The Growth-Poverty-Inequality Nexus in Malawi: A recomputation

Chirwa, Themba

Malawi Millennium Development Trust

5 May 2022

THE GROWTH-POVERTY-INEQUALITY NEXUS IN MALAWI: A RECOMPUTATION

BY

Themba Gilbert Chirwa, Ph.D.

Department of Economics

Malawi Millennium Development Trust

PO Box 31513

Lilongwe

Malawi

Email: themba.chirwa@mmd.gov.mw/ tchirwa@gmail.com

MAY 2022

THE GROWTH-POVERTY-INEQUALITY NEXUS IN MALAWI: A RECOMPUTATION

Abstract

The growth-poverty-inequality hypothesis depicts an inverse relationship between economic growth and poverty and an ambiguous relationship with inequality. However, the official national statistics in Malawi reveal a positive association between economic growth and poverty, falling from 52.4% in 2004 to 50.7% in 2019. The official results also depict that Malawi faces big income gaps, evidenced by inequality measures of more than 0.40. We noted several computational errors in estimating the official poverty and inequality measures. They include zero standard errors, significant outliers in consumption aggregates, fewer primary sampling units used, and the median approach to deal with outliers. After observing inconsistent trends between the official National Statistics Office and economic growth, the study aimed to recompute poverty and inequality measures to correct such errors. Contrary to the official results, poverty in Malawi increased significantly from 48.3% in 2004 to 61.5% in 2019, depicting an inverse relationship with economic growth. The approach adopted also downgraded the inequality trends from a big income gap to adequate income inequality. The results confirm that the growth-poverty-inequality hypothesis holds in Malawi. The study, therefore, supports the need to ensure data reliability to aid policymakers in making sound policy decisions.

Keywords: Growth; Poverty; Inequality; Malawi

JEL Classification: O47, I32, D63, N17, N37

1. Introduction

Growth, poverty, and inequality are important metrics for developing countries worldwide (Haughton and Khandker, 2009). However, many studies that focus on understanding the determinants of economic growth mainly focus on macroeconomic covariates with little linkages to poverty and inequality as deterrents to growth (Chirwa and Odhiambo, 2019; 2020). As economies grow over time, one expects the same transition into increases in average incomes, particularly those at the bottom percentile. However, this is usually not the trend in many economies facing high poverty and inequality. Understanding, linking, and decomposing economic growth into poverty metrics is very important as it also helps nations understand who to target in their quest for economic growth and development (Araar *et al.*, 2010; Ravallion, 2016).

The key fundamental metrics used widely to determine the extent of poverty are consumption expenditure aggregates as a proxy of average incomes and measures of inequality within such income distributions (Haughton and Khandker, 2009; Ravallion, 2016). The 2030 Sustainable Development Goals (SDGs) are clear that for economies to achieve inclusive growth, nations worldwide need to end all forms of poverty and reduce inequality within and among countries (Adeleye *et al.*, 2020). For economies to achieve this, they need meaningful economic growth of not less than an average of 6.0% per annum (Bassanini *et al.*, 2001). The key indicators guiding this goal include per capita consumption expenditure growth (poverty indicators), Gini coefficients (inequality measures), and real Gross Domestic Product (GDP) growth (economic growth measures).

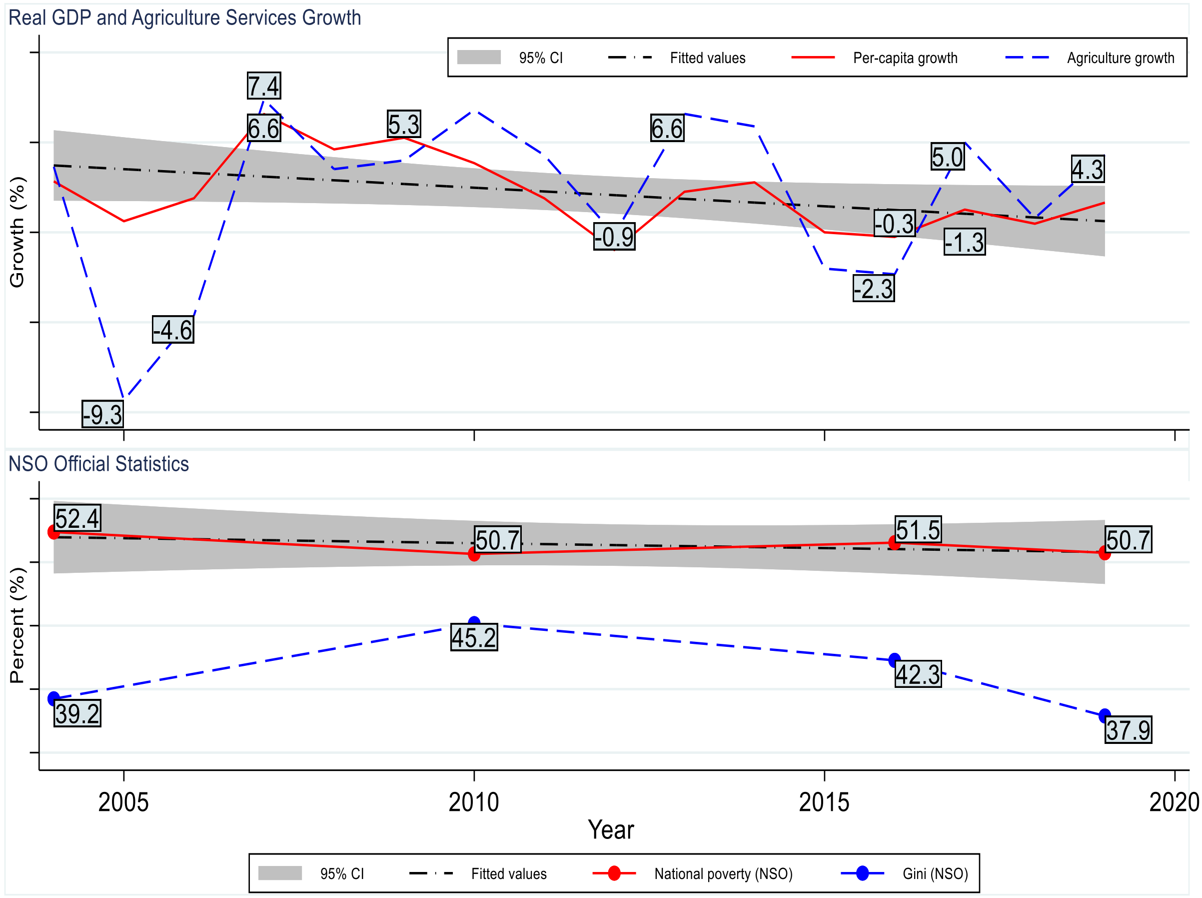
The generation of such underlying data used to estimate the trilemma statistics is usually through the collection of surveys from each country's National Statistics Office (NSO). Periodically the NSO officially publish different types of publications containing such data, mainly through National Accounts, Living Standard Measurement Surveys (LSMS), among others, which eventually find their way into international databases such as the World Development Indicators (WDI) or the World Economic Outlook (WEO) databases published by the World Bank (WB) and the International Monetary Fund (IMF).

As a case study, we investigate the growth-poverty-inequality trilemma in Malawi using official statistics from the National Statistics Office. We found inconsistent results with the general philosophy guiding the growth-poverty-inequality nexus from several fronts. Firstly, from 2004 to 2019, the Malawi economy experienced declining growth rates, especially negative growth in the agricultural sector, which supports more than 80 percent of Malawians. Malawi experienced the lowest growth rate in the agricultural sector of -9.3% at the end of 2005 and continued in 2006 when the sector registered a negative growth rate of -4.6% (World Bank, 2021). The positive growth rates that Malawi experienced from 2007 to 2010 were not significant to overcome the effects of such declines. Malawi experienced similar negative growth incidences with severe floods and droughts during 2015/16 (World Bank, 2016). The impacts of COVID-19 and subsequent flooding that Malawi continued to experience post-2019/20 household surveys are likely to negatively impact Malawi's quest to reduce poverty and income inequality. Therefore, based on such evidence, we see no justification for Malawi to experience declining poverty rates between all income distributions.

Another finding contrary to the nexus is increasing inequality inconsistent with declining poverty rates. Figure 1 supports our observations. The official NSO statistics showcase that inequality rose while poverty declined, especially from 2004-to 2010. There are several underlying factors that we found contributing to such misleading results. Firstly, we found that the number of Primary Sampling Units (PSU) used in computing the official poverty indices was inconsistent with the enumeration areas used in the survey. For instance, for the second Integrated Household Survey (IHS2), the PSUs used to compute poverty indices were 221 against 447 PSUs, IHS3 used 281 against 768, IHS4 used 291 against 791 PSUs, and the IHS5 used 328 against 717 PSUs. Secondly, the NSO reports a standard error for the fifth Integrated Household Survey for the 2019/20 period equal to zero, indicating a scenario with no variation between the sample and population poverty measurement.

Figure 1: Malawi Growth-Poverty-Inequality Nexus

Official National Statistics Office Statistics, 2004-2019



Source: World Development Indicators, 2019 & IHS, 2004/05, 2010/11, 2016/17, 2019/20

Against this backdrop, the study aims to recompute the poverty indices for Malawi from the second Integrated Household Survey of 2004/05 to the fifth Integrated Household Survey of 2019/20 to correct such anomalies. Our methodology in recomputing the consumption expenditure aggregates follows a new approach proposed by Charles P. Winsor that avoids truncating outliers by replacing extreme values with top values of a percentile in a distribution tail (Barnett and Lewis, 1994). The proposed approach is also superior to replacing extreme values with median values as the latter contributes toward overstating poverty in surveys with heavily leftward skewed data. In addition, we standardize conversion factors in this study to ensure that the four surveys are comparable.

We believe our study is the first of its kind. The study will use our newly computed consumption aggregates to answer the three following questions. Firstly, understanding the true poverty headcount and how poverty has changed in Malawi since 2004. Secondly, the level and extent of inequality in Malawi, particularly rural/urban and regional trends since 2004. Thirdly, the extent to which real consumption aggregates have grown and redistributed since 2004, focusing on the pro-poorness of growth in Malawi since 2004.

We structure the rest of the paper: Section 2 reviews existing literature on poverty measurement. Section 3 briefly outlines the methodology used and estimation techniques and their results. Section 4 discusses the survey results. Lastly, Section 5 concludes and provides policy recommendations.

2. Review of Theoretical and Empirical Literature

2.1 Theoretical Literature Review

2.1.1 Welfare Measurement Theory

The measurement of poverty involves multiple dimensions of deprivation and has relied on several summary measures of well-being for decades. The common well-being measurements include income and consumption aggregates, a form of monetary measure. Other non-monetary welfare measures related to income and consumption include freedom, life expectancy, education levels, and health status (Deaton and Zaidi, 2002; Ravallion, 2016). Using income aggregates as a welfare measure is common in developed countries where self-employment is sporadic. Conversely, consumption aggregates are measured mainly in developing economies like Malawi as a welfare measure, with many self-employed households mainly engaged in agriculture (Deaton and Zaidi, 2002).

The approach taken to guide the measurement of consumption aggregates is called an Agriculture Household Model (AHM). Bardhan and Udry (1999) define such households as having enterprises where they make production and consumption decisions separately. Over eighty percent of households in Malawi are mainly engaged in agriculture (World Bank, 2017; 2018). The AHM assumes that households operate in an environment with perfect or near-perfect markets. This allows the householder to maximize profit from production before maximizing their consumption utilities, subject to a budget constraint (Bardhan and Udry, 1999). The conventional AHM assumes that the householder's problem is to maximize consumption utility subject to a budget constraint:

$$\begin{gathered} Max\ U(\boldsymbol{c}_i, \boldsymbol{l}_i) \\ Subject\ to \\ p\boldsymbol{c_i} + w\boldsymbol{L_i^h} + r\boldsymbol{A_i^h} \leq \boldsymbol{F}(\boldsymbol{L}, \boldsymbol{A}) + w\boldsymbol{L_i^m} + r\boldsymbol{A_i^m} \end{gathered} \tag{1}$$

Equation (1) is called the canonical Agriculture Household model, where a household maximizes its utility function subject to a concave production function and income generated from supplying factors on the market. The budget constraint stipulates that the sum of cash expenditures on a vector of consumption goods, hired labor, and rented land cannot exceed cash revenues from enterprises, labor supplied on the market, and land rented out (Bardhan and Udry, 1999). The main prices faced by the household include a price for goods and services (p), wages (w), and ground rents (r). For this reason, consumption measures are an important element in any well-being assessment. Two critical concepts emphasize the importance of consumption.

2.1.2 Samuelson's (1974) Money Metric Utility

The first is the *Money Metric Utility* (*MMU*), a measure of well-being based on the money required for sustenance (Samuelson, 1974). In the MMU, a system of indifference curves represents consumer preferences, where higher indifference curves represent high consumer preference and vice versa. Given that each indifference curve also denotes some level of well-being, we can allocate each household based on consumer preference and its associated indifference curve (Deaton and Zaidi, 2002). In the MMU model, the budget constraint faced is a cost or expenditure function that the householder should minimize to reach its desired consumer preference or indifference curve. That is,

$$u_m^h = c(u^h, p^0) \approx p^0 . q^h \quad (2)$$

Equation (2) states that the MMU Function is the minimum cost of reaching a household's utility (u^h) at a given vector of prices (p^0). Therefore, the MMU is the sum of all consumer consumption bundles valued at base prices. We can also represent the MMU function using a Paasche Price Index (PPI), where equation (2) becomes

$$p_p^h = \frac{p^h . q^h}{p^0 . q^h} \quad (3a)$$

$$u_m^h \approx \frac{p^h . q^h}{p_p^h} = \frac{x^h}{p_p^h} \quad (3b)$$

2.1.3 Blackorby and Donaldson (1987) Welfare Ratios

According to Blackorby and Donaldson (1987), the challenge with the MMU comes when policymakers intend to use consumption aggregates to measure inequality in cases where household income distribution is of paramount importance. In such scenarios, the concavity of the utility function is crucial, which is not the case for the MMU utility function. Blackorby and Donaldson (1987) propose a new welfare measure that expresses the welfare measure relative to a baseline indifference curve; both expressed prices facing the household. They define such a base as a welfare level marking the boundary between poor and non-poor households. Therefore, the welfare ratio cost function is given by

$$wr_r^h = \frac{c(u^h, p^h)}{c(u^z, p^h)} \approx \frac{p^h . q^h}{p^z . q^h} \quad (4a)$$

They define the money measure of the welfare ratio function as the product between the welfare ratio, and the poverty utility are reference prices, represented by

$$u_{wr}^h = \frac{c(u^h,p^h)}{c(u^z,p^h)} \times (u^z, p^0) \approx \frac{p^h.q^h}{p^z.q^h} \times p^0.q^z \qquad (4b)$$

The Laspeyres index approximates the cost of the living price index for the welfare ratio, represented as

$$p_L^h = \frac{p^h.q^z}{p^0.q^z} \qquad (4c)$$

This means the Money measure of the welfare ratio is

$$u_{wr}^h \approx \frac{p^h.q^h}{p_L^h} = \frac{x^h}{p_L^h} \qquad (4d)$$

In the case of the monetary welfare measure, the Laspeyres index is proportional to the poverty indifference curve. Thus, changes in the total expenditure patterns do not affect the weights implying that the money-measure welfare ratio utility is proportional to the selected consumption bundle providing a direct link between redistributive policy and the measurement of its effects (Deaton and Zaidi, 2002).

Lastly, equation (1) advocates aggregating the value of time and leisure in computing consumption. However, valuing leisure has proved problematic to researchers and has introduced more problems than it solves (Deaton and Zaidi, 2002). Therefore, we do not include the computation of leisure when estimating consumption aggregates from survey data.

2.2 Empirical Literature Review

2.2.1 Measuring Consumption Aggregates and Weaknesses

Deaton and Zaidi (2002) provide guidelines for constructing nominal consumption aggregates from any integrated household survey. The standard computation of such consumption aggregates includes four main consumption aggregates classes. The first category consists of consumption aggregates related to food items that contain various foods consumed within a given reference period and from all possible sources. The second category consists of non-food items that the household uses daily, such as clothing, education, and health expenses. The third category consists of consumer durables accounting for the value of service derived from consuming a durable good. The last category estimates the value of service a household enjoys from consuming a dwelling, either rented or owned, and utilities such as water and electricity.

One of the important measurement problems they always emphasize to check is how to deal with 'gross' outliers. We can use several graphs to view them, but the most common are either 'oneway' or 'box' graphs. The biggest challenge of outliers is that they affect the sample's mean value and the dataset's range. Another significant challenge with outliers is the validity of the $t-$test by reducing the probability of Type I errors and substantially increasing the probability of Type II errors.

There are several options used to deal with outliers, and one of the most common ones used is the Inter Quartile Range (IQR), whose value is the difference between the 25^{th} and 75^{th} percentiles multiplied by a factor of 1.5. Analysts can use the IQR value to replace all extreme values beyond the IQR value with a median value since medians and modes are not affected by outliers (Deaton and Zaidi, 2002). There are several challenges we have encountered with this approach. Firstly, computed consumption aggregates from Agriculture Household Models are heavily skewed to the left and have a long-tail distribution. This means that replacing outliers at both tails will be impossible, given that the lower tail values will be negative, rendering the approach useless. Secondly, the IQR approach may overstate any computed statistic using median values. The approach replaces extreme values with a median value that is likely below the mean of a long-tailed distribution.

We propose an alternative approach to dealing with outliers proposed by Charles P. Winsor when dealing with long-tailed distributions (Tukey, 1962). The Winsor approach, also known as Winsorizing, is simply a procedure that replaces an outlier with an extreme value of a certain percentile in a given distribution. The approach is superior to truncating as there is no loss of observations, as it preserves the validity and improves the efficiency of the sample distribution (Dixon and Massey, 1957).

2.2.2 The Growth-Poverty-Inequality Trilemma

Understanding the growth-poverty-inequality trilemma is an important concept that should guide the computation of consumption aggregates to check whether they are meaningful. This requires generating trends to portray whether the fundamental axioms guiding the trilemma are true. Three important axioms guide the trilemma. Firstly, there is an inverse relationship between economic growth and poverty. Economic growth should exhibit poverty reduction properties: high economic growth reduces poverty, and low growth increases poverty. Second, a positive relationship exists between inequality and poverty growth: inequality growth should intensify poverty. Thirdly, there is an ambiguous relationship between economic and inequality growth: the relationship between growth and inequality is either positive or negative, depending on the adopted modeling approach (Adeleye *et al.*, 2020).

On the other hand, inequality growth intensifies poverty, and the higher the inequality, the greater the impact on the growth of poverty (Araar and Duclos, 2007). Therefore, the growth-poverty-inequality trilemma is an important philosophy in which inequality dampens any positive impact of economic growth on poverty reduction (Adeleye *et al.*, 2020).

3. Methodology and Computation Techniques

3.1 Data Sources and Cleaning

We used STATA 17 MP as the basic software to recompute the consumption aggregates in Malawi. The main data sources for our exercise are a series of Integrated Household Surveys (IHS) collected by the Malawi National Statistics Office since 2004. They comprise the Second IHS of 2004/05 (IHS2), the third IHS of 2010/11 (IHS3), the fourth IHS of 2016/17 (IHS4), and the fifth IHS of 2019/20 (IHS5).

We've made several improvements regarding the Deaton and Zaidi (2002) computational approach to creating consumption aggregates. First, we use the Winsorization approach to deal with outliers using percentiles as cut-off points. Second, based on the revealed preference axiom, we avoid computing imputed rentals on the housing module for two reasons: the surveys improved the collection of own rentals over time, reduced the number of missing values, and found out the imputed rentals marginally affected the computed statistics. Thirdly, we've included all primary sampling units in the computation of the consumption aggregates to ensure we generate statistics representing the entire population.

3.2 Methodology

The main objective of our study is to conduct a distributive analysis of the recomputed consumption aggregates. We will use the most popular statistics to analyze the growth-poverty-inequality nexus using a Distributive Analysis Statistical Package (DASP) developed by Araar and Duclos (2013). The following are the poverty metrics used in our study.

3.2.1 Poverty and Inequality Indices

We will use the Foster-Greer-Thorbecke (FGT) poverty index extensively, a per capita measure that looks at three dimensions of welfare shortfalls of the poor relative to a poverty line (Foster et al., 1984). The first metric is the poverty headcount ratio which defines the proportion of the population below the poverty line. The second metric used is the poverty gap index, measuring the average poverty intensity as a proportion of the poverty line. The third metric is the poverty severity index, measuring income inequality as an average of the squared poverty gap. The poverty severity index puts more weight on the poorest households further away from the poverty line (Araar and Duclos, 2013). The Gini index, developed by Corrado Gini in 1912, is an important income inequality measure, representing the income concentration of a population. The Gini coefficient is also important as the graphs it generates, the Lorenz curves, depict the spread of income inequality in a population.

3.2.2 Income Growth and Redistribution

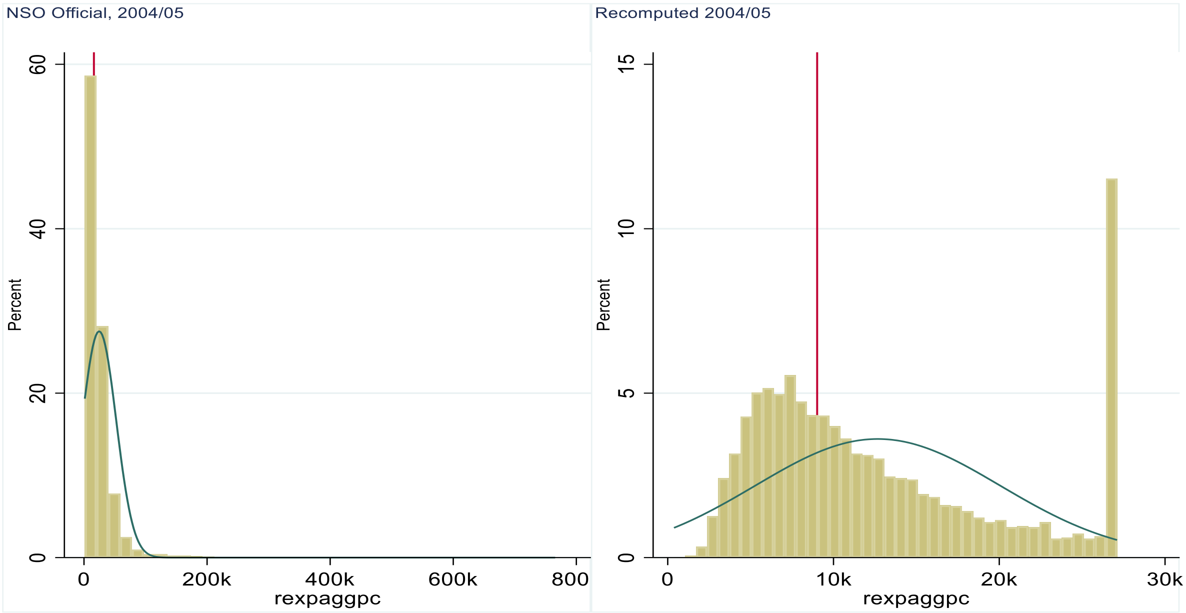
We will also concentrate on defining difference curves, that is, looking at the growth of income distributions. Ravallion and Chen (2003) define distributional income growth as the proportional change in income observed at various percentiles. They state that if the distributional changes are positive everywhere, social welfare increases for all first-order social welfare indices called "*first-order absolutely pro-poor*." This means the basis for poor benefits in absolute terms is on the observed positive distributive change. Income distribution can also be *first-order-relatively pro-poor* if income growth is higher than the mean income growth at each percentile (Araar *et al.*, 2010). This section focuses on absolute pro-poor growth.

4. Estimation Results

4.1 Descriptive Statistics – Histograms

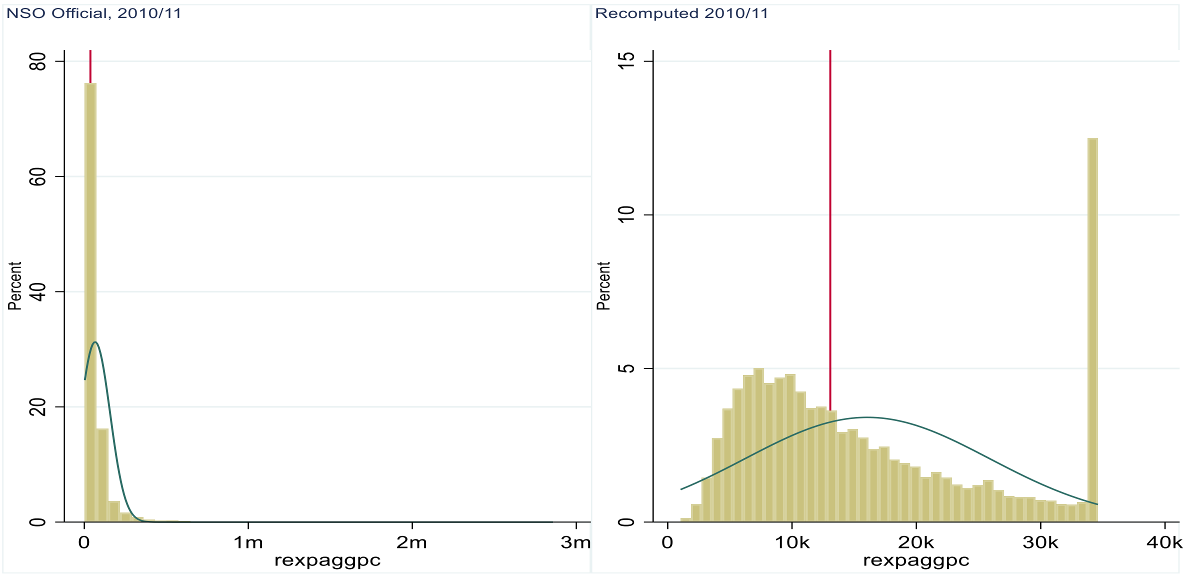
This section presents differences in histograms and normal distribution plots for the official versus our recomputed poverty and inequality measures. Figure 2a-d illustrates these graphs.

Figure 2a: Descriptive Statistics
Per capita Consumption Aggregate Histograms, 2004/05



Source: Second IHS, 2004/05

Figure 2b: Descriptive Statistics
Per capita Consumption Aggregate Histograms



Source: Third IHS, 2010/11

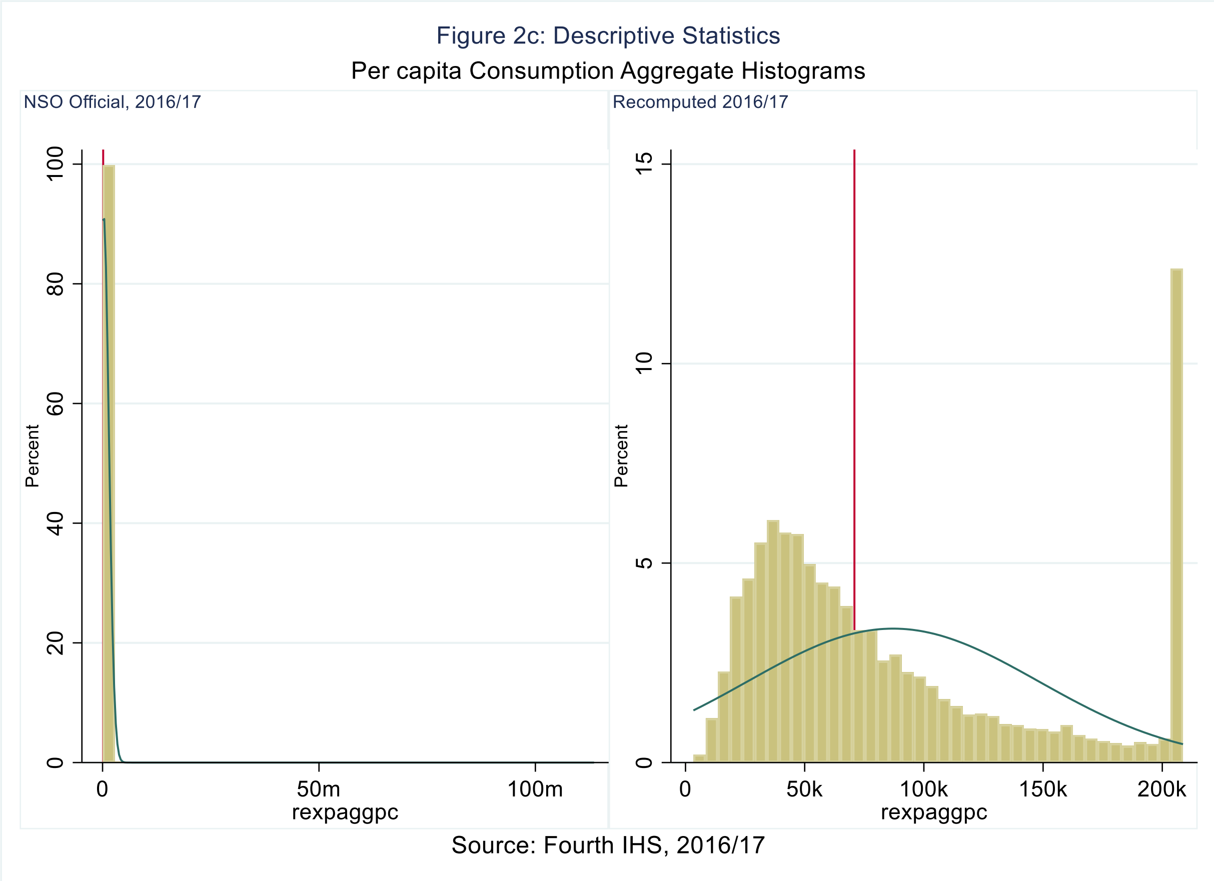


Figure 2c: Descriptive Statistics
Per capita Consumption Aggregate Histograms
NSO Official, 2016/17
Recomputed 2016/17
Percent
0
20
40
60
80
100
0
50m
100m
rexpaggpc
0
5
10
15
0
50k
100k
150k
200k
rexpaggpc
Source: Fourth IHS, 2016/17

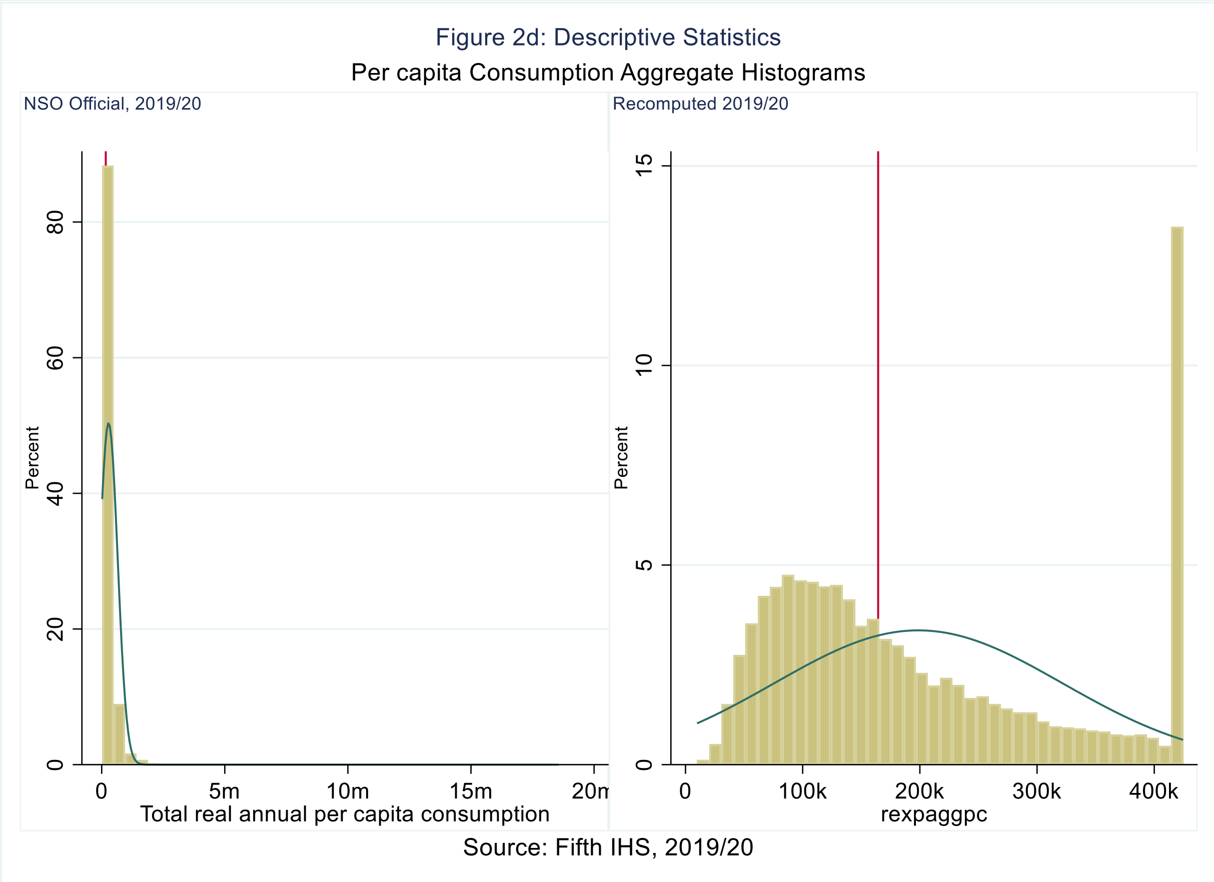


Figure 2d: Descriptive Statistics
Per capita Consumption Aggregate Histograms
NSO Official, 2019/20
Recomputed 2019/20
Percent
0
20
40
60
80
0
5m
10m
15m
Total real annual per capita consumption
0
5
10
15
0
100k
200k
300k
400k
rexpaggpc
Source: Fifth IHS, 2019/20

The left-hand graphs represent the official NSO consumption aggregates, while the right-hand side illustrates our recomputed values. Histograms are important as they visually display the data distribution and their frequencies. The normal distribution, which also plays a similar role, is vital in statistics. It helps us visualize the probability of observations in a data distribution that fall above or below a given value. In our case, we use the estimated poverty lines to see how spread the various income distribution have been since 2004/04. In Figure 2a-d, the red lines are the estimated poverty lines. For the official NSO statistics, the poverty lines are IHS2 2004/05 (MK16,165), IHS3 2010/11 (MK37,001.68), IHS4 2016/17 (MK137,428), and IHS5 2019/20 (MK165,879). The estimated poverty lines for our recomputed values are IHS2 2004/05 (MK9,014), IHS3 2010/11 (MK13,078), IHS4 2016/17 (MK70,872), and IHS5 2019/20 (MK164,422).

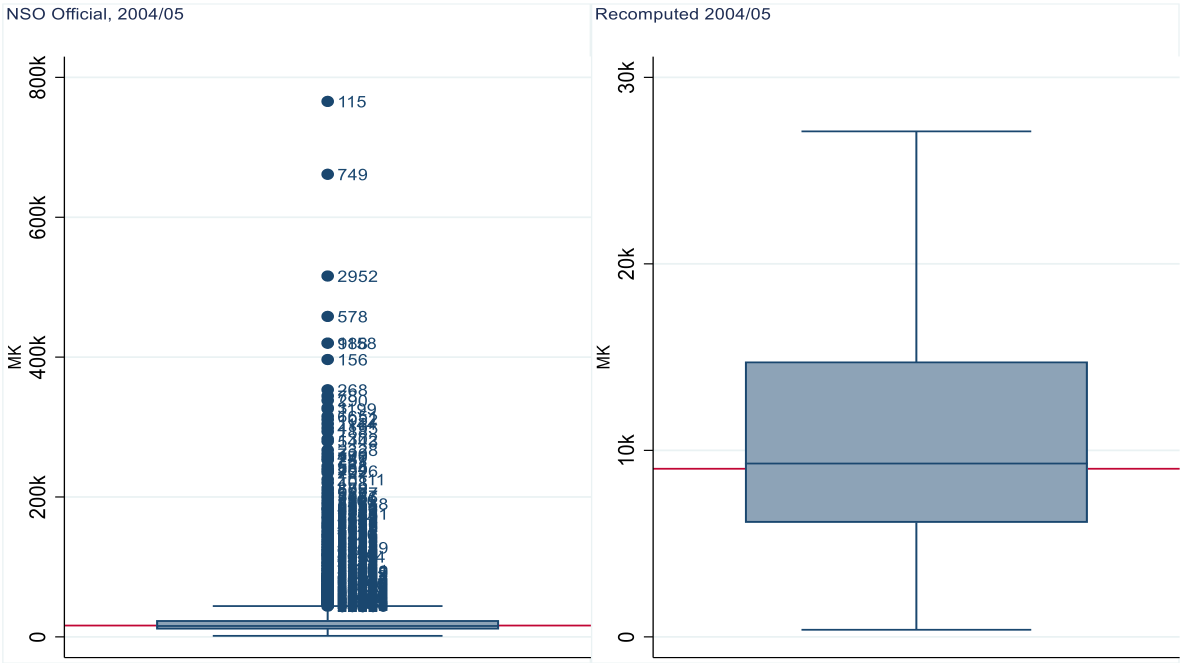
Based on the official NSO consumption aggregates, they all have a common distribution centered around the estimated poverty line. Therefore, the official poverty measures for these income distributions will always be around 50% and signify computational errors. On the other hand, our recomputed consumption aggregates adhere to the normal distribution patterns showcasing that most of the per-consumption aggregates are below the mean income of the estimated samples. Another important requirement to ensure that we are dealing with outliers in our samples is to apply the Winsorization principle that converts all outliers to the maximum value of the highest percentile used. We use the (0 95) cut-off points to Winsorize our per capita consumption aggregates. We explain the impact of such an approach in the next section.

4.2 Descriptive Statistics – Box Plots

This section presents box plots for the per capita consumption estimates, which helps us understand a variable's distributional characteristics. Figures 3a-d present such plots for both the recomputed and official NSO consumption aggregates. The left-hand graphs represent the official NSO consumption aggregates, while the right-hand side illustrates our recomputed values.

Figure 3a: Descriptive Statistics

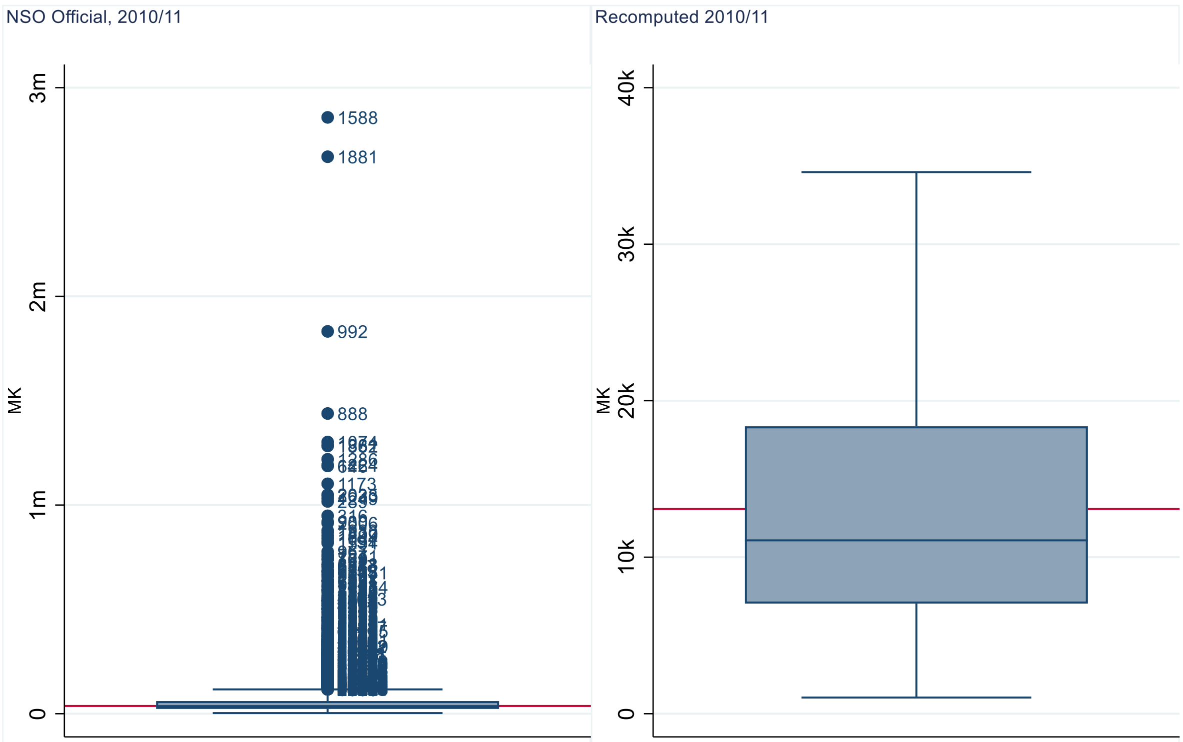
Per capita Consumption Aggregate Box Plots, 2004/05



Source: Second IHS, 2004/05

Figure 3b: Descriptive Statistics

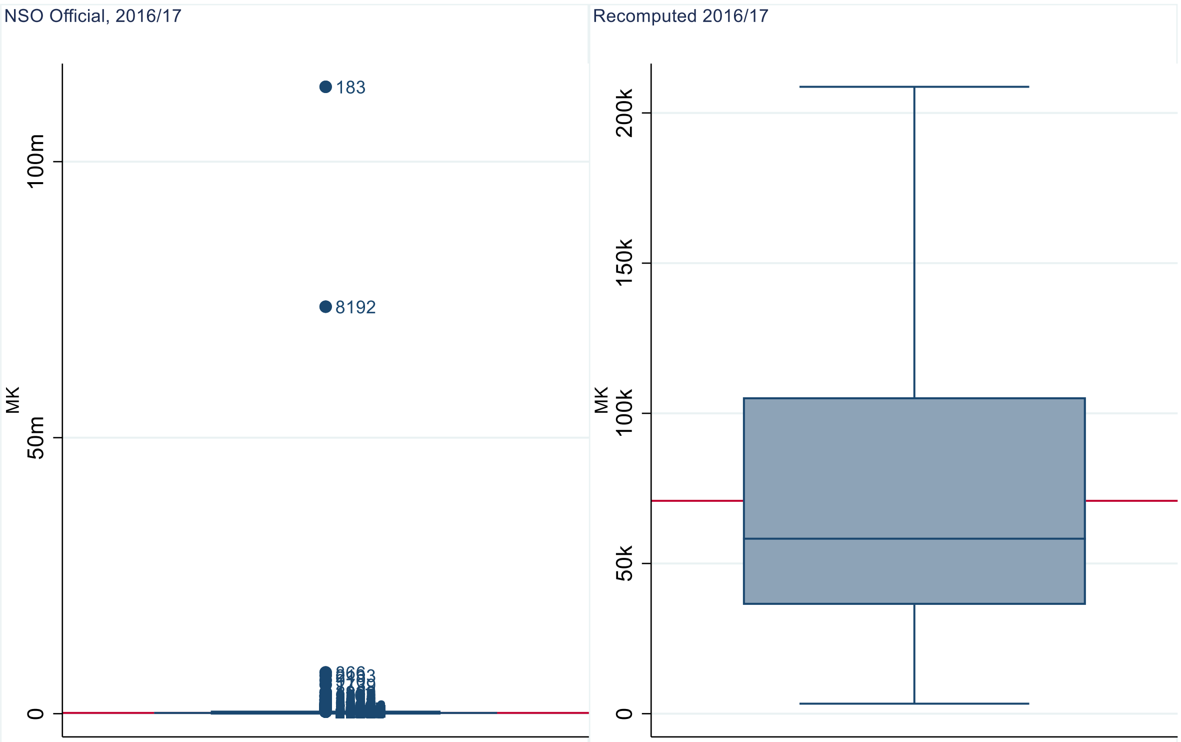
Per capita Consumption Aggregate Box Plots



Source: Third IHS, 2010/11

Figure 3c: Descriptive Statistics

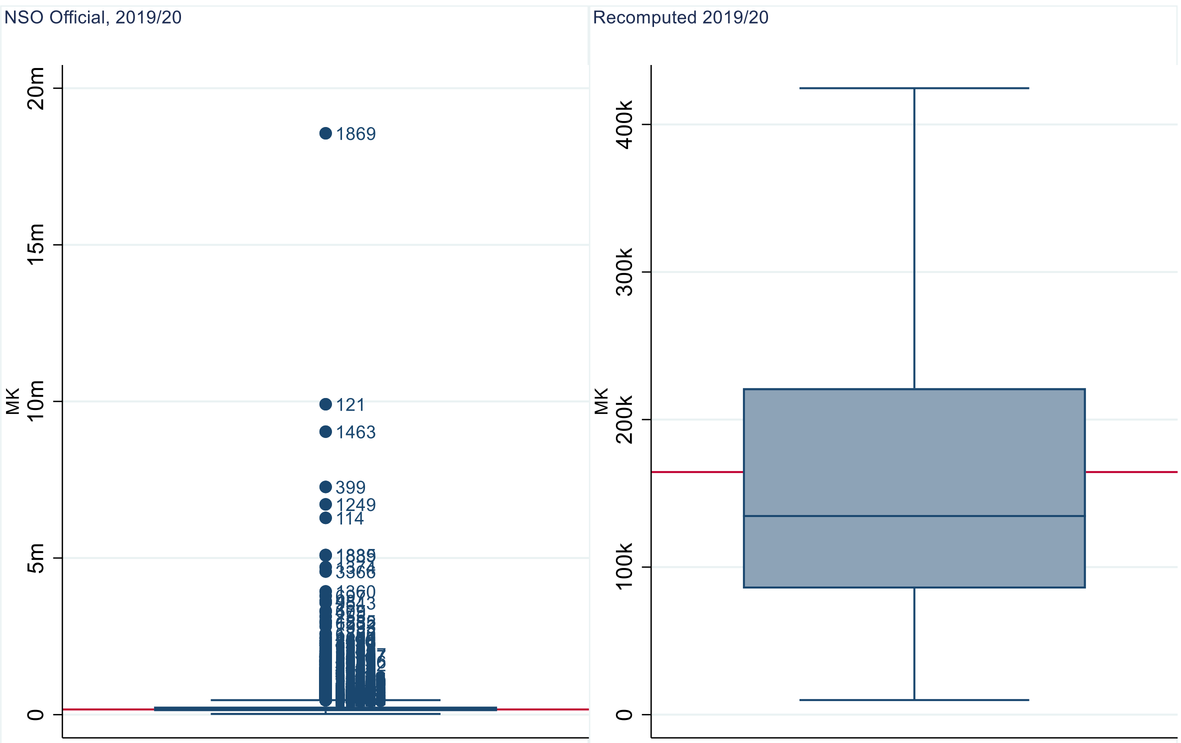
Per capita Consumption Aggregate Box Plots



Source: Fourth IHS, 2016/17

Figure 3d: Descriptive Statistics

Per capita Consumption Aggregate Box Plots



Source: Fifth IHS, 2019/20

We also include the estimated poverty lines on the y-axis in red font. The official NSO statistics consumption aggregates have significant outliers, as illustrated in the four Figures. Outliers lead to the following effects: firstly, outliers affect the mean, and since poverty and inequality estimates are means, they have compromised values. As shown in Figures 2a-d, the official NSO poverty line estimates are almost similar to the normal distribution mean value compared to our recomputed statistics on the right. Secondly, outliers generate large standard errors affecting the mean values range, eventually leading to overpredicting inequality measures.

4.3 How Poor is Malawi?

This section presents the results of our recomputed per capita consumption aggregates compared to the official NSO statistics. Table 1 presents poverty headcount and inequality estimates, including upper and lower bounds and the statistical significance of the poverty headcount and inequality differences between income distributions. We present the differences in a symmetrical matrix with lower values representing differences between income distributions from 2004 while the upper values denoting subsequent years.

We present the differences in a symmetric matrix. The first difference we note between the recomputed and official statistics is the differences in the estimated poverty lines. Based on our recomputed consumption aggregates, the results show that since 2004 national poverty in Malawi has increased significantly, evidenced by the statistically significant differences at the 1% and 5% significance levels. Compared to the 2004 income distribution, the national poverty headcount in 2019 by basis points increased by 0.52 in 2010, 2.35 in 2016, and 14.19, all statistically significant at the 1% and 5% significance levels, except for 2010, which was not statistically significant. Contrary to the official NSO estimates, the results show no significant change in national poverty between income distributions since 2004.

Table 1: Poverty and Inequality Estimates, 2004-2019

Poverty and Inequality Statistics levels						Difference			
	Estimate	Std. Error	Lower Bound	Upper Bound	Poverty Line	2004	2010	2016	2019
Poverty Headcount – National level									
Poverty Headcount									
2004	48.3	0.011	46.1	50.6	9,014		-13.66***	-11.84***	-14.19***
2010	59.1	0.009	57.3	60.9	13,078	13.66***		1.83	-0.52
2016	59.5	0.008	57.9	61.0	70,872	11.84***	-1.83		-2.35**
2019	61.5	0.008	59.9	63.1	164,422	14.19***	0.52	2.35**	
Poverty Headcount - Official NSO			-	-	-				
2004	52.4	0.013	49.7	55.0	16,165		1.73	0.85	1.64
2010	50.7	0.011	48.4	52.9	37,002	-1.73		-0.88	-0.09
2016	51.5	0.011	49.3	53.8	137,428	-0.85	0.88		0.79
2019	**50.7**	**0.000**	**50.7**	**50.7**	**165,879**	**-1.64**	**0.09**	**-0.79**	
Gini Coefficient – National level									
Gini Coefficient									
2004	33.1	0.003	32.5	33.8			-1.29***	-6.04***	-1.03***
2010	35.2	0.003	34.6	35.8		1.29***		-4.75***	0.26
2016	39.2	0.003	38.6	39.7		6.04***	4.75***		5.01***
2019	35.6	0.003	35.1	36.2		1.03***	-0.26	-5.01***	
Gini Coefficient - Official NSO									
2004	39.2	0.019	35.4	43.1			-3.53***	2.00	0.80
2010	45.2	0.013	42.5	47.8		3.53***		5.53***	4.33***
2016	42.3	0.044	33.5	51.0		-2.00	-5.53***		-1.20
2019	**37.9**	**0.000**	**37.9**	**37.9**		**-0.80**	**-4.33*** **	**1.20**	
Poverty Headcount – Rural									
Poverty Headcount									
2004	52.3	0.010	50.3	54.3	9,014		-16.3***	-16.8***	-17.2***
2010	65.6	0.009	63.7	67.4	13,078	16.3***		-0.53	-0.92
2016	68.4	0.008	66.8	70.0	70,872	16.8***	0.53		-0.39
2019	68.4	0.009	66.7	70.1	164,422	17.2***	0.92	0.39	
Poverty Headcount - Official NSO			-	-	-				
2004	56.2	0.015	53.3	59.1	16,165		-0.45	-3.27*	-0.40
2010	56.6	0.011	54.4	58.9	37,002	0.45		-2.82*	0.06
2016	59.5	0.010	57.4	61.5	137,428	3.27*	2.82*		2.88***
2019	**56.6**	**0.000**	**56.6**	**56.6**	**165,879**	**0.40**	**-0.06**	**-2.88*** **	
Poverty Headcount – Urban									
Poverty Headcount									
2004	17.3	0.023	12.8	21.7	9,014		-16.3***	-16.8***	-17.2***
2010	23.7	0.033	17.1	30.2	13,078	16.3***		-0.53	-0.92
2016	21.3	0.018	17.4	25.1	70,872	16.8***	0.53		-0.39
2019	24.0	0.018	20.5	27.6	164,422	17.2***	0.92	0.39	
Poverty Headcount - Official NSO			-	-	-				
2004	25.2	0.031	19.1	31.4	16165		7.53**	7.96*	6.05
2010	17.3	0.028	11.8	22.8	37002	-7.53**		-0.43	-1.91
2016	17.7	0.019	14.0	21.5	137428	-7.96*	0.43		-1.48
2019	**19.2**	**0.000**	**19.2**	**19.2**	**165879**	**-6.05***	**1.91**	**1.48**	

The results also show that Malawi is experiencing adequate income inequality, with Gini coefficient measures ranging from 0.33 in 2004/05 to 0.36 in 2019/20. The differences are statistically significant at the 1% significance level, except for the difference between 2010/11 and 2019/20 income distributions. In contrast, the official NSO statistics show a different pattern that assumes Malawi's inequality to represent big income gaps, with Gini coefficients peaking at 0.45 in 2010/11 before falling to 0.42 in 2016/17. Contrary to the official poverty headcount statistics in rural areas, poverty increased significantly from 52.3% in 2004/05 to 68.4% in 2019/20. This contrasts with the official statistics that show poverty not changing much since 2004/05 hovering around 56.6%.

4.4 Adherence to the Growth-Poverty-Inequality Trilemma

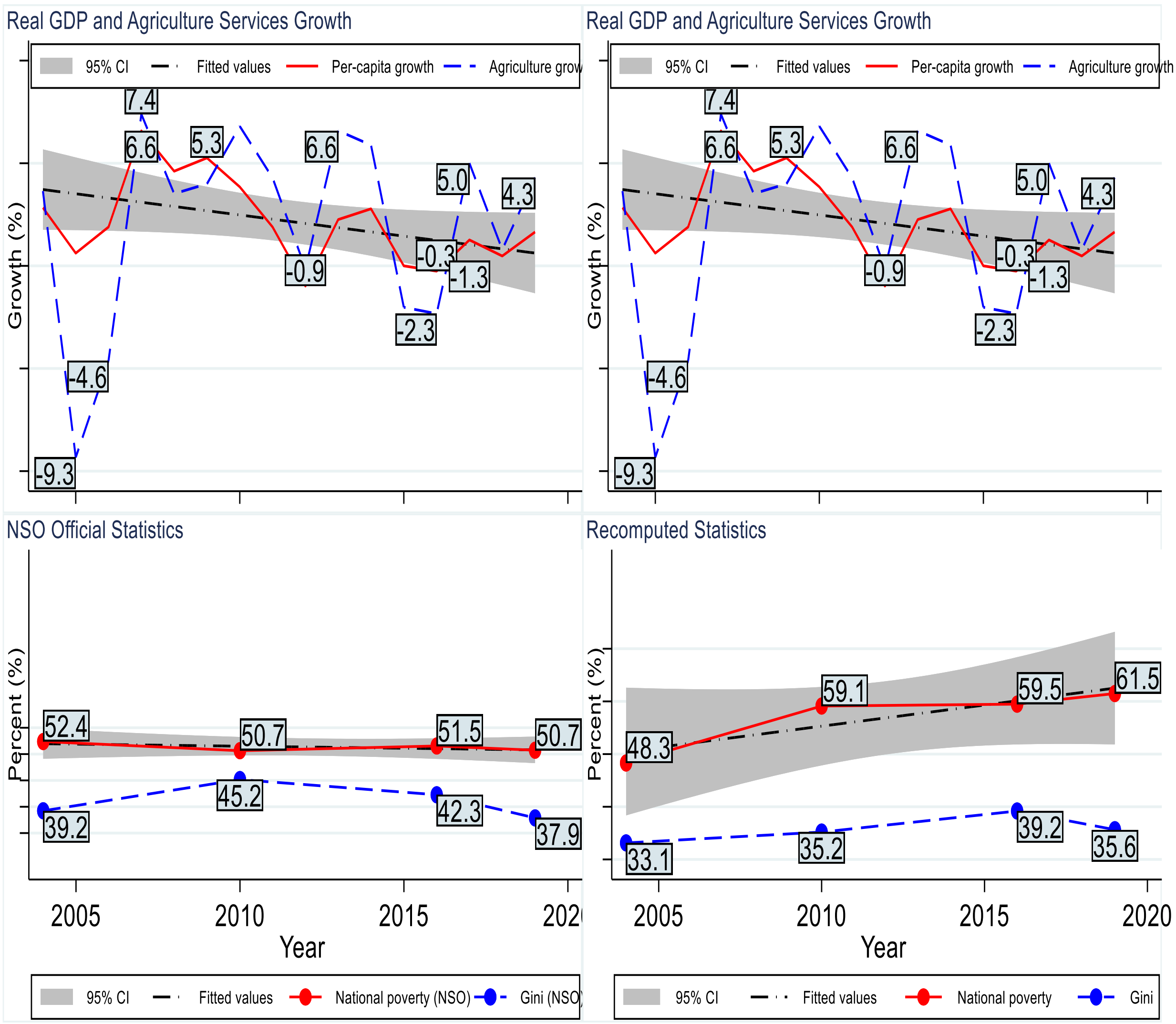
As we presented in the introduction, one weakness of the official NSO statistics is its inability to confirm the growth-poverty-inequality trilemma. Figure 4 plots our recomputed statistics on the lower right-hand quadrant against the official NSO poverty and inequality estimates on the left. As illustrated in Figure 4, our recomputed statistics adhere to the growth-poverty-inequality hypothesis. The graph's confidence intervals on the right-hand side of Figure 4 depict an inverse relationship between economic growth and poverty. The growth-inequality relationship is ambiguous: an inverse relationship for the first three income distributions and a positive relationship for the fourth income distribution.

While the official NSO statistics show that the poverty trend since 2004 declined, our statistics reveal that poverty increased from 48.3% in 2004/05 to 61.5% in 2019/20. On the other hand, inequality shows a similar trend, increasing from 0.33 in 2004/05 to 0.39 in 2016/17 and declining to 0.34 in 2019/20: all periods representing adequate income inequality. In contrast, the official NSO statistics reveal an adequate to big income inequality gap with a peak value reached in

2010/11 of 0.45, but overall declining to 0.38 in 2019/20. We attribute this result to the inability to deal with outliers in the various income distributions illustrated in Figures 2a-d and 3a-d.

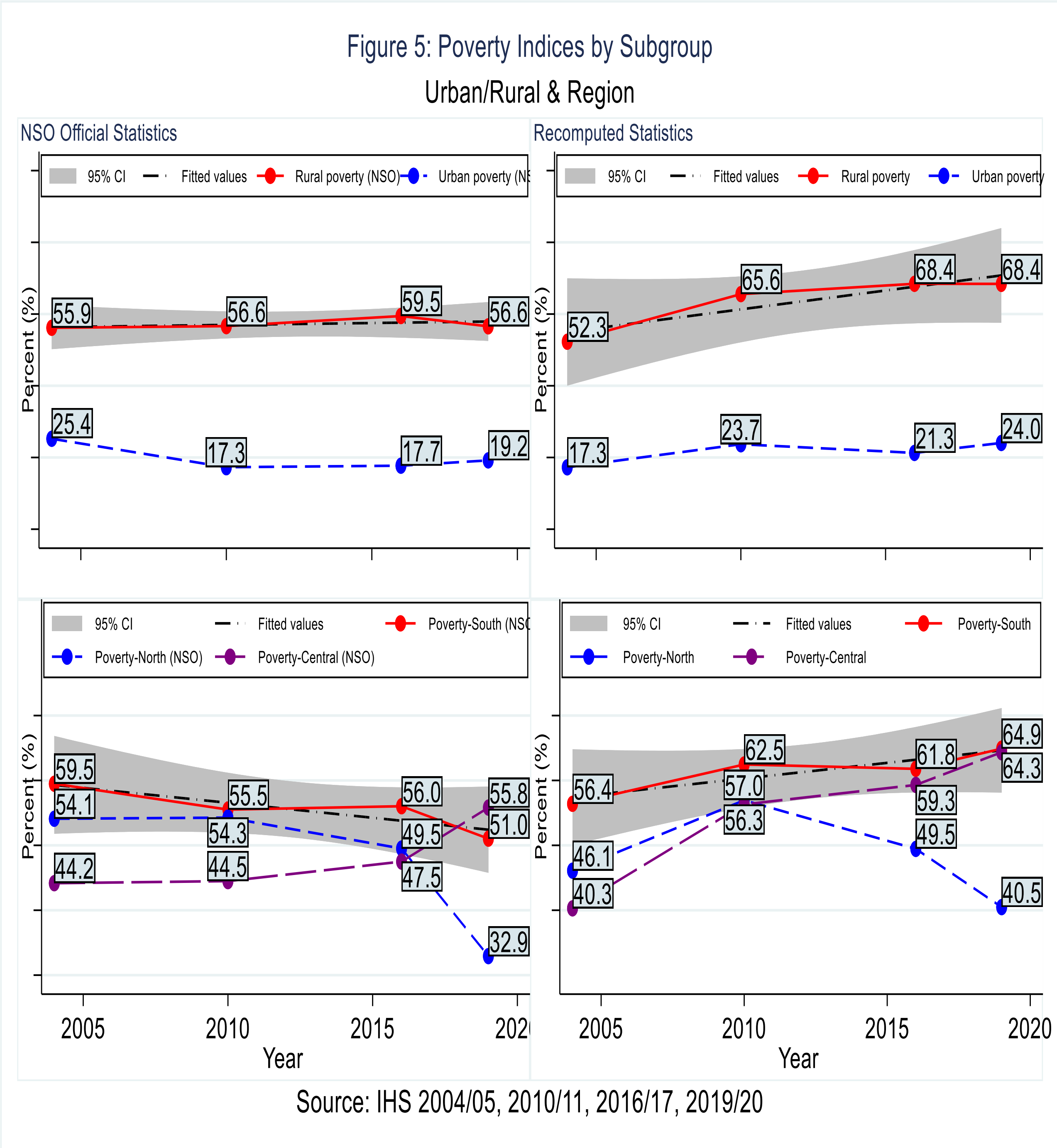
Figure 4: Poverty and Inequality in Malawi

Malawi Growth-Poverty-Inequality Nexus



Source: World Development Indicators, 2019 & IHS, 2004/05, 2010/11, 2016/17, 2019/20

Another important illustration is looking at poverty indices by subgroup, especially by location (urban and rural) and region (north, central, and south). Figure 5 illustrates these indices. The graphs on the left quadrants represent official statistics from the NSO and our recomputed poverty indices on the right.



As illustrated in Figure 5, the official NSO poverty statistics show no significant change in the growth of rural poverty in the three income distributions of 2004 (55.9%), 2010 (56.6%), and 2019 (56.6%), illustrating that poverty in Malawi is still a rural phenomenon. On the contrary, urban poverty significantly declined from 25.4% in 2004 to 17.3% in 2010, rising to 17.7% in 2016 and 19.2% in 2019.

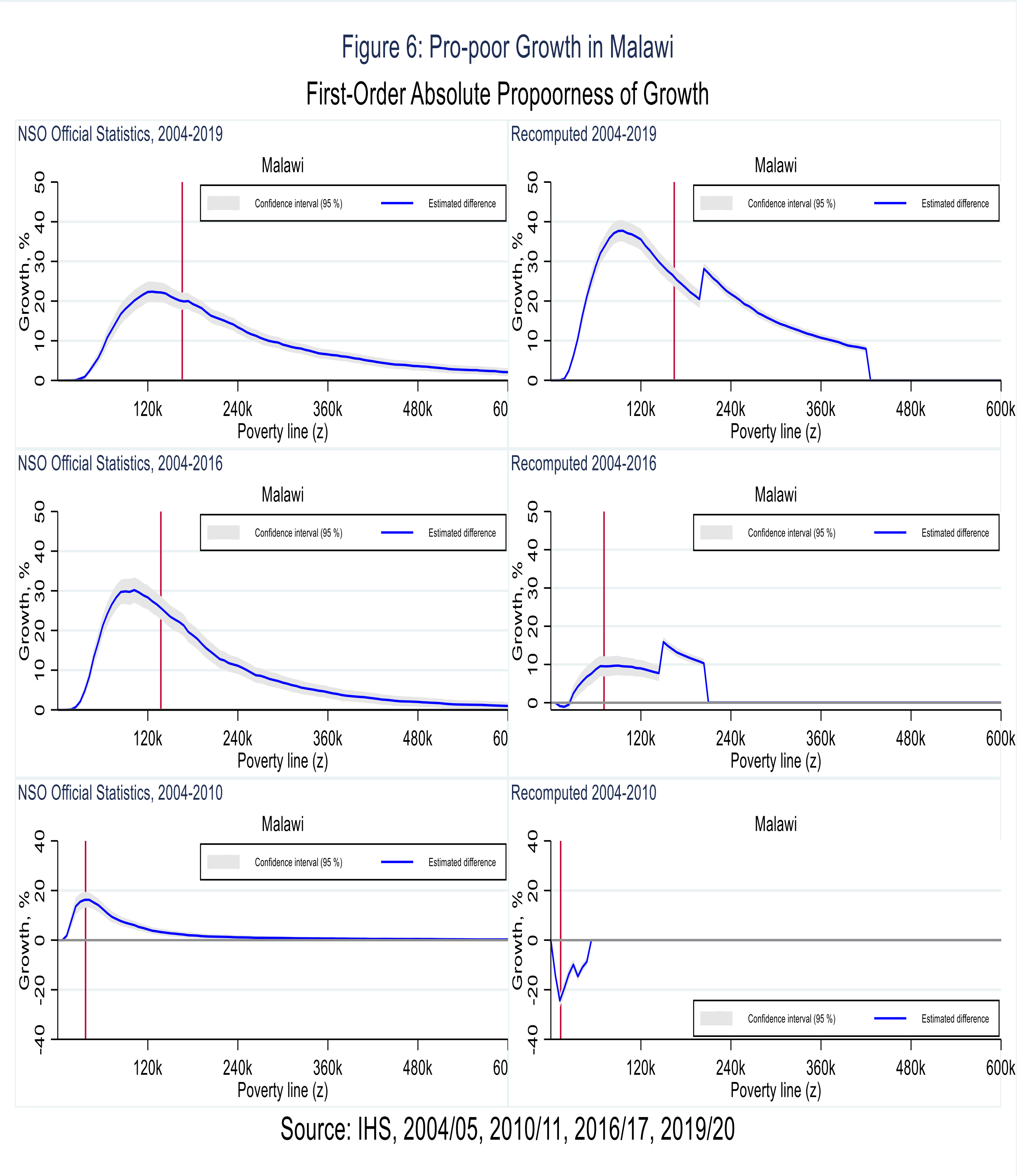
When we consider regional aspects, the official statistics reveal a significant reduction in poverty levels, particularly in the southern region, from 59.5% in 2004 to 51% in 2019. The northern region declined from 54.1% in 2004 to 32.9% in 2019. In contrast, the official statistics reveal that poverty increased in the central region from 44.2% in 2004 to 55.8% in 2019.

Based on our recomputed poverty statistics on the right-hand side quadrants, the results are contrary to the official NSO statistics. In the urban and rural settings, the results show that poverty in Malawi increased in both urban and rural areas, though varying degrees. Rural poverty significantly increased from 52.3% in 2004 to 68.4% in 2019, while urban poverty increased from 17.3% in 2004 to an average of 24.0% in 2019.

In the regional setting, the results reveal poverty increased in the southern region from 56.4% in 2004 to 64.9% in 2019, contrary to the official NSO poverty statistics. The results also reveal similar trends in the central region, where poverty increased from 40.3% in 2004 to 69.3% in 2019. Contrary to the two regions, the northern part of Malawi is the only region showing a significant decline in poverty levels rising from 46.1% in 2004 to 57.0% in 2010, falling to 40.5% in 2016, and reaching an all-time low, 41% in 2019.

4.5 Growth and Redistribution

The impact of computational errors is seen in Figure 6 when we plot first-order absolute pro-poorness of growth. The left-hand side quadrants represent income distribution growth curves based on NSO official statistics, and on the right-hand side, our recomputed statistics. It is important to deflate the consumption aggregates using rebased annual consumer price indices for each income distribution of the survey as a base year to remove the effect of inflation. As illustrated in Figure 6, the official statistics show that growth policies in Malawi were absolutely pro-poor as the results reveal a positive growth curve for all income distributions with significant growth experienced between the 2004 and 2016 surveys. Between the 2004 and 2010 income distributions, the maximum absolute pro-poor growth reached 20%, 40% for the 2004-2016 income distributions, and 20% for the 2004 and 2019 surveys.



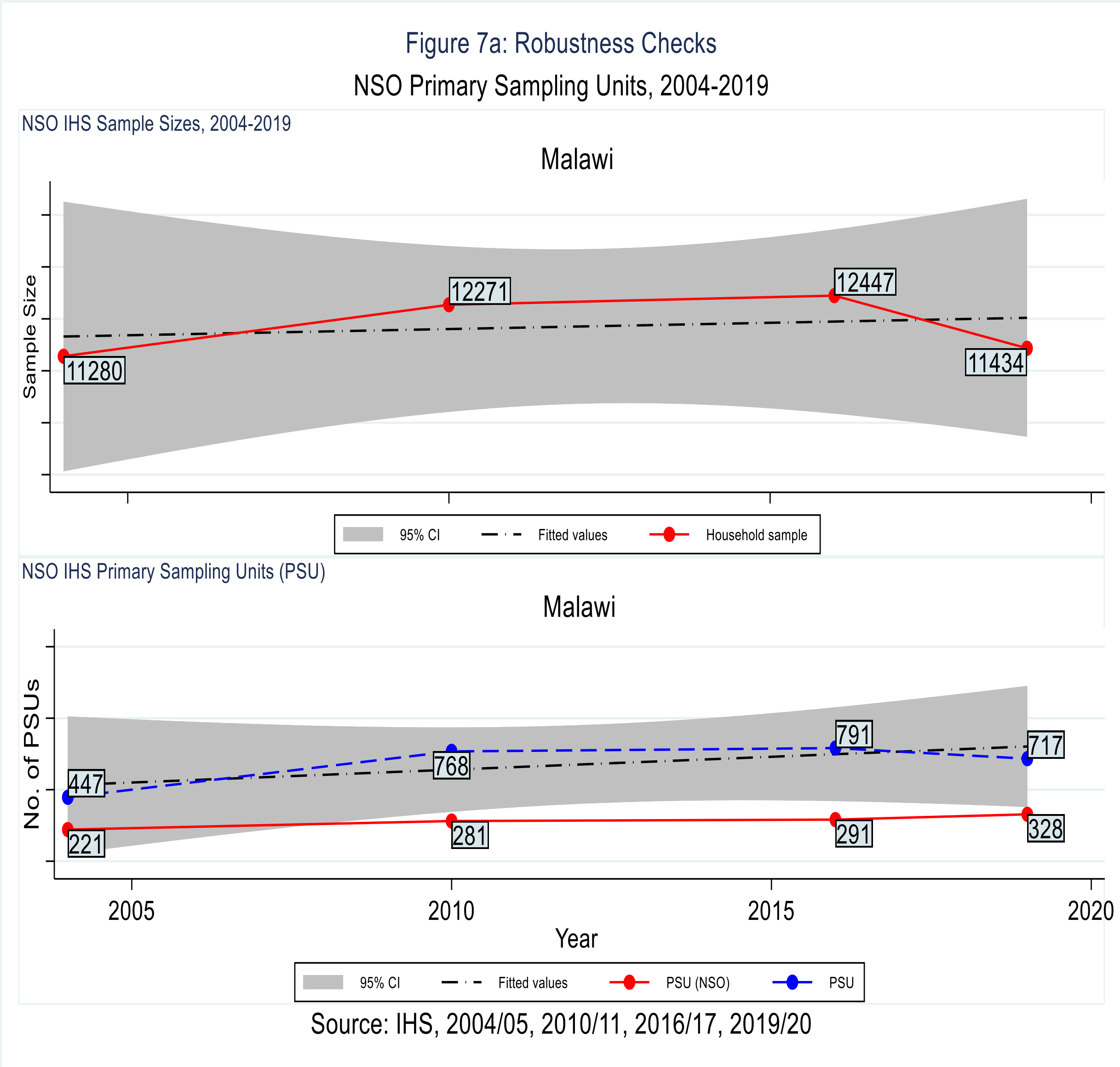
Conversely, our recomputed statistics reveal that growth was absolutely not pro-poor during the 2004-2010 income distributions with negative growth rates of approximately 20%. The results align with the observed negative real GDP and agriculture growth rates experienced during the same period. On the other hand, the 2004-2019 income distributions reveal absolutely pro-poor growth, with growth rates reaching approximately 40%. Conversely, much as 2004-2015 revealed some pro-poorness, growth was not absolutely pro-poor, with earlier estimations revealing negative pro-poor growth.

The results align with the real GDP and agriculture growth rates observed during the same period. In particular, the graphs depict a smooth transition from negative to positive pro-poor growth. Therefore, our recomputed statistics depict persistent poverty series commensurate with macroeconomic growth illustrated in Figure 4 over time.

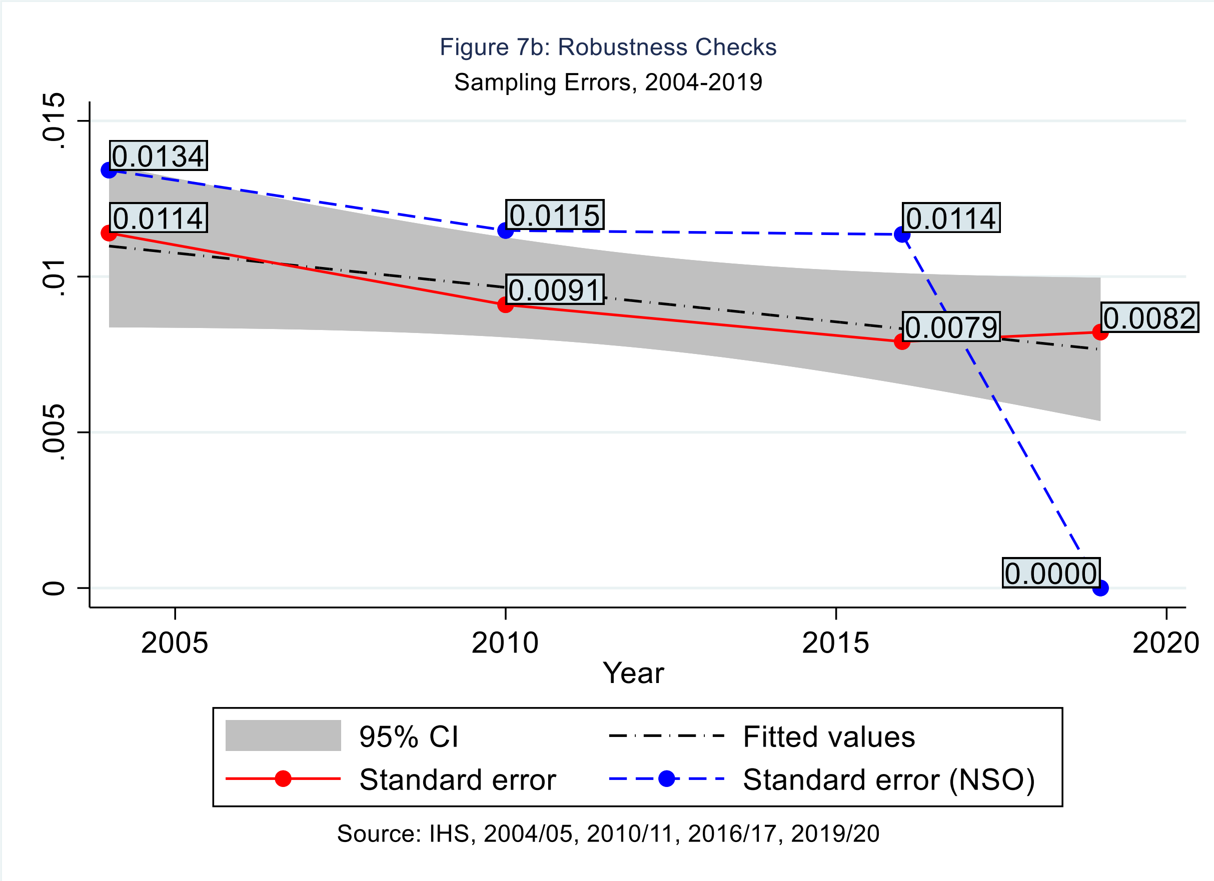
4.6 Robustness Checks

This section provides important statistics to ensure that the computed consumption aggregates are sound. Figure 7a presents distributive trends in the number of primary sampling units used by the official surveys. For sample statistics to represent the true population, it is important to ensure that we include all enumeration areas in estimating the consumption aggregates. This is not the case with the official NSO statistics that used far fewer primary sampling units in estimating consumption aggregates and still represented the total sample of each survey.

As observed in the graph, the PSUs for 2010/11, 2016/17, and 2019/20 income distributions used in estimating the official statistics were outside the confidence interval of the original Enumeration Areas. This has its consequences, as illustrated in this study. Our study used the actual number of primary sampling units to recompute consumption aggregates, as demonstrated on the right-hand side graphs of Figure 7a. This ensures that the computed statistics are representative of the entire population.



Lastly, Figure 7b presents distributive statistics related to sampling errors. A standard error is an important factor that estimates the efficiency, accuracy, and consistency and measures how precisely a sampling distribution represents a population. The lower the sampling error, the closer it is to measuring the true population value and vice-versa. In Figure 7b, our recomputed standard errors are lower than the official NSO values for all income distributions, except for the 2019/20 survey, where the NSO statistics revealed a zero standard error. We can attribute an obvious reason the official statistics have large standard errors to the low number of primary sampling units used to estimate the official poverty and inequality measures, as illustrated in Figure 7a.



This eventually leads to wider lower and upper bounds of the poverty or inequality measure. Everything we estimate using a sample is subject to random errors and should have positive standard errors (Haughton and Khandker, 2009). This should not be the case with the projected official standard error of zero for the 2019/20 income distribution. This implies that every computed poverty statistic has no lower or upper bound, as illustrated in Table 1. It implies a population and not a sample attribute.

5. Conclusion and Policy Implications

After observing unusual trends with macroeconomic growth variables, the study set out to recompute Malawi's poverty and inequality statistics. The study used raw data obtained from the World Bank microdata for Malawi. The data include STATA files for four integrated household surveys conducted in 2004/05, 2010/11, 2016/17, and 2019/20 to generate income distributions.

The income distributions are important in generating poverty and inequality estimates needed to make policy decisions.

The growth-poverty-inequality hypothesis is an essential guide to our study and states an inverse correlation between growth and poverty. Contrary to such a hypothesis, the National Statistics Office's (NSO) official poverty trends for income distributions since 2004 showed a persistent reduction in poverty in Malawi against a declining trend of economic growth. The official statistics assume a positive correlation between economic growth and poverty, rejecting the growth-poverty hypothesis. In addition, Malawi's official inequality statistics trends reveal a big income gap that peaked at 45.2% in 2010/11 and 42.3% in 2016/17 but still reveals an ambiguous relationship with economic growth.

We noted several computational errors that may have led to inconsistent poverty and inequality metrics. Firstly, the 2019/20 income distribution has a zero standard error, implying that all computed sample statistics are equal to population statistics. Second, we noted that all income distributions had significant outliers, a likely factor that affects the computation of both poverty and inequality measures, as noted in the official statistics. Third, the official NSO statistics used fewer primary sampling units to estimate poverty and inequality for all income distributions. Fourth, based on the Deaton and Zaidi (2002) guidelines, we find the median approach to replacing outliers problematic. It leads to errors in replacing those above the poverty line as poor if the median value is below the poverty line. This is likely the case in income distributions that have long tails. Our approach includes using the Winsorization approach to deal with outliers to overcome such a challenge. Lastly, we found the poverty line approximately equal to the normal distribution mean for the official per capita consumption aggregates rendering the poverty measure equivalent to 50% as per the official NSO results.

Our approach addressed these observed challenges to generate new poverty and inequality measures. Contrary to the NSO official statistics, we found that poverty in Malawi has increased since 2004, from 48.3% in 2004 to 61.5% in 2019. The recomputed statistics also reveal that the growth-poverty-inequality hypothesis holds in Malawi, as we found an inverse relationship. Similarly, the growth-inequality relationship shows ambiguous results as expected, and Malawi reveals adequate income inequality between 0.30 and 0.40.

The main policy implication drawn from this study is that data accuracy is important for decision-making. The more accurate it is, the higher the policymakers' confidence level to make sound decisions. It is, therefore, a crucial foundation that ought to have the foremost priority during data cleaning to avoid making policy decisions on erroneous poverty and inequality metrics. Second, the importance of data accuracy is seen in situations when further research is needed involving estimating complex regressions. Erroneous data used to define dependent and explanatory variables may lead to questionable results. Therefore, the study recommends employing sound data analytical tools proposed in this study to ensure that survey data is clean of any errors.

References

Adeleye B. N., O. Gershon, A. Ogundipe, O. Owolabi, I. Ogunrinola, and O. Adediran. 2020. "Comparative investigation of the growth-poverty-inequality trilemma in Sub-Saharan Africa and Latin American and Caribbean Countries." *Heliyon*, 6(12), 1-11. DOI: https://doi.org/10.1016/j.heliyon.2020.e05631

Araar, A., and J. Duclos. 2007. "Poverty and Inequality Components: A Micro Framework." *CIRPEE Working Paper No. 07-35*, 1-43. DOI: http://dx.doi.org/10.2139/ssrn.1023643

Araar, A., S. Bibi, and J. Duclos. 2010. Pro-Poor Growth. *Workshop on Poverty and Social Impact Analysis*, Dakar, Senegal, 8-12 June. Available at https://fdocuments.in/download/abdelkrim-araar-sami-bibi-and-jean-yves-2010-6-28-abdelkrim-araar-sami-bibi

Araar, A., and J. Duclos. 2013. *DASP: Distributive Analysis Statistical Package*. PEP, CIRPÉE, and World Bank, pp. 1-150. Available at http://dasp.ecn.ulaval.ca/modules/DASP_V2.3/DASP_MANUAL_V2.3.pdf

Bardhan, P., and C. Udry. 1999. *Development Microeconomics*. Oxford University Press

Barnett, V., and T. Lewis. 1994. *Outliers in Statistical Data*. 3rd Edition. Wiley

Bassanini, A., S. Scarpetta and P. Hemmings. 2001. "Economic Growth: The role of policies and institutions. Panel data evidence from OECD countries." OECD *Economics Department Working Papers*, No. 283, OECD Publishing, Paris, DOI: https://doi.org/10.1787/722675213381

Blackorby, C., and D. Donaldson. 1987. "Welfare Ratios and the Distributionally Sensitive Cost-Benefit Analysis." *Journal of Public Economics*, 34(3), 265-290. DOI: https://doi.org/10.1016/0047-2727(87)90001-6

Ravallion, M., and S. Chen. 2003. "Measuring Pro-Poor Growth." *Economic Letters*, 78(1), 93-99. DOI: https://doi.org/10.1016/S0165-1765(02)00205-7

Ravallion, M. 2016. The Economics of Poverty: History, measurement, and policy. The World Bank Group, Oxford University Press.

Chirwa, T. G., and N. M. Odhiambo. 2019. "The Nexus between Key Macroeconomic Determinants and Economic Growth in Zambia: A Dynamic Multivariate Granger Causality Linkage." *Empirical Economics*, Vol. 57(1), July, pp. 301-327. DOI: https://doi.org/10.1007/s00181-018-1439-2

Chirwa, T. G., and N. M. Odhiambo. 2020. "Determinants of gold price movements: an empirical investigation in the presence of multiple structural breaks." *Resources Policy*, Vol. 69, December, pp. 1-13. DOI: https://doi.org/10.1016/j.resourpol.2020.101818

Deaton, A., and S. Zaidi. 2002. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis." *Living Standards Measurement Study Working Paper No. 135* Available at https://openknowledge.worldbank.org/handle/10986/14101

Dixon, W.J. and F.J. Massey Jr. 1957. *Introduction to Statistical Analysis*. McGraw-Hill, New York, pp. 191- 227

Foster, J., J. Greer, and E.Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica*, 52(3), 761-766. DOI: https://doi.org/10.2307/1913475

Haugton, J. and S.R. Khandker. 2009. Handbook on Poverty and Inequality. The World Bank Group, Washington DC. DOI: https://doi.org/10.1596/978-0-8213-7613-3

National Statistics Office. 2005. *Second Integrated Household Survey 2004-2005*. STATA Database Available at https://microdata.worldbank.org/index.php/catalog/2307/get-microdata Malawi

National Statistics Office. 2011. *Third Integrated Household Survey 2010-2011*. STATA Database Available at https://microdata.worldbank.org/index.php/catalog/1003/get-microdata Malawi

National Statistics Office. 2017. *Fourth Integrated Household Survey 2016-2017*. STATA Database Available at https://microdata.worldbank.org/index.php/catalog/2936/get-microdata Malawi

National Statistics Office. 2020. *Fifth Integrated Household Survey 2019-2020*. STATA Database Available at https://microdata.worldbank.org/index.php/catalog/3818/get-microdata Malawi

Samuelson, P. 1974. “Complementarity: An Essay on the 40^{th} Anniversary of the Hicks-Allen Revolution in Demand Theory.” Journal of Economic Literature, 12(4), 1255-1289. DOI: https://www.jstor.org/stable/2722379

Tukey, J. W. 1962. “The Future of Data Analysis.” *The Annals of Mathematical Statistics*, 33(1), 1-67. DOI: https://www.jstor.org/stable/2237638

World Bank. 2016. Malawi Drought 2015-16: Post-disaster needs assessment (PDNA). The World Bank Group.

World Bank. (2017). *Malawi economic monitor 2017: Harnessing the urban economy*. The World Bank. https://openknowledge.worldbank.org/handle/10986/26763?locale-attribute=en

World Bank. (2018). *Malawi systematic country diagnostic: Breaking the cycle of low growth and slow poverty reduction*. World Bank Group, https://openknowledge.worldbank.org/handle/10986/31131

World Bank. 2021. *World Development Indicators 2021*. Retrieved from www.worldbank.org on August 8, 2021.