (45-59)	0.202***	(0.0515)	-0.151***	(0.0551)	-0.0612	(0.0476)	0.006	(0.0119)	0.0162	(0.0175)	-0.0115	(0.0212)
Time to closest health facility (60-119)	0.155***	(0.0497)	-0.0812	(0.0522)	-0.0867*	(0.0451)	-0.0177	(0.0182)	-0.0116	(0.0181)	0.0425**	(0.0186)
Time to closest health facility (120+)	0.262***	(0.0491)	-0.225***	(0.0599)	-0.0888*	(0.0458)	-0.00908	(0.0121)	0.00472	(0.0213)	0.0553***	(0.0155)
3-5 days ill	-0.143***	(0.0316)	0.0733**	(0.0308)	0.0334	(0.0260)	0.0233**	(0.00932)	-0.00264	(0.00741)	0.016	(0.0140)
6-9 days ill	-0.286***	(0.0368)	0.138***	(0.0327)	0.0949***	(0.0281)	0.0387***	(0.0114)	0.00242	(0.00921)	0.012	(0.0137)
More than 10 days ill	-0.335***	(0.0410)	0.0949**	(0.0371)	0.177***	(0.0334)	0.0519***	(0.0124)	-0.0106	(0.00963)	0.0223	(0.0152)
Observations	42156		42156		42156		42156		42156		42156	
Subpopulation	4371		4371		4371		4371		4371		4371	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Province and trimester dummies omitted

Table 4 – Multinomial logit: Where to seek care – Marginal effects

	2008-09											
Variable	Self-care		Health post		Health centre		Hospital		Other formal		TMP	
Price - health post	0.000308*	(0.000175)	-0.000492**	(0.000194)	0.000128	(0.000267)	6.66e-05	(0.000121)	-1.75e-05	(1.80e-05)	7.07e-06	(6.15e-06)
Price - health centre	0.000891***	(0.000195)	0.000959***	(0.000283)	-0.00186***	(0.000368)	-8.93e-05	(0.000206)	7.42e-05***	(2.25e-05)	2.02e-05***	(7.77e-06)
Price - hospital	-0.000537**	(0.000218)	0.000414	(0.000379)	4.32e-05	(0.000377)	0.000116	(0.000131)	-1.30e-05	(1.77e-05)	-2.25e-05***	(7.77e-06)
Price - other formal provider	8.29e-06	(0.000188)	-0.000640*	(0.000342)	0.000832**	(0.000368)	-0.000185	(0.000118)	-1.93e-05	(1.60e-05)	4.53e-06	(6.91e-06)
Price - traditional medical practitioner	-0.000185	(0.000187)	-0.000178	(0.000210)	0.000324	(0.000272)	7.58e-05	(0.000118)	-3.72e-05*	(1.95e-05)	-2.63e-08	(4.10e-06)
Per capita daily expenditures (log)	-0.0781***	(0.0182)	0.0246	(0.0214)	0.0413*	(0.0230)	0.0039	(0.0103)	0.00834***	(0.00231)	-4.46e-05	(0.000662)
Education years	-0.0282**	(0.0124)	0.0224**	(0.00949)	0.00576	(0.0128)	0.00109	(0.00515)	-0.000649	(0.000873)	-0.000427	(0.000465)
Education years squared	0.00183*	(0.00106)	-0.00212**	(0.000851)	0.000102	(0.00102)	0.00011	(0.000406)	4.22e-05	(6.28e-05)	3.49e-05	(4.20e-05)
Female	-0.0275	(0.0219)	0.00933	(0.0152)	0.0163	(0.0217)	0.00461	(0.00697)	-0.00228	(0.00199)	-0.000457	(0.000769)
Age	0.00191	(0.00139)	-0.00193*	(0.00108)	-0.00126	(0.00151)	0.0011	(0.000695)	0.000207	(0.000146)	-2.84e-05	(4.61e-05)
Age squared	-1.48e-06	(1.98e-05)	6.72e-06	(1.52e-05)	5.17e-06	(2.12e-05)	-7.75e-06	(1.07e-05)	-3.23e-06	(2.22e-06)	5.74e-07	(6.42e-07)
Ability to read and write	-0.0581**	(0.0266)	0.00797	(0.0212)	0.0375	(0.0290)	0.00938	(0.0131)	0.00488**	(0.00242)	-0.00165	(0.00133)
Household size	-0.00328	(0.00393)	0.00147	(0.00308)	0.00104	(0.00461)	0.0013	(0.00170)	-0.000543*	(0.000314)	7.95e-06	(0.000167)
Household members per room	0.011	(0.00944)	-0.00892	(0.00688)	0.00975	(0.0108)	-0.0105**	(0.00514)	-0.00125	(0.000970)	3.60e-06	(0.000315)
Inaccessible community	0.105***	(0.0395)	0.0147	(0.0367)	-0.117**	(0.0492)	-0.00639	(0.0319)	0.004	(0.00729)	0.000366	(0.00149)
Rural	0.101***	(0.0358)	0.0199	(0.0445)	0.0284	(0.0538)	-0.142***	(0.0478)	-0.00864*	(0.00470)	0.00068	(0.00139)
Time to closest health facility	0.00303	(0.00214)	-0.00245	(0.00183)	-0.00141	(0.00228)	0.000788	(0.00114)	4.54e-05	(0.000141)	-2.13e-07	(5.28e-05)
Time to closest health facility squared	-2.60e-05	(2.13e-05)	3.73e-05**	(1.74e-05)	-1.86e-06	(2.35e-05)	-9.51e-06	(1.17e-05)	-2.19e-07	(1.50e-06)	2.87e-07	(5.86e-07)
Days ill	-0.0181***	(0.00271)	0.00356**	(0.00165)	0.0114***	(0.00203)	0.00308***	(0.000656)	-6.91e-05	(0.000207)	0.000124***	(4.22e-05)
Observations	50466		50466		50466		50466		50466		50466	
Subpopulation	6146		6146		6146		6146		6146		6146	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Province and trimester dummies omitted

Table 5 – Probit with selection: Experiencing problems conditional on seeking care, with marginal effects

	2008-09							
Variable	Problems		Seek care		Pr(problems=1|seek care=1) *pcond*		Pr(seek care=1) *psel*	
Per capita daily expenditures (log)	-0.190***	(0.0556)	0.206***	(0.0582)	-0.04**	(0.019)	0.073***	(0.02)
Education years	-0.0309	(0.0346)	0.0712**	(0.0326)	-0.001	(0.012)	0.025**	(0.012)
Education years squared	0.00278	(0.00267)	-0.00387	(0.00277)	0	(0.001)	-0.001	(0.001)
Female	-0.028	(0.0543)	0.0435	(0.0527)	-0.004	(0.019)	0.015	(0.019)
Age	0.00734*	(0.00386)	-0.00464	(0.00339)	0.002	(0.001)	-0.002	(0.001)
Age squared	-6.47e-05	(5.17e-05)	2.31e-06	(4.83e-05)	0	(0)	8.18E-007	(0)
Ability to read and write	-0.0459	(0.0876)	0.115	(0.0736)	0	(0.03)	0.041	(0.026)
Household size	-0.0278**	(0.0111)	0.00475	(0.0106)	-0.01***	(0.004)	0.002	(0.004)
Household members per room	-0.0136	(0.0324)	-0.0289	(0.0236)	-0.01	(0.011)	-0.01	(0.008)
Inaccessible community	0.0578	(0.135)	-0.394***	(0.0974)	-0.038	(0.042)	-0.14***	(0.034)
Rural	0.280***	(0.104)	-0.421***	(0.0861)	0.042	(0.036)	-0.149***	(0.031)
Time to closest health facility	0.0142**	(0.00627)	-0.0109*	(0.00568)	0.004*	(0.002)	-0.004*	(0.002)
Time to closest health facility squared	-0.000160**	(6.25e-05)	9.99e-05*	(5.59e-05)	-0**	(0)	0*	(0)
Days ill			0.0496***	(0.00837)			0.0177***	(0.003)
Number of primary health units divided by the population of the district (log)	-0.163*	(0.0835)	0.151*	(0.0876)	-0.039	(0.03)	0.054*	(0.031)
Constant	-1.112	(0.757)	1.542*	(0.787)				
athrho			-0.826***	(0.296)				
Observations	50,711		50,711		50,711		50,711	
Subpopulation	6391		6391		6391		6391	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Province and trimester dummies omitted

Table 6 – Predicted and actual behaviour of ill individuals

	2008-09		
	Experienced problems		
Conditional probability	No	Yes	Total
Lower than the mean	0.41	0.08	0.49
Higher than the mean	0.34	0.17	0.51
Total	0.74	0.26	1

Table 7 – Imputed probability of experiencing problems for ill individuals who did not seek care for "other" reasons

2008-09

Imputed probability of experiencing problems in access or treatment	Percentage of ill individuals not seeking care for "other" reasons
≥ 20%	0.72
≥ 30%	0.31
≥ 40%	0.11