May 2012

Poverty in Mozambique: new evidence from recent household surveys[1]

Abstract

The scope of this paper is threefold: to investigate potential problems affecting the last Mozambican nationally-representative household survey on consumption expenditure, to verify robustness and reliability of official poverty figures, and finally to provide alternative estimates of poverty and welfare indicators. Two main issues affect the official poverty rate estimates: measurement errors in consumption data and in methodology used to calculate poverty lines (PLs) i.e. cost of basic needs approach based on provincial food bundles with entropy correction. In the present work, observations likely affected by high measurement errors have been identified, which in turn negatively impacted PLs. Our strategy is to provide methods to correct the consumption distribution by re-calculating the PLs based on a single national food basket, unlike the official PLs estimates based on provincial-specific food baskets. Results show that the revised poverty incidences are weakly correlated to the official estimates on poverty across provinces, and the former is more consistent with other socio-economic indicators. With the use of three cross-sectional data of consumption surveys (*IAF* 1996/97, 2002/03 and *IOF* 2008/09, from their Portuguese acronyms) and recalculating welfare measures, results show that poverty was more tapered off in the period of 1996/97-2002/03 than between 2002/03-2008/09. Also, poverty seems to be highly concentrated in some areas with dramatically high rates found in Central and Northern Mozambique, as well as in rural areas, and relatively low rates in Southern Mozambique.

1. Introduction

[1] Authors: Federica Alfani, Carlo Azzarri, Marco D'Errico and Vasco Molini. Corresponding author Vasco Molini vmolini@worldbank.org

* Comments on earlier versions by Peter Lanjouw, Kathleen Beagle, Bart van den Boom, Gero Carletto, Andrew Dabalen and Julio Revilla are gratefully acknowledged. A special thank to Antonio Nucifora for having initiated and strongly supported the whole process of poverty figures revision. Usual disclaimer applies.

Three consecutive "National Poverty Assessments" (NPA) in Mozambique (MPD-DNEAP, 2010; MPF/IFPRI/PU, 2004; MPF/UEM/IFPRI, 1998) have provided a wealth of information over the last decade, on the poverty patterns and the changes thereof. The recently published NPA based on the new 2008/09 Mozambique Household Budget Survey (known as IOF, *Inquérito sobre Orçamento Familiar*) shows that, although welfare has generally increased, poverty is still widespread in the country. The national poverty ratio, based on total household per-capita consumption, was estimated at 54.7% according to the IOF 2008/09, as opposed to 54.1% 2002/03, with an increasing rural poverty (from 55.3% to 56.9%), and a decreasing urban poverty (from 51.5% to 49.6%).

However, the poverty profile that emerged from the latest NPA (MPD-DNEAP, 2010) computed using a consumption-based money metric of welfare, does not seem to be entirely consistent with a number of alternative indicators ranging from, child anthropometrics (see Azzarri *et al.*, 2011), asset ownership, access to basic services and GDP estimates. These, and other elements, trigger concerns about the robustness of poverty statistics in the latest NPA. This discrepancy becomes even more striking at the provincial level[2] . In the last NPA, rural Niassa and Cabo Delgado, historically considered poor according to various indicators, show the lowest poverty rates in the country. In contrast, Southern provinces and urban areas record very high poverty rates while traditionally facing low incidence of child malnutrition and higher GDP.

Three sources of error might have contributed to these counter-intuitive results. First, measurement errors on expenditures may be due to consumption underreporting or flaws in the questionnaires. Second, the methodology used to compute poverty lines[3] (PLs), which is not the standard one used in many other developing countries, might have caused unexpected provincial patterns. Third, the interaction between consumption underreporting and the particular methodology used to compute the PLs might have caused the PLs to be highly dependent on flawed data.

The present paper aims to analyze these three potential sources of bias in the poverty figures, and to isolate their effect on the final estimates. New tentative estimates of poverty figures are also proposed, which are more consistent with poverty correlates and historic geographical welfare distribution in the country. The paper is organized as follows. Section 2 presents an overview on measurement error in the 2002/03 and 2008/09 survey and presents our strategy to identify observations affected by measurement errors and re-estimate them, section 3 deepens the relationship and interaction between underreporting and PLs and proposes an alternative methodology to compute the PLs; section 4 presents the results from the revised PLs and re-estimation of observations likely

[2] North, Center and South regions are grouped according to the official country definition. *North* includes the provinces of Niassa, Cabo Delgado and Nampula; *Center* includes Tete, Zambezia, Manica and Sofala; and *South* includes Gaza, Inhambane, Maputo provincia, Maputo city. For an overview of Mozambican provinces, see Map 1 below

[3] All the three NPA report a different poverty line, based on local prices and consumption patterns observed in the cross-sectional rounds of, respectively, the IAF (*Inquérito aos Agregados Familiares sobre Orçamento Familiar*) 1996/97, the IAF 2002/03 and the IOF 2008/09 survey, by grouping provinces and areas according to 13 locations: Niassa and Cabo Delgado, rural and urban; Nampula, rural and urban; Sofala and Zambezia, rural and urban; Manica and Tete, rural and urban; Gaza and Inhambane, rural and urban, Maputo Province, rural and urban; Maputo City.

affected by underreporting, and discusses the resulting difference between official and revised poverty levels and trends; and finally section 5 concludes.

2. Measurment error in consumption

2.1 Underreported consumption

This section covers the most relevant general findings on consumption underestimation and provides a background and rationale for the present work. The analysis proceeds in two directions: first, it seeks to identify a subset of likely downward biased observations in food consumption in the IOF 2008/09 data and, subsequently, to look at the possible sources of this bias. We explore the consumption data from three perspectives: over all calorie intake, dietary diversity patterns, and food share patterns.

About 30% of households report to consume less than an average of 1,000 daily k-calories per capita per day, considered below the minimum standard for survival by the World Health Organization; moreover, in some Southern regions median calorie intake is even below this threshold. We follow the strategy to run an OLS regression model where per-capita consumption expenditure is regressed against poverty correlates. A clear pattern of the error terms (not shown) can be identified: observations with calorie intake below 1000 tend to show high concentration of negative residuals. This result suggests that, for households showing calories per capita/day below 1000, the model predicts higher fitted expenditures, based on characteristics highly correlated to consumption, than the original expenditures.

To further check the incongruities in the data, dietary diversity is analyzed. Unexpectedly, the share of cereals, a food category typically consumed in big quantities by the poor, increases with expenditure, whereas consumption of nutritionally superior items such as meat, fish, and dairy does not increase with expenditure.

Finally, we find that, contrary to normal patterns in consumption data, the share of food consumption is very low in the South, the region with the officially estimated highest poverty rate. Even more surprising, it seems that the Engel's law does not hold in rural areas. The negative correlation between share of food consumption and expenditure level is hardly detectable.

When looking at the causes of consumption underreporting, three elements emerge. The first is the pre-printed food item list in the consumption questionnaires. In the 2008/09 survey, a list of 18 (for purchases) and 20 (for own consumption) items had been used. Both item lists seem to be too restrictive to adequately capture the dietary diversity and composition across provinces. The enumerators' apparent tendency to underreport transactions could have been facilitated by an overly restricted questionnaire, which may not have left enough room, nor room for reporting additional food items at the bottom of the module. Also, the fact that the pre-typed list included only primary agricultural commodities, but not processed food, could have prevented an exhaustive collection of food consumed in urban areas, where households devote a larger share of consumption to processed food. The 2002/03 survey was also affected by similar problems and this caused a potential bias in poverty figures, particularly in Southern Mozambique.

The second issue is poor management of data collection activities. Food consumption data in the IAF 2002/03 and IOF 2008/09 have been collected on the field differently from what was initially envisaged. In fact, food transactions should have been collected by the households through a diary where every two days they would keep track of the daily transactions occurred, as three visits were initially envisaged. However, this procedure was not implemented since, instead, only one visit was completed in 93 percent of the cases in the IOF 2008/09. It should not come as a surprise that many households showing low calorie consumption also reported fewer than ten transactions during the recall period (15 days); this can reveal limited supervision during the survey fieldwork activities.

Finally, the lack of official conversion factors for non-standard measurement units could have highly affected calorie calculations. Two main assumptions are posed in computing calorie consumption from the 2008/09 survey, namely 1) one piece/unit of bread is set to weigh 200 grams in every province, although it is uncertain whether this arguably holds across the entire country; 2) one coconut is assumed to contain 800 grams of edible parts.

The problems related to caloric intake, and to the overall limited concordance between well-being indicators and poverty figures, also seem to characterize the 2002/03 data of well-being indicators. These drawbacks in data collection and analysis produce an extremely volatile and time inconsistent poverty profile with a continuous re-ranking of provinces in terms of poverty (Figure 1). On the contrary, a different survey design of the IAF 1996/97 guaranteed higher reliability and, as a consequence, greater coherence of poverty figures with other welfare indicators[4]. The more complete food list in 1996/1996/97, accounting for 50 items, most likely led to better consumption measurements, through a reduced underreporting and a more limited concentration of reported transactions on the pre-printed items with the consequent decrease of diet polarization.

NPA's authors (DNEAP 2010) carefully examine the problem, coming to similar conclusions regarding the magnitude of underreporting and its spatial concentration. It is also proposed as preliminary solution, the re-estimation of expenditures for households whose calorie intake falls below a certain threshold. Adjustments produce substantial reduction in poverty figures at regional level in the South and a reduction of 3% in the poverty rate at national level. However, since also 2002/03 data show similar problems in the South, the NPA's authors argue that even though there is some regional variation, the trend between the two surveys results very stable.

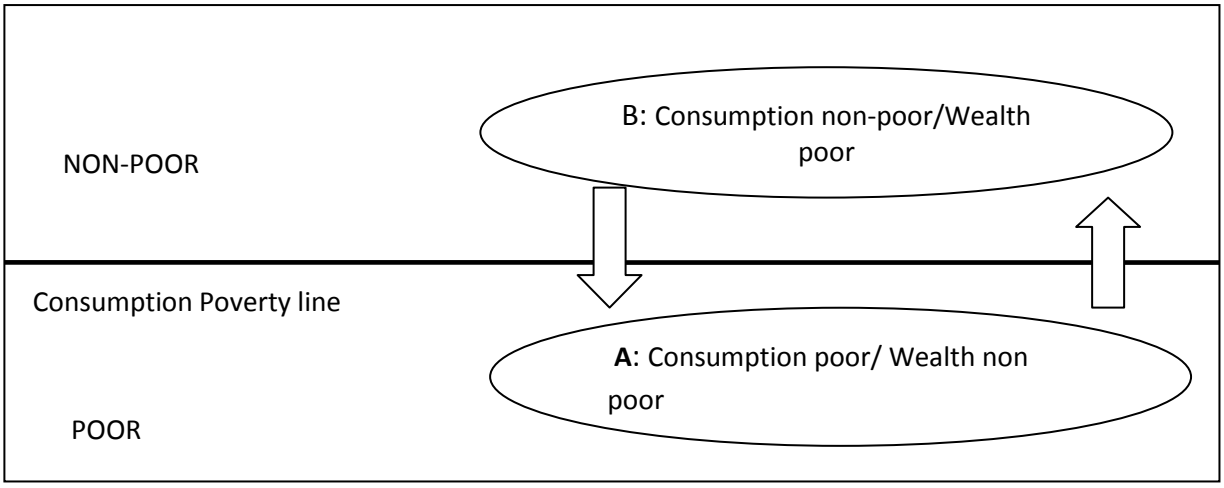
[4] In 1996/97 provincial poverty rates are negatively and significantly correlated with all welfare and education indicators.

2.2 Identifying likely unreliable observations: 2002/03 and 2008/09

In order to cope with measurement error in consumption and, in a second stage, with the 'error-poverty line interaction' (section 3), we opted for a method that saves most of the original information while correcting only manifest mis-measured observations. Our approach is to carefully screen all consumption data and set defined and specific criteria to determine their degree of reliability[5]. In a second stage the unreliable observations are re-estimated using out of sample predictions from 1996/97coefficients model (see section 2.3). Since we identified also for 2002/03 clear signs of measurements error, after having tested the methodology for 2008/09, we applied to this survey as well.

Original consumption data are ranked and compared to a wealth index[6]. Observations below the poverty line (54th percentile of income distribution) and at the same time above the 70th wealth index percentile or observations above the poverty line and below the 30th wealth index percentile are singled out as 'suspicious'[7]. Scheme 1 illustrates the definition of these two different sets: a set of households (A) considered poor based on original data but classified as non-poor based on the wealth index and a set of households (B) considered as non-poor based on original consumption data, but classified as poor based on the wealth index.

Scheme 1



[5] The first best approach to identify consumption measurement error would be to administer a small consumption-based survey questionnaire in the areas where underreporting and over-reporting are deemed more serious, revisiting the same households sampled in the IOF 2008/09. The comparison with the original data would provide the extent to which questionnaire design and fieldwork operations can impact consumption data quality, and in turn diet composition and caloric intake. Information could then be used to adjust consumption data in the IOF 2008/9.

[6] With this asset index, suggested by Filmer and Pritchett (2001), principal components analysis is used to calculate the weights of the index. The first principal component, the linear combination capturing the greatest variation among the set of variables, can be converted into factor scores, which serve as weights. This rationale for using this index is that it captures, better than consumption, the household permanent welfare dimension, and can provide a more reliable rank among households.

[7] The use of 30th and 70th percentile is not coincidental: by choosing extreme and symmetrical values (+/- 20 percentiles form the median) of the wealth distribution we aimed at screening consumption observations that were reasonably inconsistent with their corresponding wealth index value.

Further, we controlled whether general characteristics[8] of the two sets were significantly different from those of other observations: set A compared with all other consumption poor and set B compared to all the other consumption non poor (*see appendix*), set A compared to set B and finally set A compared to the consumption non-poor. This cross comparison substantially validates our hypothesis. General characteristic of set A are significantly superior (higher education, more access to services, assets etc) than those of other consumption poor, whereas they closely resemble those of consumption non-poor if we exclude set B.

Set A observations also show a clear regional pattern (see table 1)**.** In both survey rounds, about 70**%** of set A households are located in the South (Gaza, Inhambane and Maputo area), and this concentration is not consistent with population distribution as this Southern population accounts for only 25% of that of Mozambique (see Map 1). This suggests that, in the South, the interaction between location-specific PLs and predicted consumption might change the calculation of poverty figures.

Regarding B, i.e the consumption non poor/wealth set, around 80% of observations are concentrated in Central and Northern rural areas in both years (see table 1)**.** Their general characteristics are associated with lower and statistically significant outcomes than those of the other non-poor households and high resemblance to poor households not belonging to the set A (*see appendix*). This could suggest that consumption data might also be affected by overestimation, mostly due to prices, given that the share of own-consumption is also higher than in the rest of the country.

For these households own-consumption shadow prices of key items (maize, rice, other cereals, cassava, groundnuts, chestnuts, bananas and tomatoes) are significantly higher than those for the rest of the households, also due to their higher share of own-consumption in total food consumption. Also, these prices are greater than purchase prices, unlike for the rest of households, something that runs counter to the general behavior of farmers to market higher quality commodities while keeping the lower quality food for them. A possible source of discrepancy is that while households reported purchase prices, information on own-consumption prices was collected at the community level and associated to all households in the same community.

In sum, both analyzed survey show problems of measurement error; in 2002/03 and 2008/09 respectively 27% and 21% of the observations show a poverty status inconsistent with all their other characteristics. Among them, the set (A) is likely affected by underreporting; and the set B by overestimation, mostly due to prices. The remaining households seem to show more concordance among various welfare indicators. In section 4 we proceed to re-estimate these observations with out of sample estimates based on the 1996/97 survey.

2.3 Using out of sample estimates to re-estimate unreliable observations

The fact that the 1996/97 survey seems to be less affected by consumption underestimation, and more consistent with other welfare indicators, suggests that consumption prediction models can

[8] Non-food share, head's years of education, highest level of household members education, ownership of assets, access to protected water, electricity and toilet, housing conditions and dietary variety

produce a less-biased fitted measure. Thus, for each survey (1996/97, 2002/03, 2008/09), an identical consumption model having as dependent variable household per capita consumption (y_i), deflated with a national consumer price index *CPI*, has been estimated (1). Then, out-of sample predictions for both 2002/03 and 2008 are obtained by substituting original coefficients with those obtained applying the 1996/97 model parameters. Normally distributed errors are randomly assigned (2)

$$\frac{y_i^{97}}{CPI_{97-j}} = \alpha + \beta * X_i^{97} + \varepsilon_i \qquad for \;\; j = 2002/03, 2008/09 \tag{1}$$

$$\widehat{y_i^j} = a + \hat{\beta} X_i^j + f(\hat{\epsilon}_i) \qquad for \;\; j = 2002/03, 2008/09;\; f(\hat{\epsilon}_i) \sim N(0,1) \tag{2}$$

Results illustrate the hypothetical scenario had the underlying relationship between consumption and its correlates remained constant over the period 1996/97 to 2008/09. Coefficients' time invariance is arguably a strong assumption[9]. However, bearing in mind the evidence that the model estimated on the IAF 1996/97 is less affected by measurement error, the out-of-sample prediction in 2002/03 and 2008/09 produces for mis-measured observations a revised consumption that is less biased than the actual. Furthermore, using the 1996/97 parameters, variation in consumption figures over time are fully explained by variation of the independent variables (see appendix).

Figure 2 shows median consumption per capita in 2002/03 and 2008/09 using official data and the revised consumption data. Whereas in 2002/03 the difference between original and revised figures seem to be remarkable only in the South, the 2008/09 model prediction shows higher median consumption in Southern and urban provinces and much lower prevalence in rural as well as Northern and Central provinces. For example, in the Urban South (basically the Maputo urban area) median consumption is slightly higher than in 2002/03 and the gap dramatically widens in 2008/09. By contrast, in rural Nampula and Niassa-Cabo Delgado the predicted median consumption is lower than the official.

The full sample re-estimation of 2002/03 and 2008/09 data, confirms what discussed previously. For 2008/09 and to lesser extent 2002/03, data are affected by underestimation problems but also by substantial 'overestimation' in some Central and Northern areas of the country. In these two areas the model indicates that general household characteristics would predict a lower consumption per capita and thus, keeping the official PLs, higher poverty rates.

Although providing useful insight into the temporal dynamic of expenditures and poverty, the full re-estimation of the two surveys is clearly an extreme measure; not all consumption data are affected by measurement error. Using the out-sample prediction and combining it with the observations

[9] An alternative could be to re-estimate the mis-measured observations using a so-called 'hybrid model' (see appendix). The model is estimated on a sample of reliable observations and the estimated β coefficients are then applied to the whole sample; the fitted consumption is assigned only to the biased observations and errors are estimated by random assignment of normally-distributed errors from equation. Both methodologies come to a cost; estimates based on a reduced sample are not regionally representative (regional pattern of dropped observations) while estimates based on 1996/97 coefficients assume their invariance over time. Being the results for the two methodologies quite similar, we opted for the 1996/97 based model since we judged that the lack of regional representation was a much bigger problem.

selection discussed in section 2.2, the fitted consumption is assigned only to the observations of set A and B (equation 3), for both 2002 and 2008 (3)

$$\overline{y_i^j} = \widehat{y_i^j} \qquad for\ i \ \in \ A \cup B;\ for \ \ j = 2002/03, 2008/09; \qquad (3)$$

The partial prediction using 1996/97 estimates embeds a number of desirable properties. First, overall provincial poverty rates using the show a similar pattern to those obtained using other correction techniques (*see footnote 9 and appendix*). Second, even using bootstrapped error terms in equation 2, results are statistically unchanged, witnessing the robustness of results to the sample selected. Third, the model specification is relatively parsimonious, with the consequent low noise in the fitted variable.

3. Poverty lines and measurement error interaction

3.1 The official PLs

Most national poverty estimations define a single poverty line using a single food basket, pricing that basket and adding to it a cost of non-food items. The PL is a minimum threshold (Cost of Basic Needs-CBN) below which one individual is considered poor and is calculated *for all individuals in the country*.

The three NPAs take advantage of information on local prices, from reported household unit values as well as community markets, and local dynamic consumption patterns -the so-called "adjusted flexible bundles" (AFB)- to create 13 location-specific PLs varying over time (Tarp *et al.*, 2002; Arndt and Simler, 2010). Location-specific flexible PLs are designed to reflect more accurately local tastes and preferences, under the assumption that in such a wide and diverse country such as Mozambique, it is reasonable to think that consumer baskets as well as commodity prices show high temporal and geographic variation.

The methodology aims at capturing the substitution effect, a dimension that single basket PLs hardly address: household facing different prices can adjust their basket in order to keep the same level of utility. This is valid both spatially and inter-temporally; in the AFB methodology, baskets used to calculate PLs can change over time to reflect modifications in the preferences or adjustments to price variations. However, when moving from a single CBN-basket to multiple ones, the risk is that bundles result utility inconsistent, thus violating the fundamental premise that PLs should represent the same level of welfare. Ravallion and Lokshin (2006) proposed as solution a scalar correction that equalizes baskets in terms of utility, but warned "that there is no guarantee that such a scalar adjustment exists".

To address the problem Arndt and Simler (2010), after having calculated different CBN provincial baskets, develop a methodology to correct their value, ensuring that the PLs bundles produce the same level of welfare. The correction is undertaken by means of a constrained minimization where the objective is the distance between the adjusted and original budget shares and the constraints are

represented by the revealed preference condition[10] and a minimum calorie requirement (around 2150 calories per capita per day). The only variable parameter is represented by the quantities of each item in the bundles while the price vector, obtained from the survey, is fixed.

The theoretical discussion of the methodology goes beyond the purpose of this paper. In principle the entropy correction methodology seems to provide a valid tool to produce utility consistent bundles and, thus, might enable the construction of location specific PLs. Our point is different. We question the opportunity of constructing location specific poverty line -with or without entropy correction- in presence of measurement error with a geographic pattern[11].

Any location specific PL, being constructed on a smaller sample than a national poverty line, can potentially be more influenced by measurement error and produce counterintuitive results. There is a clear trade-off between the advantage of creating more locally-representative PLs and their robustness and consistency. Specifically on the entropy correction, since to our knowledge the procedure doesn't foresee any correction mechanism for the measurement error, if the input information based on location specific bundles from raw data is flawed, the entropy adjusted bundles are bound to be affected by the same problem. As we will see in details in section 3.3, when we construct location specific PLs without entropy correction, our outcomes closely reproduce the official PLs, constructed with entropy correction[12]. It follows that that the impact of entropy correction is minimal but also that the problems, we show affect these PLs, can affect also the official ones.

3.1 Why focusing on PLs?

Some measurement error is always present in household survey data and Mozambican ones are not an exception. Recalling what written in section 2, we could eventually identify 27% (2002/03) and 21 % (2008/09) of the observations (set A and B) that were likely mis-measured. This was a very restrictive definition; early attempts to circumscribe the problem produced even broader sets of observations. However, in all these cases the pure re-estimation of these observations didn't produce the expected results in particular regarding the regional distribution of poverty: after re-estimation Southern provinces still looked poorer than Northern provinces while all the other well-being indicators would indicate the opposite. The counterintuitive spatial distribution of poverty figures pointed at analyzing also the PL's and their high spatial variation

[10] Revealed preference condition rely essentially on the non-satiation condition: consumers always prefer consuming more rather than less. By applying this condition it guarantees that the cost of the bundle for the domain based PLs is the same taking relative price into account, so that each bundle delivers the same utility level.

11 Here we implicitly assume that measurement error inside every location is only in one direction, say either negative or positive. Section 2 confirms that this is the case for Mozambique. In the South and urban areas we have a problem of underestimation of food consumption while in few areas of the rural North we have a less severe problem of overestimation

12 The re-construction of official PLs with entropy correction is a complicated computer procedure we were not able to fully reproduce. However, when constructing the 13 province level bundles we followed the same procedure used in the poverty assessment, using the same calorie intake and the same procedure to identify most consumed food items and to eliminate outliers from prices and quantities consumed. As table shows, the variation from official is minimal and the geographical pattern is fully reproduced.

High PLs differentials crucially condition the calculation of poverty figures; two examples corroborate this assertion. Figure 2 (left hand graph) compares the consumption per capita distribution of urban Maputo province and rural Niassa Cabo Delgado (N&C) while Figure 2 (right hand graph) compares rural Maputo to rural Nampula[13]. Urban Maputo and rural Maputo are stochastically dominant throughout the distribution, even if, these are the areas where underestimation affects about 50% of the observations. Yet, when looking at poverty figures the picture seems opposite: while urban Maputo registers a poverty rate of 63.66% rural Cabo Delgado only registers a rate of 32.7%, and rural Maputo 76.33 % and rural Nampula 56.67% (table2, column 2).

What make the difference are the PLs, in both cases (the vertical lines in the two graphs). The PLs in urban Maputo is 93% higher than that in rural N&C while that in rural Maputo exceeds the one in Nampula by 73% (see table 4, column 6). As shown in figure 2 the magnitude of PLs (the red vertical lines) completely offsets the dominance Maputo areas have in terms of nominal income. Are these big differentials between PLs reflected by any price index directly calculated on raw data? To our knowledge this is not reflected.

We first analyzed the price differential between urban Maputo and rural N&C using a Paasche index[14] constructed on survey prices which indicates a differential of 18% while that based on SIMA[15] prices (but here Maputo is not differentiated between rural and urban) a 30% differential. Similar calculations (Sohnesen, 2011) show a price index ratio between 5%-20%[16]. As a further check, by computing Paasche indexes, we used only observations in a range of 20% below and above the PL in order to see whether the differential increased the closer we moved to the PL. The Paasche index based on food and non food items from the household survey register a differential of 30%; based only on food it is 5%, and finally the Paasche based on SIMA is 16%. Similar results are obtained when only the observations below the 55th percentile[17] are used.

The areas of Maputo and Nampula are more comparable since both are rural; yet the big PL spread is not underpinned by any price index differential. Sohnesen's simulation indicates a price differential between 31%- 40%, while our deflator either based on survey prices or on SIMA indicates no difference between the two areas. As in the previous case, we compute the Paasche index 'for the poor' and even in this case the differences are minimal: 13% for the food, non-food index, 20% using only food items and finally a -2% when using SIMA.

[13] In both cases, we truncated the upper side in order to close up on the lower percentiles of the distribution: in this way we simply disregard the upper 1st percentile. What concerns us is the rest of the distribution, in particular the part where the poor are located

[14] The Paasche index has been computed for 1060 enumeration areas and 49 food and non food items using as base price the national price for every item.

[15] SIMA is the Portuguese acronym for Agricultural Marketing System of Mozambique. Is a monthly database of agricultural prices at market level with a good national coverage.

[16] A recent policy note by Sohnesen (2011), focuses on Mozambican spatial price distribution. The note examines price distribution at the local level by calculating product prices and spatial Paasche indices under different scenarios. The basic rationale is that all the scenarios proposed to compute unit prices should not impact overall prices considerably. The fact that some scenarios show high impact on prices leads the author to argue that price data are not stable enough to support the relative high number of PLs used by MPD-DNEAP (2010).

[17] All these results are available upon request

To sum up, these findings suggest that the price spatial distribution reflected in the official PL is controversial and not necessarily reproduces the prices faced by the poor. Since poverty headcounts are calculated with these PL's, it suggests that the peculiar spatial distribution of poverty is a consequence of some problem related to construction of the PL's, which we recall, is a non standard one.

3.2 Sensitivity of PLs to measurement error

In this section we analyze, using 2008/09 data[18], how the construction of location-specific PLs is prone to be affected by the measurement error present in the data and might produce PLs that do not reflect the baskets consumed by poor people (see also Maia and van den Berg, 2010). Findings here suggest that PLs calculated using location-specific baskets are highly conditioned by the presence of measurement error. In particular, given the high concentration in these parts of the country, Southern Mozambique and urban areas are highly influenced by observations whose food component is excessively low and non-food component extremely high. These observations have an impact both on the food PLs and on the non food PLs. Having in mind the two sets of observations (A and B) we defined in section 2, we focus our attention on set A- say the underestimated observations- since they enter into the PL's estimation.

The high non-food share of these observations makes the non-food PLs higher than what, in fact, they should look like when comparing them with various baskets of local non-food prices (see appendix). In Maputo areas, for instance, the non food component accounts for about 1/3 of the total PLs causing the final lines to be twice as high compared to all the other lines in the rest of the country, yet this differential cannot be reproduced by any spatial price index (see appendix).

Regarding the food component, the impact of underestimated consumption is less clear. Broadly speaking, relying on these observations to construct 'local baskets' might produce counterintuitive results. Two types of measurement errors typically occur.

The first type of error is that many food transactions are not reported; this causes the aggregate level of food expenditures to be lower than expected. Consequences are clear in terms of food/non food *shares* but less in terms of impact of food components on the food poverty line. This is because the food PL relies on the *pattern* of consumption and not on the level[19]. Baskets relying on under-reported consumption may have a very different dietary pattern than more precise consumption. A priori, it is not clear whether food PL will be is biased upward or downward. It depends on the type of items not reported on; if relatively cheap items are not reported, the basket consumed is biased upward (as will be the food PL) and vice-versa.

The second type of error is a discrepancy between food expenditures and food quantities consumed. The household correctly reports expenditures on one item but is less capable to report the right quantities. The error clearly conditions the unitary price of the item. In case of auto-consumption,

[18] As an additional check, we reproduced the test and procedures for 2002/2003 obtaining similar results. All these results are available upon request

[19] This because PLs are rescaled to meet a minimum (2150 calories per capita per day) calories requirement

prices are directly imputed from local market surveys and, in theory, they should have passed through a screening procedure. For goods purchased prices the problems remains.

In order to compare the different sensitivity to errors of PLs, we constructed a national basket PL, and a urban rural (not reported) deflated with a location based Paasche-index and compared them with location specific PL, where location means 13 regions, a subdivision of the country that groups homogenous urban and rural areas (see table 1).

Our hypothesis is that location-based PLs are more affected by measurement error in consumption than national and urban rural. Sensitivity tests to expenditures data classified as biased (section 3) have been constructed. Two tests are reported on; 1) non-food component of the PLs; 2) total PLs.

Table 2 illustrates our first sensitivity test on non-food PLs. These are constructed exactly like those in the three Poverty Assessments[20]. Column 4 of Table 2 shows, for the 13 regions, the non-food share of observations used to calculate the non-food PLs. These households spend on non-food items the amount that becomes the non-food poverty line; their share of non-food consumption is the value reported. As noted above, the non food PLs rely heavily on observations that, due to food consumption underestimation, show a very high share of non food consumption.[21] To what extent these observations condition the non food PL construction is shown next.

Columns 5, 6, 7 of Table 2 report some simulated results obtained by eliminating underestimated observations we identified in section 3 (Set A) from the calculation of the non-food PL. The procedure is rather simple. Before constructing the various non-food PLs we eliminated the observations we singled out as suspicious, next we ran the whole procedure and once final results were obtained we compared these with the baseline scenario, notably the results obtained using the full sample[22]. As expected, both simulated non-food shares and non food PLs result in much lower figures than the baseline case in Southern Mozambique and urban areas, which suggest that the construction of the non food PL was severely affected by underestimated observations.

The direction of the bias is quite straightforward. The elimination of observations characterized by excessively high non-food shares produces lower non-food PLs and consequently, total PLs tend to be lower. Comparing columns 2 and 3 with columns with 5 and 6 of Table 2 this effect is clear. Due to the simple modification of the non-food PL, the total PLs will decrease of at least 1 to 2 Meticais (MZN) per capita per day in several urban areas, of 4 MZN in Maputo urban and in Maputo City. Do other types of PLs minimize this effect? Our simulations indicate that non-food PLs calculated at national or urban rural level are rather insensitive to underestimated observations. Therefore, it is a matter of number of

[20] We took a simple weighted average of non-food expenditures for people with expenditure between 80-120% of the food poverty line. A triangular weighting scheme is then used: weights are higher the closer a household's expenditure is to the poverty line.

[21] As discussed in section 3, the non food share of these observations is significantly higher than all the other observations

[22] At this stage of the analysis, we limit our simulation to a rather simple and mechanical procedure. We reiterate that a generalization of the results is beyond the scope of this contribution.

observations used to calculate the poverty line; the more in detail the methodology goes the higher the risk of depending on mis-measured observations[23].

In Table 3 we extend the sensitivity test to the whole construction of the PLs and compare results with a national-basket alternative; results show the PLs and their variation. As pointed out in the outset, whereas for the non-food PLs the impact of underestimation is easily detectable, for the food PLs the impact can be ambiguous: the direction of the bias can be in both directions, depending on what piece of information is missing.

The findings in Table 3 support the hypothesis we formulated above. Provincial PLs tend to be more affected by the presence of underestimated observations in their calculation. The elimination produces big variations in the PLs in the South. With the notable exception of urban Gaza and Inhambane, whose baseline value differed from the official more than other PLs, the direction of change in the area affected by underestimation is towards lower PLs and, thus, lower poverty figures[24].

In conclusion, underestimated observations tend to bias the province-specific PLs upwardly, due to both food and non-food components. By contrast, single national seems to be less affected: the adjustment for underreporting affects PLs only slightly, less than 1 Mt per capita per day in every location. Sensitivity of PLs seems to be correlated only to the number of excluded observation sets. The higher the sample of consumption observations, the lower the effect of underreporting on the PLs: in results not reported even urban rural PL's result more affected by underestimation than the national basket PL's.

4 Poverty trends: official vs. revised

4.1 Revised poverty line construction

In a recent contribution on the possible alternative to the official poverty figures, van den Boom compares the official poverty figures with figures obtained by averaging the context specific poverty line and creating a unique national poverty line. The simulation leads to striking changes in the poverty patterns in Mozambique. Poverty looks more prevalent in the rural population (65% in 2008/09, previously 57%), while the urban population is less poor (39% in 2008/09). The dynamics of poverty reduction seems to be more consistent with other data sources, including macro-economic growth and anthropometric indicators both at provincial and national level. Besides this, there are dramatic improvements in the stability of poverty trends over years and quasi-complete elimination of those puzzling poverty rate swings.

Using the information contained in the 1996/97 survey, we construct a single basket poverty line:

[23] Similar results are reported in a recent paper on Mali (Delarue et al., 2010). The use of provincial level PLs, also calculated on a limited amount of observations, leads to construct very high PLs in areas of the country traditionally richer, such as the Sikasso Province. It follows that poverty rates result to be higher in Sikasso than in less developed areas of the country.

$$\underline{q}_{97} = \sum_i w_{97i}\, q_{97i} \qquad for\ i\ \in P = \{y_i : y_i \leq \bar{y}\} \tag{4}$$

where $\underline{q}_{97}$ is a bundle typically consumed by poorest segment of the population in 1996/97 (*see appendix for a list of items and food shares*), resulting from taking the population weighted average $(\sum_i w_{97i})$ of bundles q_{97i} consumed by household below a certain consumption threshold $\bar{y}$ -in our case the 60th percentile- and scaled up to obtain 2150 calories per capita per day. The bundle $\underline{q}_{97}$ is then multiplied by a vector of prices $\underline{p}_t$

$$\underline{p}_t = \sum_i w_{ti}\, p_{ti} \qquad for\ i\ \in P = \{y_i : y_i \leq \bar{y}\}\ ;\ for\ t = 1996/97, 2002/03, 2008/09 \tag{5}$$

Combining 4 and 5 and adding $\partial \in (0,1)$ to account for the non-food share and p_{ts} , a spatial Paasche index for the 13 regions, we obtain

$$z_{ts} = p_{ts}(1 + \partial_t)\, \underline{p}_t \underline{q}_{97} \quad for\ t = 1996/97, 2002/03, 2008/09; for\ s = 1, \ldots 13 \tag{6}$$

A national fixed-basket spatially deflated poverty line z_{ts} for three survey years. Results are reported in table 4 with official ones.

4.2 Comparison between official and revised PLs

The comparison between official and revised PL's confirms a number of hypotheses discussed in previous paragraphs. First, revised PL's result spatially more stable over the years, likely reducing the 'poverty number swings' shown by official data. The Coefficient of Variation (CV) [25] calculated on the revised PLs –higher values equal higher variability- goes from 0.19 in 1996/97 to 0.20 in 2002/03 to 0.14 in 2008/09 whereas for the official PL's are 0.29, 0.46 and 0.28 respectively. When looking at the ratio of the highest PL (Maputo city in all cases) over the lowest PL, it results rather stable for the revised ones (around 1.6). To the contrary, the ratio for the official one peaks to 3.57 in 2002/03 from 2.54 and falls back to 2.31 in 2008/09.

Second, keeping the basket fixed it produces a substantial reduction in the PLs in Southern Mozambique and urban areas both in 2002/03 and 2008/09, the areas where we identified the interaction with measurement error; in 1996/97, less affected by measurement error the gap is smaller. Also interesting, in these areas, the gap between revised and official PL's widens between 1996/97 and 2002/03 and stabilizes on 2002/03's value or slightly decreases in 2008/09. Again, this might be additional evidence of the problem: the gap widens when the error is introduced and then stabilizes since also error distribution doesn't change much between 2002/03 and 2008/09 (see table 1).

One can object that different gaps in 1996/97 and the other two surveys are not necessarily related to the measurement error. In fact, it might be interpreted as consequence of fixing the basket in 1996/97: the gap between revised and official PL's widens because of changes in the baskets, i.e.

[25] $CV = \sigma/\mu$ where σ= standard deviation and μ the mean. Lower values mean lower dispersion

substitution[26]. Yet, this argument is not entirely convincing. When constructing national basket PLs *with baskets changing every year* (*see section 3.2 and appendix*), values are closer to our revised PL'S suggesting that the inter-temporal basket substitution is minimal. The gap, thus, is either capturing some spatial basket substitution or, as we argue, is an artifact of interaction local-based PL's and measurement error.

Finally, the comparison between the two groups presents another interesting piece of information, notably their rate of growth. This provides, at 13 provincial level, a measure of the change in the cost, measured in Meticais, of living at exactly the poverty line level. By design, this particular measure of inflation is most relevant for the poor as the contents of the food bundle reflect items that are consumed by the poor. Whereas the rate of growth of PL's between 1996/97 and 2002/03 is different, leading to an increased gap in 2002/03, revised and official PL's between 2002/03 and 2008/09 show a similar geographical pattern of growth (Figure 4) and high and significant rank correlation (0.74). As extensively discussed in the PA, the pattern of price increase between 2002/2003 and 2008/09 is consistent with all the available price data and helps to explain some of the poverty outcomes (DNEAP, 2010). Notably, the sharp price increase, together with a stagnant agricultural productivity, contributed to the bad poverty performance of Central provinces. Likewise, according to the PA, the relatively slower growth of PL's in Northen urban areas triggered a fast reduction of poverty and, *we add,* in general in urban areas and in the South (compare figure 4 and table 7, last column).

To further analyze the robustness of revised PLs vis a vis the official, we constructed a correlation matrix that shows poverty rates and consumption proxies for 2002/03 and 2008/09 at the 13 provinces level, in Table 5. Using the revised PL, the 13 provincial ranks of poverty rates seem more consistent with the one of poverty correlates, given that the correlations are negative and statistically significant with proxies associated to lower poverty (higher wealth index, more access to services, head's years of education), and positive and statistically significant with proxies associated to higher poverty (higher % of stunted and underweight children in the household). By contrast, official poverty rates show a non-significant rank correlation with all correlates, again casting doubts on their validity.

In addition a standard consumption prediction model is presented in Table 6 and can be expressed as:

$$\log\left(\frac{y_i}{z_{jk}}\right) = \beta_{1k}HEAD_i + \beta_{2k}EDU_i + \beta_{3k}SHOCKS_i + \beta_{4k}ASSET_i + \beta_{5k}AGRIC_i + MONTH + \varepsilon_{ik} \quad for\ k = 1,2 \qquad (7)$$

where y_i refers to total household per-capita/day official consumption expenditure, z_k the two different groups of PLs, and subscript *i* and *j* households and locations, respectively.

Since the scale of the dependent variable across the two models slightly differs, due to the level of the PLs used (see table 3), the simple difference in the t-statistics of each parameter is reported, instead of the statistical difference in the coefficient usually reported in the seemingly unrelated

[26] Revised PLs are constructed fixing the basket at national level in 1997, the whole inter-temporal variation, thus, is explained by prices

regression equations (SURE) models. A higher absolute difference of each parameter can be interpreted as better capacity of the ratio consumption/PL to be explained by conventional poverty correlates. Results indicate that the model using revised PLs (B) fits standard poverty prediction better. The adjusted R-square of this model is 22% higher than the one of the model using the official PLs. Moreover, the difference in *t* statistics of the variables highly associated to poverty (assets and access to services) are all positive and relatively high, except the one associated to household ownership of bike, as the latter is an indicator of relative poverty (Hanlon and Smart, 2008).

4.3 Revised poverty outcomes: 1996/97-2002/03-2008/09

Combining section 3 and section 4 results, we calculated revised poverty figures using revised PLs and prediction from the 1996/97 model (section 2.3). Table 7 compares the revised poverty figures with official ones. In figure 5 we grouped poverty figures by urban-rural macro areas.

Three elements are worth noting. First at national level the revised figures show a different trend compared to the official. Rather than fast reduction of poverty in the first period (1996/97) and stagnation in the second (2002/03), a sort of L shaped trend, we observe a change in the gradient of poverty reduction; from 2002/03 to 2008/09 the poverty reduction trend *slowed down* but did not *stop* (see also van den Boom 2010). As shown by the confidence intervals in the last row of table 7, the revised values of poverty at national level in 2002/03 and 2008/09 are statistically different and indicate a declining trend.

Second, official and predicted poverty rate *levels* are across regions similar in 1997 (South richer than the other two macro-regions) but differ noticeably in the other two rounds. According to the official figures, in 2002/03 poverty is concentrated in the South: rural Maputo is the poorest rural area and urban Maputo the second poorest urban area. The rest of the country, with the exception of rural North, shows poverty rates below 50%. In 2008/09, poverty figures face a sharp re-ranking and Central provinces become, together with Maputo area the poorest in the country.

Our revision, instead, shows that in 2002/03 the Southern region, again both urban and rural, seems to be better-off than the corresponding Central and Northern regions, with urban faring always better than rural areas. This pattern is confirmed in 2008/09. Poverty is low in urban areas and in South while is high in rural Central and Northern Mozambique.

Third, official and predicted poverty rate *changes* also diverge quite strikingly across regions. According to the official one, poverty ratio drops substantially only in Niassa and Cabo Delgado, both urban and rural, while there is a moderate reduction in rural and urban South. Poverty rates stagnate in rural Nampula while increase in Central region -both in urban and rural areas-and in Nampula urban.

Our revision indicates a rather different pattern (see figure 5). Already starting from a lower level of poverty, urban areas in the South reduce poverty rates to 15% in 2008/09. There is some spill over from urban to rural areas; poverty, thus, drops from 70% in 1996/97 to about 45% in 2008/2009: rural South becomes the least poor rural area in the country and one of the fastest reducer. In the rest of the country, urban areas fail to trigger substantial improvement in the neighboring rural areas. The

important difference that distinguishes Southern Mozambique and more specifically the Maputo area from the rest of the country is the capacity of urban areas to spread the benefits to the hinterland, whereas in the Center and the North, urban areas remain island of relative well-being surrounded by big pockets of extreme poverty. The lack of spillover from urban areas relatively rich to the poorer rural areas indicates dramatic market incompleteness. In conclusion, with the notable exception of urban Central, poverty reduction is driven by the stable reduction of poverty in urban areas and the good performance of the Southern part of the country.

Analysis of the distribution of real consumption (Figure 6) overcomes the potential disadvantage of looking at only one measure of welfare -the poverty line- and overlooking the rest of the distribution. The horizontal axis represents revised consumption measured in terms of percent of the poverty line. The vertical axis represents the population percentage. Each point on the distribution function shows the share of the population falling below a specified consumption level[27]. The poverty levels in the three survey years can be read from the distribution functions at the point where the function crosses the vertical line that indicates 100% of the poverty line (DENEAP, 2010). The distribution can be compared in welfare terms; distributions laying on the right to the others are said statistically dominant and can be considered a welfare improvement.

Across the national distribution (top left graph), the main conclusions on the poverty trends at national and macro-level but also at macro-region level hold. 2002/03 survey saw a dramatic poverty reduction and a general improvement of the welfare. However, higher income grew faster than lower income, therefore, inequality increased. It is worth noting, that this unequal growth has a clear spatial connotation. The welfare improvements in the South between 1996/97 and 2002/03 (figure 5, bottom right graph) were twice as higher than in the other two macro-regions.

Between the last two surveys, the rate of poverty reduction decreased but not stopped. The distribution in 2009 is still superior but only above the 60% of the PLs; it confirms that the poorest were not benefited by any type of growth. These are likely located in Central Mozambique (figure 5 bottom left graph) where conditions for most of the household went back to 1996/1997 levels. Finally, the performance of South is less pronounced than in the previous period but able to bring down poverty to about 30% and clear improvements can be seen all over the distribution.

5. Conclusions

In Mozambique over the last decade the debate around poverty rates, trends but also data quality has been intense. The aim of our analysis is to take part to the debate by providing alternative methods to compute poverty rates. According to our analysis, two intertwined factors have contributed to the official reporting of poverty stagnation between 2002/03 and 2008/09: underreporting of consumption expenditures on one side, and use of location-specific PLs on the other. In the present work methods are proposed to cope with underreporting and construct a more standard PLs based on a single national basket constructed with data from the first available survey (1996/97) combined with

[27] Given the interest on the performance of the poorest, we cut the distribution above 200% of the poverty line

price data from each round. This substantially differs from the official poverty lines based on province-specific and over-time-flexible bundles corrected with an entropy-based adjustment (MPD-DNEAP, 2010; MPF/IFPRI/PU, 2004).

Looking at the construction of province-specific PLs, we detected the likely cause of the high provincial differential. Observations affected by underreporting, characterized by a high non-food consumption share, tend to bring PLs upward. As an alternative to the provincial-specific official PLs, we constructed a single national PL and two PLs by urban and rural, respectively. Besides being more consistent with standard poverty correlates, both at the regional and household level, national and urban/rural PLs seem to be less affected by measurement error since they rely on a larger sample in calculating the baskets.

Results of our revision highlight that the substitution of province-specific PLs with a national 1996/97 basket PL combined with a partial re-estimation of observations (about 1/4 in the two years) likely to be affected by underreporting yields in 2002/03 and 2008/09 poverty levels more consistent with other socio economic indicators. Poverty is relatively lower in Southern Mozambique and in urban areas while dramatically high in Central and Northern Mozambique and in rural areas. Compared to the official figures, this represents a complete re-ranking of provinces based on poverty incidence. Also, our calculations do not seem to support the enormous spatial price variation underpinning large changes in provincial PLs and, in turn, the poverty headcounts swings between 2002/03 and 2008/09.

Looking at poverty dynamics, our revision indicates a different trend of poverty reduction. Rather than stagnation, Mozambique is experiencing a decrease of poverty in urban areas where only 30% of population resides and a substantial invariance in the most populated rural areas. Yet, the rural urban divide is not homogenous. The important difference that distinguishes Southern Mozambique, specifically the Maputo area, from the rest of the country is the capacity of urbanized areas to *pull* some of the surrounding rural areas, whereas in the Center and the North, urban areas remain island of relative well-being surrounded by big pockets of extreme poverty, indicating a serious problem of market incompleteness.

References

Arndt C., K. Simler (2010). Estimating Utility-Consistent Poverty Lines with Applications to Egypt and Mozambique, *Economic Development and Cultural Change*, 58(3): 449-474.

Arndt C., A. Garcia, F. Tarp and J. Thurlow (2010). Poverty reduction and economic structure: Comparative path analysis for Mozambique and Vietnam, *UNU-WIDER Working Paper*, n. 122, United Nations University, World Institute for Development Economics Research, Helsinki.

Azzarri, C., Carletto, G., Davis, B., Nucifora, A., Sohnesen T.P. (2011). Child Undernutrition in Mozambique, World Bank/UNICEF, Maputo.

Boom van den B. (2010). Analysis of poverty in Mozambique, household poverty status, child malnutrition and other indicators 1996/97, 2002/03, 2008/09, study commissioned by the Royal Dutch Embassy in Maputo, mimeo.

Cunguara, B. and Joseph Hanlon, J. (2010). Poverty is not being reduced in Mozambique, *Crisis States Working Papers Series*, no.2 (74), DESTIN, London School of Economics, London.

Delarue J, Mesple-Somps S., Naudet J.D, Robilliard A.S. (2009). The Sikasso Paradoxe: Cotton and Poverty in Mali, *Document de Travail*, n. 2009-09, Développement, Institutions et Mondialisation (DIAL), Paris.

Deaton A. (1997). The Analysis of Household Surveys: A Microeconometric Approach to Development Policy, World Bank

Filmer, D., and Pritchett, L. (2001). Estimating Wealth Effects Without Expenditure Data or Tears: An Application to Educational Enrollments in States of India, *Demography,* 38(1): 115–32.

Hanlon, J. and Smart, T. (2008). Do Bicycles Equal Development in Mozambique*,* Woodbridge: James Currey.

Maia C., van der Berg S. (2010). When the remedy is worse than the disease: Adjusting survey income data for price differentials, with special reference to Mozambique, *MPRA Working Paper,* n. 26571.

MPD-DNEAP (2010). Poverty and Wellbeing in Mozambique: Third National Poverty Assessment, Ministry of Planning and Development - National Directory of Studies and Policy Analysis-, Maputo.

MPF/IFPRI/PU (2004). Poverty and Wellbeing in Mozambique: Second National Poverty Assessment, Ministry of Planning and Finance/Purdue University/International Food Policy Research Institute, Maputo.

MPF/UEM/IFPRI (1998). Understanding Poverty and Well-being in Mozambique: The First National Assessment, Ministry of Planning and Finance/Eduardo Mondlane University/International Food Policy Research Institute, Maputo.

Ravallion, M. and Lokshin, M. (2006). On the consistency of poverty lines, in *Poverty, Inequality and Development: Essays in Honor of Erik Thorbecke*, ed. A. de Janvry and R. Kanbur, Amsterdam: Kluwer Academic Publishers.

Sohnesen T. (2011). Are price data from IOF 2008/09 stable enough to support multiple poverty lines and entropy adjustment?, mimeo.

UN (2010). Millennium Development Goals Report, United Nations, New York.

World Development Indicators (2011), accessed online at http://data.worldbank.org/indicator, The World Bank, Washington, DC.

Maps, graphs, and tables

Map 1: Population density by regions and districts

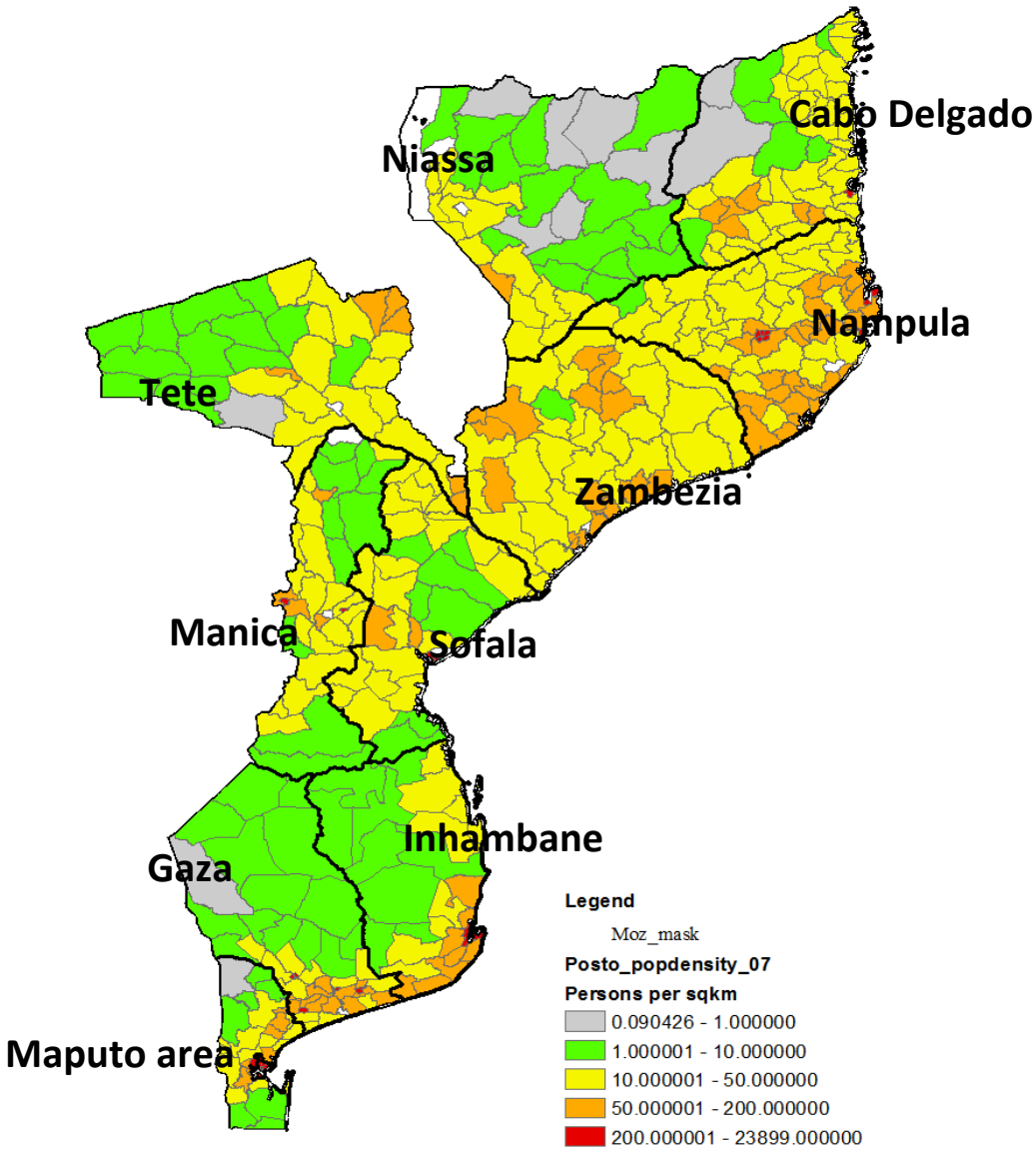
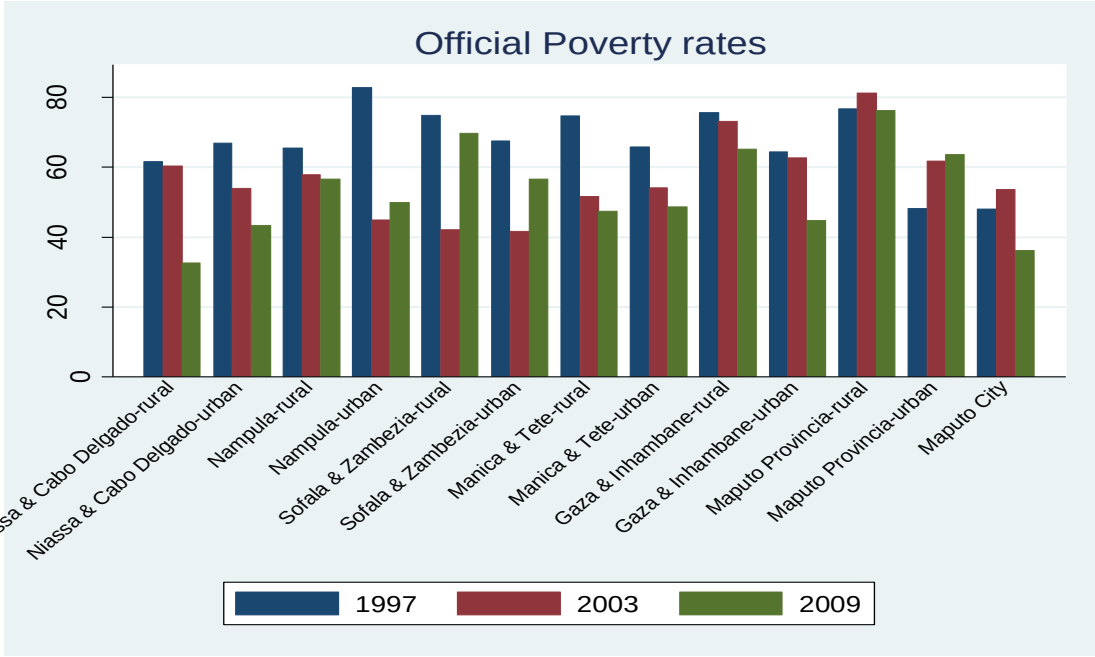


Figure 1



Official: 2008/09 PLs (MPD-DNEAP, 2010); **Simulated:** 1996/97 official PLs deflated by the consumer price index from World Development Indicators (2011)

Figure 2

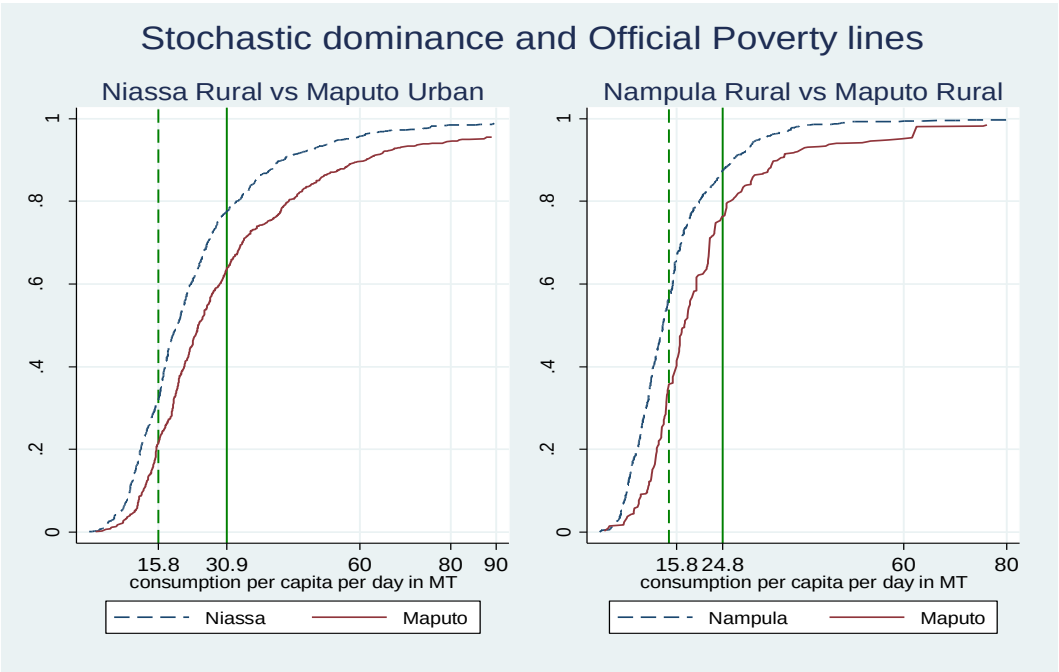
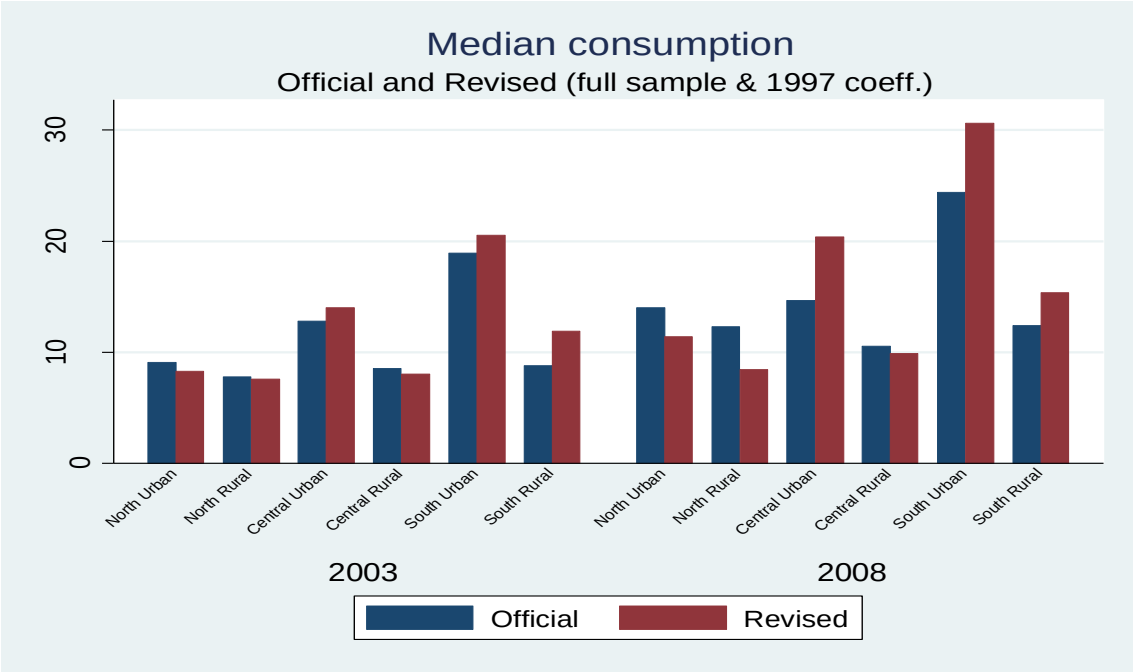


Figure 3



Official: IAF 2002/03 and IOF 2008/09 original household survey consumption data; **Revised**: out-of-sample predicted consumption using 1996/97 coefficients

Table 1: Sets A, B observations and shares (IAF 2002/03-IOF 2008/09)

	2002/03				2008/09			
13 provinces	Set A		Set B		Set A		Set B	
	Obs.	%	Obs.	%	Obs.	%	Obs.	%
Niassa & Cabo Delg.-rur	5	1	231	24	2	0.1	128	13
Niassa & Cabo Delg.-urb	59	10	33	6	39	9	18	3
Nampula-rural	1	0.1	117	22	8	2	232	23
Nampula-urban	29	13	12	6	56	12	35	6
Sofala & Zambezia-rural	10	1	310	30	17	2	291	19
Sofala & Zambezia-urban	87	17	11	2	168	25	27	3
Manica & Tete-rural	19	2	228	25	21	2	233	22
Manica & Tete-urbana	134	20	25	4	95	21	22	4
Gaza & Inhambane-rural	142	14	89	9	126	19	64	7
Gaza & Inhambane-urban	171	31	9	2	151	27	12	2
Maputo Province-rural	84	29	3	1	53	43	1	1
Maputo Province-urban	261	48	0	0	332	57	1	0.1
Maputo City	381	41	0	0	313	35	0	0
Total	**1383**	**15**	**1068**	**12**	**1381**	**12**	**1064**	**9**

Set A: Consumption poor/ Wealth index non-poor; **Set B:** Consumption non-poor/ Wealth index non-poor;

Table 2: Sensitivity of Official Non-Food Poverty lines (Mt per capita per day) to Set A

13 provinces	Official			Simulated eliminating underestimated (set A)		
	Non-Food P.L	Poverty Line	Non-food share	Non-Food P.L	Poverty Line	Non-food share
Niassa & Cabo Delgado-rural	3.43	15.95	0.30	3.75	16.25	0.30
Niassa & Cabo Delgado-urban	5.01	18.91	0.39	5.17	19.17	0.37
Nampula-rural	3.21	14.33	0.33	3.63	14.73	0.33
Nampula-urban	4.22	16.72	0.39	4.06	16.56	0.33
Sofala & Zambezia-rural	3.30	14.35	0.29	3.12	14.52	0.27
Sofala & Zambezia-urban	5.39	19.07	0.42	4.15	17.85	0.31
Manica & Tete-rural	4.21	19.39	0.29	4.34	19.54	0.28
Manica & Tete-urban	5.93	21.47	0.40	4.97	20.57	0.32
Gaza & Inhambane-rural	5.31	18.37	0.43	5.07	18.17	0.39
Gaza & Inhambane-urban	6.32	20.31	0.47	5.78	19.88	0.41
Maputo Provincia-rural	7.45	24.84	0.43	6.62	24.52	0.39
Maputo Provincia-urban	10.25	30.86	0.50	6.25	26.95	0.31
Maputo City	12.34	33.14	0.59	7.05	27.95	0.35

Official: 2008/09 PLs and shares (MPD-DNEAP, 2010)**; Set A:** Consumption poor/ Wealth index non-poor households.

Table 3: Sensitivity of Poverty lines (Mt per capita per day) to underestimated observations (Set A)

13 provinces	Set A		National basket P.L			Location sp. basket P. L		
	Obs	Share	Full sample	Without set A	**Diff.**	Full sample	Without set A	**Diff.**
Niassa & C. Delgado-rural	2	0.00	18.79	18.11	**0.68**	18.13	18.14	**-0.01**
Niassa & C. Delgado-urban	39	0.09	20.61	19.87	**0.74**	22.89	22.16	**0.73**
Nampula-rural	8	0.02	14.57	14.05	**0.52**	10.36	10.30	**0.06**
Nampula-urban	56	0.12	16.55	15.96	**0.59**	13.16	12.50	**0.66**
Sofala & Zambezia-rural	17	0.02	15.81	15.25	**0.56**	13.36	13.21	**0.15**
Sofala & Zambezia-urban	168	0.25	17.62	16.99	**0.63**	19.13	17.68	**1.45**
Manica & Tete-rural	21	0.02	20.87	20.12	**0.75**	19.11	19.05	**0.06**
Manica & Tete-urban	95	0.21	20.19	19.46	**0.73**	21.40	20.00	**1.40**
Gaza & Inhambane-rural	126	0.19	15.99	15.42	**0.57**	19.18	18.77	**0.41**
Gaza & Inhambane-urban	151	0.27	17.04	16.43	**0.61**	15.87	16.52	**-0.65**
Maputo Provincia-rural	53	0.43	18.74	18.07	**0.67**	28.37	27.75	**0.62**
Maputo Provincia-urban	332	0.57	22.59	21.78	**0.81**	32.07	28.02	**4.05**
Maputo City	313	0.35	22.31	21.51	**0.80**	33.78	28.79	**4.99**
National	**1381**	**0.12**	**17.80**	**17.20**	**0.60**	**19.20**	**18.40**	**0.80**

National: single national PL and 13 provincial price deflators; **Location specific**: 13 PLs based on local food basket, without entropy correction.

Table 4: Official and Revised PLs (Mt per capita per day): 1996/97, 2002/03, 2008/09 by 13 provinces and national level

13 provinces	1996/97		2002/03		2008/09	
	Official	Revised	Official	Revised	Official	Revised
Niassa & C. Delgado-rural	4.02	4.33	7.10	7.72	15.95	17.33
Niassa & C. Delgado-urban	5.43	4.47	10.23	10.60	18.91	19.01
Nampula-rural	3.36	4.52	5.97	6.37	14.33	13.44
Nampula-urban	4.95	5.13	6.66	7.52	16.72	15.27
Sofala & Zambezia-rural	4.85	4.41	5.47	7.01	14.35	14.58
Sofala & Zambezia-urban	7.60	6.19	8.77	9.12	19.07	16.25
Manica & Tete-rural	4.71	4.45	6.93	7.99	19.39	19.25
Manica & Tete-urban	7.41	6.02	9.69	9.69	21.47	18.62
Gaza & Inhambane-rural	6.43	5.99	9.01	6.74	18.37	14.75
Gaza & Inhambane-urban	7.83	5.78	10.72	8.00	20.31	15.71
Maputo Provincia-rural	7.32	6.10	16.76	9.40	24.84	17.28
Maputo Provincia-urban	8.71	7.19	18.30	11.33	30.86	20.83
Maputo City	8.54	7.21	19.52	11.77	33.14	20.57
National	**5.27**	**5.03**	**8.47**	**7.94**	**18.41**	**16.41**

Official: 13 flexible basket PLs based on location specific baskets, with entropy correction **Revised**: national 1996/97 based basket PL with 13 location price deflators;

Figure 4

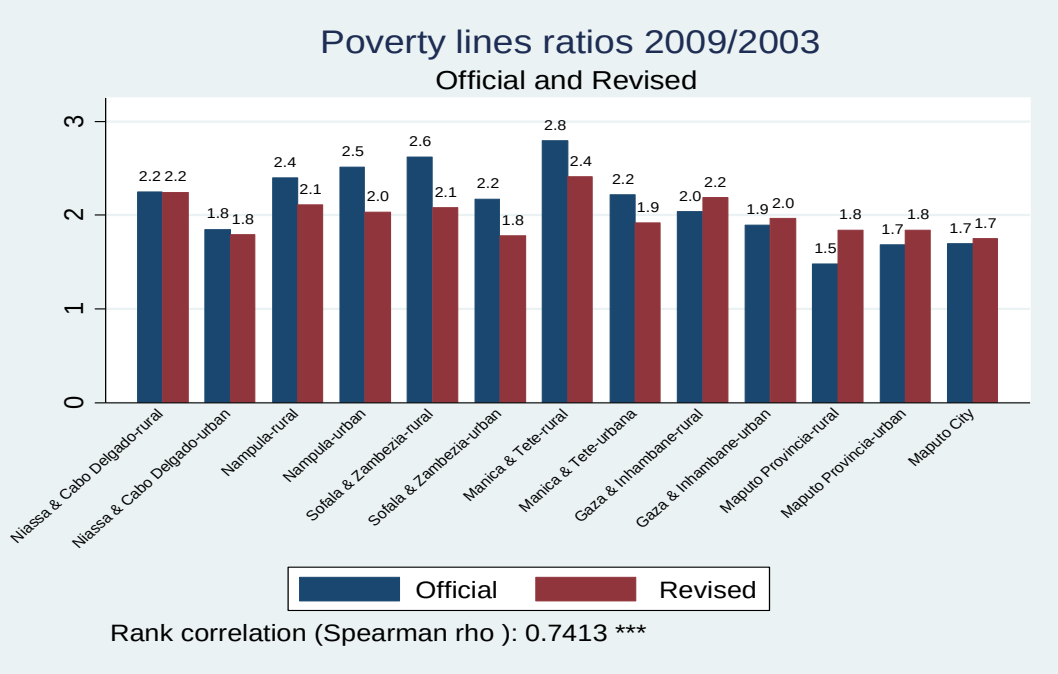


Table 5: 13 province rank correlation (Spearman rho) between poverty rate and its correlates using different PLs (IAF 2002/03 and IOF 2008/09)

Year	2002/03		2008/09	
	Official	**Revised**	**Official**	**Revised**
Wealth Index	-0.06	-0.95**	-0.17	-0.70**
Piped water (%)	-0.27	-0.70**	-0.24	-0.66**
Toilet (%)	-0.13	-0.92**	0.0	-0.68**
Electricity (%)	-0.20	-0.89**	-0.21	-0.64**
Concrete walls (%)	0.09	-0.95**	0.05	-0.59**
Household head education (years)	-0.31	-0.49	-0.36	-0.73**
Kids stunted (%)	-	-	-0.19	0.56**
Kids underweight (%)	-	-	-0.12	0.54*
note: *** p<0.01, ** p<0.05, * p<0.1				

Official: 13 location-specific PLs (MPD-DNEAP, 2010); **Revised:** national 1996/97 based basket PL with 13 location price deflators;

Table 6: Official and revised consumption/PL OLS regression models (dependent variable: ratio between total household consumption per-capita/day and PLs; difference=national - official)

	Model with official PLs		Model with revised PL		Difference of t- values
	Coef	**t**	**Coef**	**T**	
Household characteristics					
Household head age	0.009***	5.500	0.010***	6.342	**0.842**
Hh head age squared	-0.000***	-5.837	-0.000***	-6.668	**0.832**
Head is female	0.006	0.398	0.016	1.007	**0.609**
Head is widow/er	-0.081***	-3.654	-0.072***	-3.327	**-0.327**
Hh head is polygamous	0.106***	5.205	0.090***	4.554	**-0.651**
Household size	-0.123***	-35.800	-0.123***	-36.599	**0.799**
Avg adult yrs of age	0.002*	1.852	0.002**	2.517	**0.664**
Dependency ratio	-0.087***	-11.925	-0.089***	-12.543	**0.618**
Education and work					
Unemployment ratio + discouraged and seasonal	-0.076***	-2.769	-0.098***	-3.658	**0.889**
Education of household head	0.032***	14.901	0.029***	13.711	**-1.190**
Hh not engaged in agr. Activity	0.024	1.067	0.094***	4.305	**3.238**
Shocks					
Hh had a relevant natural shock during past 5 years	-0.025*	-1.760	-0.023*	-1.665	**-0.095**

Hh had a relevant death-related shock during past 5 years	-0.010	-0.660	-0.027*	-1.734	**1.074**
Assets and access to services					
Hh has piped or protected water source	0.033*	1.827	0.085***	4.824	**2.997**
is the make of roof good?	-0.042**	-2.372	0.075***	4.352	**1.980**
is the make of floor good?	0.069***	3.871	0.119***	6.799	**2.928**
Hh has any toilet	0.310***	10.787	0.333***	11.881	**1.094**
Hh has electricity?	0.186***	7.833	0.178***	7.686	**-0.147**
Hh has cell phone?	0.201***	11.289	0.250***	14.433	**3.144**
Hh owns tv	0.198***	8.564	0.213***	9.420	**0.856**
Hh has any bike	0.167***	12.197	0.119***	8.893	**-3.304**
Hh has car?	0.849***	22.328	0.856***	23.071	**0.744**
Agricultural assets					
Tropical Livestock Units: cattle	0.005	0.839	0.005	1.017	**0.178**
Tropical Livestock Units: sheep	0.315***	3.038	0.314***	3.102	**0.064**
Tropical Livestock Units: goats	0.065***	3.545	0.062***	3.501	**-0.044**
Tropical Livestock Units: pigs	0.069***	3.424	0.077***	3.895	**0.471**
Tropical Livestock Units: chicken	0.092**	2.434	0.089**	2.410	**-0.024**
Tropical Livestock Units: ducks	0.310	1.528	0.379*	1.916	**0.388**
Tropical Livestock Units: other animals	0.598***	4.293	0.501***	3.686	**-0.607**
Hh sold maize	0.126***	5.461	0.102***	4.564	**-0.898**
Hh cultivated maize	0.175***	11.129	0.165***	10.776	**-0.353**
Hh sold cassava	0.171***	5.820	0.168***	5.844	**0.024**
Hh cultivated cassava	-0.016	-1.121	-0.004	-0.260	**-0.861**
Hh sold potatoes	0.055	0.952	0.034	0.608	**-0.345**
Hh cultivated potatoes	-0.005	-0.199	-0.008	-0.342	**0.143**
R2	0.427		0.532		
Adj-R2	0.425		0.53		
F-statistics (46, 10715)	173.54		264.73		
Observations	10762		10762		
note: *** p<0.01, ** p<0.05, * p<0.1					

*Monthly dummies are not reported

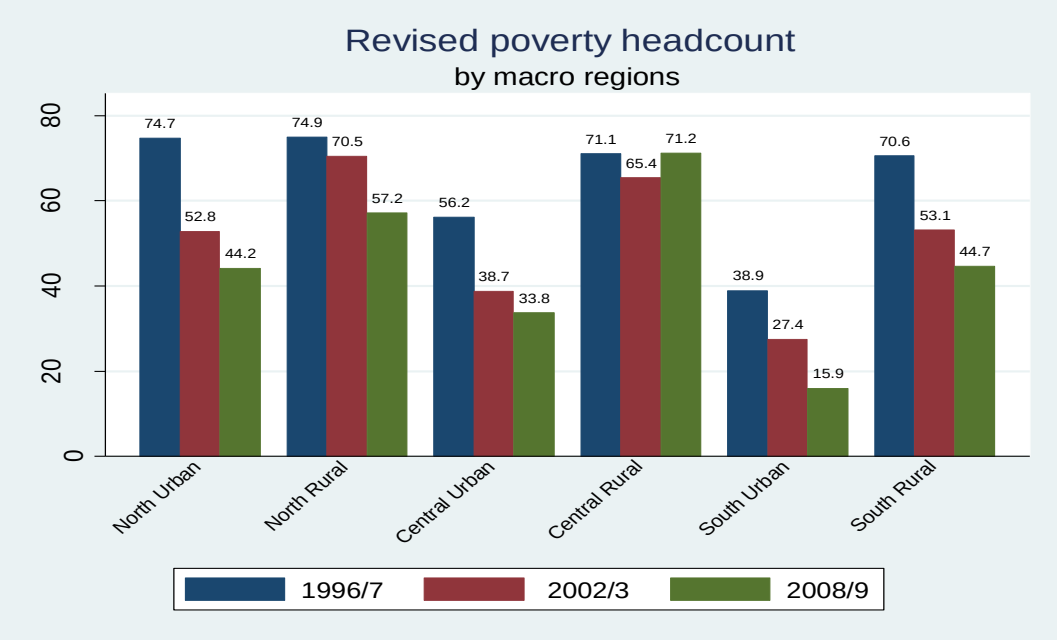
Official: 13 location-specific PLs (MPD-DNEAP, 2010); **Revised :** single 1996/97 basket national PL and 13 provincial price deflators

Table 7: Poverty headcounts in 1996/97, 2002/03, 2008/09 by 13 provinces and national level

13 Regions	1996/97		2002/03		2008/09	
	Official	Revised	Official	Revised	Official	Revised
Niassa & C. Delgado-rural	61.63	66.34	60.37	71.74	32.70	46.44
Niassa & C. Delgado-urban	66.88	56.43	53.93	51.10	43.38	40.06
Nampula-rural	65.44	81.16	57.81	69.32	56.67	65.68
Nampula-urban	82.76	82.90	44.91	53.47	49.90	46.38
Sofala & Zambezia-rural	74.83	70.69	42.07	64.28	69.69	76.70
Sofala & Zambezia-urban	67.18	58.04	41.72	37.72	56.66	35.67
Manica & Tete-rural	74.60	71.98	51.62	67.68	47.51	62.27
Manica & Tete-urban	65.57	53.46	54.09	40.27	48.67	30.01
Gaza & Inhambane-rural	75.31	71.68	73.14	54.49	65.19	45.76
Gaza & Inhambane-urban	63.77	42.75	62.72	31.61	44.86	21.17
Maputo Provincia-rural	76.81	66.39	81.18	46.12	76.33	39.26
Maputo Provincia-urban	48.24	41.35	61.83	30.40	63.66	19.96
Maputo City	47.84	37.14	53.60	23.20	36.15	9.53
National	**69.38**	**68.38**	**54.07**	**56.40**	**54.69**	**52.08**
Confidence intervals: National	**68.4-70.46**	**67.4/ 69.4**	**53.0/ 55.1**	**55.4/ 57.5**	**53.7/ 55.6**	**51.1/ 53.0**

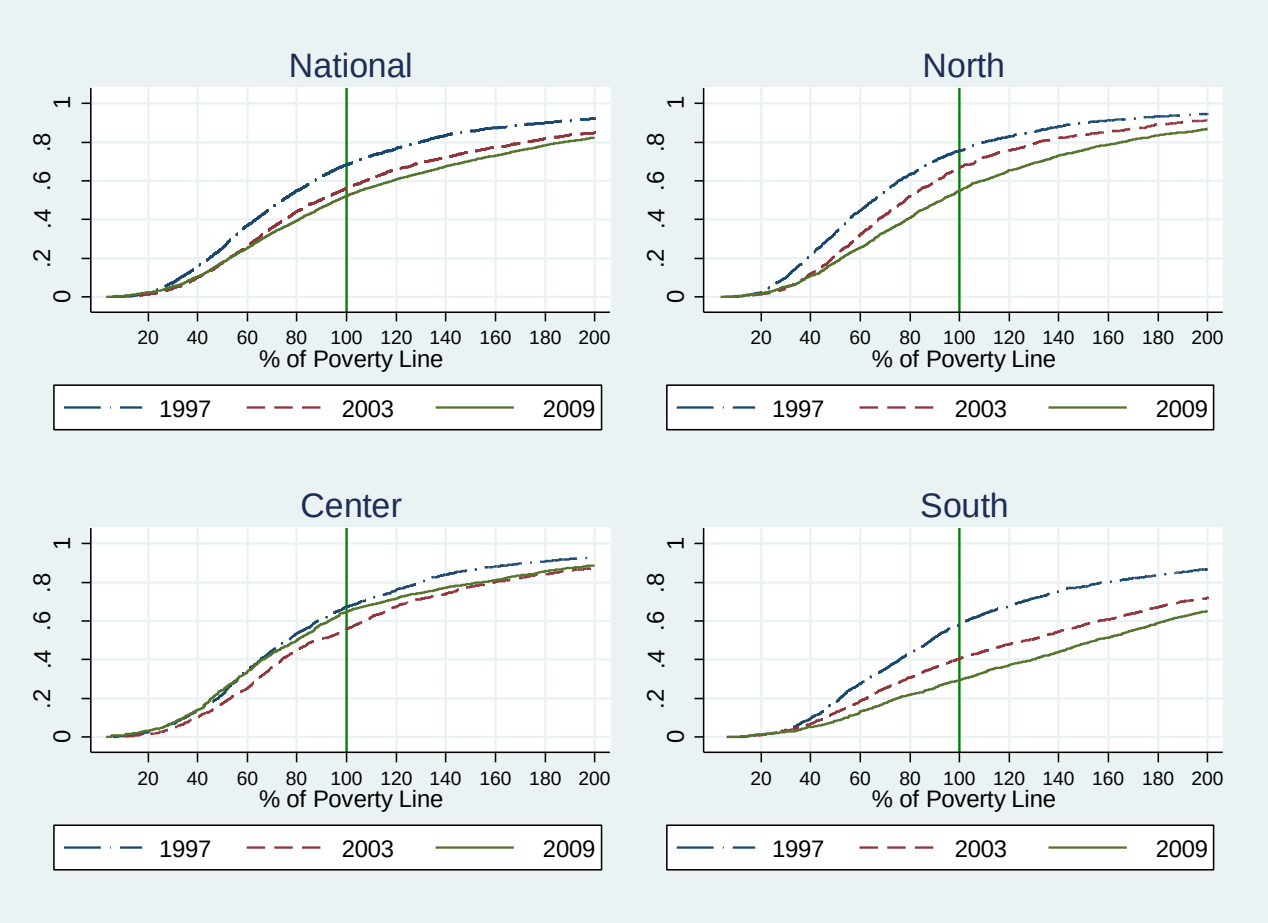
Official: 13 flexible basket PLs based on location specific baskets, with entropy correction (MPD-DNEAP, 2010)**; Revised:** national 1996/97 based basket PL with 13 location price deflators and 2002/03 and 2008/09 data partially re-estimated (set A and B table1)

Figure 5



Revised: national 1996/97 based basket PL with 13 location price deflators and 2002/03 and 2008/09 data partially re-estimated (set A and B table1)

Figure 6: Distribution function of revised real consumption



Revised: national 1996/97 based basket PL with 13 location price deflators and 2002/03 and 2008/09 data partially re-estimated (set A and B table1)