

POLICY RESEARCH WORKING PAPER

7758

All That Glitters Is Not Gold

Polarization Amid Poverty Reduction in Ghana

Fabio Clementi

Vasco Molini

Francesco Schettino



WORLD BANK GROUP

Poverty and Equity Global Practice Group

July 2016

Abstract

Ghana is an exceptional case in the Sub-Saharan Africa landscape. Together with a handful of other countries, Ghana offers the opportunity to analyze the distributional changes in the past two decades, since four comparable household surveys are available. In addition, different from many other countries in the continent, Ghana's rapid growth translated into fast poverty reduction. A closer look at the distributional changes that occurred in the same period, however, suggests less optimism. The present paper develops an innovative methodology to analyze the distributional changes that occurred and their drivers,

with a high degree of accuracy and granularity. Looking at the results from 1991 to 2012, the paper documents how the distributional changes hollowed out the middle of the Ghanaian household consumption distribution and increased the concentration of households around the highest and lowest deciles; there was a clear surge in polarization indeed. When looking at the drivers of polarization, household characteristics, educational attainment, and access to basic infrastructure all tended to increase over time the size of the upper and lower tails of the consumption distribution and, as a consequence, the degree of polarization.

This paper is a product of the Poverty and Equity Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at vmolini@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

All That Glitters Is Not Gold: Polarization Amid Poverty Reduction in Ghana

Fabio Clementi¹, Vasco Molini^{2,†}, and Francesco Schettino³

¹ *University of Macerata, Macerata, Italy*

² *World Bank, Washington DC, USA*

³ *Second University of Naples, Naples, Italy*

Keywords: Ghana; consumption expenditure; poverty and inequality; polarization; relative distribution; decomposition analysis.

JEL classification: C14; D31; D63; I32.

[†] Corresponding author. E-mail addresses: fabio.clementi@unimc.it (Fabio Clementi), vmolini@worldbank.org (Vasco Molini), francesco.schettino@unina2.it (Francesco Schettino).

Note: The authors acknowledge financial support from the World Bank. We thank Dan Pavelesku (World Bank) for excellent assistance with data preparation. We also thank Christoph Lakner, Pierella Paci and Andrew Dabalen (World Bank) for comments on an earlier version of the manuscript. Of course, we are the sole responsible for all possible errors the paper may contain.

1 Introduction

Over the last two decades, several African countries experienced stable and sustained growth that did not translate, nonetheless, into rapid poverty reduction. Compared to other regions, particularly in the last decade, the growth elasticity of poverty of Sub-Saharan African (SSA) countries has been lower than in the rest of developing world (Molini and Paci, 2015). The causes of this limited inclusiveness are numerous, but typically scholars point to the excessive reliance of many SSA economies on a limited basket of raw materials and the limited trickle down of this growth to households' consumption.

At first glance, Ghana is an exception compared to the rest of Sub-Saharan countries. Ghana's rapid growth did translate into fast poverty reduction. Inequality increased over the last two decades (Cooke et al., 2016; Aryeetey and Baah-Boateng, 2015) but, compared to other SSA countries, Ghana still fares relatively well. When ranking SSA countries according to the average Gini index over the last 20 years Ghana occupies the bottom 20 percent of the Gini distribution and despite some deterioration, in 2012, it was still below the median and among the lowest of rapidly growing African economies (Molini and Paci, 2015).

A closer look at the distributional changes occurred in the same period, suggests, however, less optimism. Like many other developing countries (Clementi et al., 2014, 2015; Clementi and Schettino 2015) Ghana is experiencing a fast increase in polarization. Whereas inequality relates to the overall dispersion of the distribution and provides clues to whether a society's prosperity has been shared broadly or not, polarization is concerned with the division of society into subgroups. In the context of income distribution, this concept is typically equated with the "hollowing out of the middle", a situation where the society has a sizeable group of poor persons and there is also a non-negligibly sized group of persons with very high income and, in contrast, the size of the group occupying the center of the income distribution is rather low (see, for instance, Foster and Wolfson, 1992, and Wolfson, 1994, 1997). Within each group there is increasing "identification", which means income homogeneity and often declining income inequality, while between the two groups we have instead increasing "alienation" (Duclos et al., 2004). The combined effect of the forces of alienation and identification between two significantly sized groups would tend to lead to effective opposition, a situation that might give rise to social conflicts and tensions (Esteban and Ray, 1999, 2008, 2011).

The contribution of this paper is twofold. First, it uses a very intuitive yet little explored method, the relative distribution, (Handcock and Morris, 1998, 1999) to analyze the recent distributional changes occurred in the country. The strength of this method consists in providing a non-parametric framework for taking into account all the distributional differences that could arise in the comparison of distributions over time and space. In this way, it enables to summarize multiple features of the

expenditure distribution that would not be detected easily from a comparison of standard measures of inequality and polarization. Second and most important, the paper develops within the relative distribution framework a novel methodology to identify the drivers of distributional changes and quantify their impact on the welfare distribution; the main value added being it enables a very granular analysis of the distributional changes that an analysis based on standard inequality decompositions would not allow.

The paper is organized as follows. Section 2 discusses the data and presents the methodology. Section 3 provides the results. Section 4 concludes.

2 Data and methodology

2.1 The Ghanaian household survey data

The data used in this paper come from the Ghana Living Standard Survey (GLSS), a nation-wide survey conducted by the government-run Ghana Statistical Service that provides information for assessing the living conditions of Ghanaian households.

The GLSS has emerged as one of the most important tools for the welfare monitoring system in Ghana. It provides detailed information on approximately 200 variables, including several socio-economic and demographic characteristics, and information on household consumption of purchased and home-produced goods as well as asset ownership. Each of the waves is organized into 4 modules, which are stored in the individual, the labor force, the household and the household expenditure files, for which survey questionnaires are readily available.

The Ghana Statistical Service has conducted six rounds of the GLSS since 1987, thereby providing over 20 years of comparable data. The second, third, fourth and fifth rounds were carried out, respectively, in 1988, 1991/92, 1998/99 and 2005/06. Recently, data for the sixth round of GLSS have also become available, so that the proposed case study paper will be one of the first studies using this data set. However, only the last four rounds, from 1991/92 (GLSS-3) to 2012/12 (GLSS-6), have been based on the same questionnaire and are therefore fully comparable.

The availability of comparable and extensive information represents a success on its own. Ghana is one of the few countries in Africa that has produced comparable, high-quality household data covering over two decades. This is an important achievement because the availability of such rich and comparable information beginning in 1991, as well as the quality improvements of the surveys over the years and the fact that they collect data on both the monetary and the non-monetary dimensions of welfare, permit the establishment of an accurate picture of inequality and polarization over time, including the drivers behind these phenomena.

As a measure of well-being we will use household consumption for 1991/92 (GLSS-3), 1998/99 (GLSS-4), 2005/06 (GLSS-5) and 2012/13 (GLSS-6). In that, we depart from the literature using income as a measure of well-being. Our choice is mainly motivated by measurement issues, which play a very relevant role in the case of Ghana and have to do with the quality of the income measures that one can obtain from the GLSS data. In economies where agriculture is an important and established sector, consumption has indeed proven preferable to income because the latter is more volatile and more highly affected by the harvest seasons, so that relying on income as an indicator of welfare might under- or over-estimate living standards significantly (see, for instance, Deaton and Zaidi, 2002, and Haughton and Khandker, 2009). On the theoretical ground, as consumption gives utility to individuals, the analysis of its distribution should be the most natural approach to study well-being. Income matters insofar as it gives access to consumption, which is the ultimate source of individual welfare. Consumption is a better measure of long-term welfare also because households can borrow, draw down on savings, or receive public and private transfers to smooth short-run fluctuations. The GLSS collects sufficiently detailed information to facilitate estimates of the total consumption of each household. It relies on consumption per adult equivalent¹ to capture differences in need by age and economies of scale in consumption. Scales of consumption by age and sex are computed by the Ghana Statistical Service.

The GLSS is based on a two-stage (non-stratified) sample design. Therefore, when the data are analyzed, sampling weights are used to account for the survey design. Besides, to enhance comparability of consumption data over the four waves, all expenditures have been deflated across both space² and time and expressed in 2005 constant prices – as well as converted, when necessary, from Ghanaian second cedi (GHC) to Ghanaian third cedi (GHS), i.e. for GLSS-3 to GLSS-5.

A summary of distributional statistics obtained from the GLSS data sets is given in Table 1. Besides the growth of the real mean and median consumption expenditures, the most notable feature is the picture that emerges across different indicators of inequality. The consumption shares of the poorest percentiles of the population decreased between approximately 0.9 and 1.4 % a year in the period examined, in contrast to what is observed for the richest percentiles, whose shares experienced average yearly increases of around 0.2 %. Inequality in household consumption was initially constant, but widened considerably between 1998/99 and 2005/2006 – a jump of about 7 % in the Gini's coefficient and 20 % in the Theil's index.³ Inequality has remained constant at the higher level after 2005/06, but

¹ We use adult equivalent scales because also the official consumption, poverty and inequality figures are expressed in adult equivalent terms.

² The price deflator differs across the ten regions in which Ghana is divided and within each region by urban and rural areas.

³ Running a simple *t*-test of the difference between Gini and Theil indices from the 1998/99 and 2005/06 samples yields a *p*-value of around zero, which confirms the finding that points to increasing inequality over the 1998-2005 period at any of the usual significance levels.

the trends in the shares of consumption of the bottom and top quintiles have continued in the same direction.

However, the narrative about inequality is more nuanced than the summary measures suggest. The summary measures of inequality analyzed above only partially capture the changes at various points of the consumption distribution. The results of a simple inter-quantile analysis can provide more detailed information on the changes occurring at all points of the distribution (see Table 2). They show that the ratio of average consumption among the top 10 of the distribution to the average consumption among the bottom 10 had risen considerably even before 1998/99, suggesting that the more well-off had benefited more than the poorest decile from the economic growth in 1991-98. Over the years, the consumption levels of the top and the bottom of the distribution continued to diverge at a steady rate so that the gap expanded by 30 % over the full period.⁴ The divergence was widening because the bottom 10 was being left behind, rather than because the top 10 was gaining disproportionately compared with the rest of the population. The average consumption of the 90-th percentile rose little relative to the median, while the average consumption of the bottom 10 had deteriorated by nearly 20 % by 2005/06. The bottom 10 appears to be losing ground also compared with other households in the bottom 25, who are also losing ground to the median but only half as quickly.

These preliminary findings denote a clear tendency towards rising polarization in household consumption over the period. The notion of “polarization” commonly refers to the case where there is a significant number of individuals who are very poor but there exists also a non-negligible share of the population that is quite rich. Such a gap between the poor and the rich implies evidently that there is no sizeable middle class.⁵ As we will see later when applying relative distribution methods, the distributional changes that occurred between 1991/92 and 2012/13 hollowed out the middle of the Ghanaian household consumption distribution and increased the concentration of households around the highest and lowest deciles, hence leading to an increase of polarization.

2.2 Relative distribution methods

2.2.1 Basic concepts

⁴ The gap between 90-th and 10-th deciles is probably a lower bound of the real one. In general, household surveys do not contain good estimates of upper percentiles of welfare (Alvaredo and Piketty, 2010). When using consumption to rank welfare, as it is normally done in low/er middle income countries, the situation is further aggravated. Consumption is very accurate in capturing the well-being of poorer people, yet it is rather imprecise in capturing that of people living in upper percentiles.

⁵ In this paper we will analyze the median-based approach to the measurement of polarization. Since it subdivides the population into two subgroups – those above the median and those below the median, respectively – we refer to this as the case of “bi-polarization”. For a detailed explanation of the main differences not only between the study of inequality and that of polarization, but also between the concept of bi-polarization and that of “multi-polar” polarization, see e.g. Chakravarty (2009, ch. 4), Deutsch et al. (2013) and Chakravarty (2015).

To address the question of the “hollowing out of the middle” in Ghana, we use *relative distribution* methods. Developed by Handcock and Morris (1998, 1999), these techniques based on the relative distribution powerfully assist in the description of distributional change and enable counterfactual comparison of location-adjusted distributions.

Basically, relative distribution methods can be applied whenever the distribution of some quantity across two populations is to be compared, either cross-sectionally or over time.⁶ For our purposes, the relative distribution is defined as the ratio of the density in the comparison year to the density in the reference year evaluated at each decile of the consumption distribution, and can be interpreted as the fraction of households in the comparison year’s population that fall into each tenth of the reference year’s distribution.⁷ This allows us to identify and locate changes that have occurred along the entire Ghanaian household consumption distribution.

To formalize, let:

$$p_i^t = \frac{1}{m} \sum_{j=1}^m \mathbf{I} \left(c^0 [i-1] \leq y_j^t < c^0 [i] \right), \quad i = 1, \dots, 10, \quad (1)$$

be the proportion of households in year t ’s comparison sample falling into each decile and:

$$p_i^0 = \frac{1}{n} \sum_{j=1}^n \mathbf{I} \left(c^0 [i-1] \leq y_j^0 < c^0 [i] \right), \quad i = 1, \dots, 10, \quad (2)$$

be the proportion of households in year 0’s reference sample falling into the same deciles, where m and n reflect the comparison and reference sample sizes and:

$$\mathbf{I} (S) = \begin{cases} 1 & \text{if the event } S \text{ is true} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

denotes the indicator function. The cut points $c^0[i-1]$ and $c^0[i]$ for each interval are estimated as deciles of the reference sample, hence the proportion of the sample from the reference distribution falling into each decile is exactly 1/10. The relative distribution is given by the proportion of year t ’s households

⁶ Here we limit ourselves to illustrating the basic concepts behind the use of relative distribution methods. Interested readers are referred to Handcock and Morris (1998, 1999) – but see also Hao and Naiman (2010, ch. 5) – for a more detailed explication and a discussion of the relationship to alternative econometric methods for measuring distributional differences. A method very similar in spirit to the relative distribution one has recently been developed by Silber et al. (2014).

⁷ To keep the notation simple and the graphical displays informative, we will focus throughout the paper on *group-level* data from underlying continuous distributions. We will also assume that discretization is based on decile ranges with respect to the reference year’s distribution. By extending the fundamental concepts of the relative distribution approach to the grouped data context, Handcock and Morris (1999, ch. 11) allow the analysis of discretized distributions to retain the tractability and interpretability of its continuous counterpart based on ungrouped data. For an application of relative distribution methods to ungrouped samples, see e.g. the Brazilian and Nigerian case studies by, respectively, Clementi and Schettino (2015) and Clementi et al. (2014, 2015).

whose consumption expenditures fall into each decile estimated from the reference distribution, divided by the proportion in the reference year:

$$g^t(i) = \frac{P_i^t}{P_i^0}, \quad i = 1, \dots, 10. \quad (4)$$

When the fraction of the comparison population in a decile is higher (lower) than the fraction in the reference year, the relative distribution will be higher (lower) than 1. When there is no change, the relative distribution will be flat at the value 1. Therefore, in this way one can distinguish between growth, stability or decline at specific points of the consumption distribution.

2.2.2 The location/shape decomposition of the relative distribution

One of the major advantages of this approach is the possibility to decompose the relative distribution into changes in *location*, usually associated with changes in the median (or mean) of the distribution, and changes in *shape* (including differences in variance, asymmetry and/or other distributional characteristics) that could be linked to several factors such as, for instance, polarization. The decomposition can be represented in the following terms:

$$g^t(i) = \underbrace{\frac{P_i^{0L}}{P_i^0}}_{\text{Location effect}} \times \underbrace{\frac{P_i^t}{P_i^{0L}}}_{\text{Shape effect}}, \quad i = 1, \dots, 10, \quad (5)$$

where:

$$p_i^{0L} = \frac{1}{n} \sum_{j=1}^n \mathbf{I} \left(c^0 [i-1] \leq y_j^{0L} < c^0 [i] \right), \quad i = 1, \dots, 10, \quad (6)$$

denotes the proportion of households in each estimated decile range of the original reference distribution whose consumption expenditures have been median-adjusted by an additive shift to yield identical centers of the comparison and reference distributions, while the shapes of the two distributions remain the same.⁸

⁸ In formal notation, the *median-adjusted* reference variable is $Y^{0L} = Y^0 + \rho$, where Y^0 denotes the year 0's consumption variable and the value ρ is the difference between the medians of the comparison and reference distributions. Median adjustment is preferred here to mean adjustment because of the well-known drawbacks of the mean when distributions are skewed. A *multiplicative* median shift can also be applied. However, the multiplicative shift has the drawback of affecting the shape of the distribution. Indeed, the equi-proportionate changes increase the variance and the rightward shift of the distribution is accompanied by a flattening (or shrinking) of its shape – see e.g. Jenkins and Van Kerm (2005).

The first ratio term in the right hand side of Equation (5) is an estimate of the “location effect”, i.e. the pattern that the relative distribution would have displayed if there had been no change in distributional shape but only a location shift of the consumption distribution over time. When the median-adjusted and unadjusted reference populations have the same median, the ratio for location differences will have a uniform distribution. Conversely, when the two distributions have different median, the location effect is increasing (decreasing) in i if the comparison median is higher (lower) than the reference one.

The second term (the “shape effect”) represents the relative distribution net of the location effect and is useful to isolate movements (re-distribution) occurred between the reference and comparison populations. For instance, one could observe a shape effect with some sort of (inverse) U-shaped pattern if the comparison distribution is relatively (less) more spread around the median than the median-adjusted reference distribution. Thus, it is possible to determine whether there is polarization of the consumption distribution (increases in both tails), “downgrading” (increases in lower tail), “upgrading” (increases in the upper tail) or convergence towards the median (decreases in both tails).

The graphical display provides a useful visual summary of the relative size and nature of the three components of the decomposition in Equation (5).

2.2.3 Relative polarization indices

Another relevant feature of these methods is that one can use summary measures to quantify the observed pattern of changes. Morris et al. (1994) and Handcock and Morris (1998, 1999) developed a measure of polarization that captures the degree to which there is divergence from, or convergence toward, the center of the distribution, and is thus ideally suited to addressing the question of the “hollowing of the middle”. For group-level data, the *median relative polarization index* (MRP) takes the form (Morris et al., 1994, p. 217; Handcock and Morris, 1999, p. 190):

$$\text{MRP} = \frac{4}{Q-2} \sum_{i=1}^Q \left| \frac{i - \frac{1}{2}}{Q} - \frac{1}{2} \right| g^t(i) - \frac{Q}{Q-2}, \quad (7)$$

where $g^t(i)$, $i = 1, \dots, Q$, are the relative proportions in (4) and the adjustment by $1/2$ establishes the mid-point for each group. The expression for a decile aggregation is easily obtained from Equation (7) by setting $Q = 10$. The index varies between -1 and 1. It takes the value of 0 when there has been no change in the distribution of household consumption relative to the reference year. Positive values signify relative polarization (i.e. growth in the tails of the distribution) and negative values signify relative convergence toward the center of the distribution (i.e. less polarization).

The median relative polarization index can be decomposed into the contributions to distributional change made by the segments of the distribution above and below the median, enabling one to distinguish “upgrading” from “downgrading”. For grouped data, the *lower relative polarization index* (LRP) and the *upper relative polarization index* (URP) are calculated as:

$$\text{LRP} = \frac{8}{Q-2} \sum_{i=1}^{Q/2} \left| \frac{i-\frac{1}{2}}{Q} - \frac{1}{2} \right| g'(i) - \frac{Q}{Q-2}, \quad (8)$$

$$\text{URP} = \frac{8}{Q-2} \sum_{i=Q/2+1}^Q \left| \frac{i-\frac{1}{2}}{Q} - \frac{1}{2} \right| g'(i) - \frac{Q}{Q-2}. \quad (9)$$

They have the same theoretical range as the MRP and decompose the overall polarization index in the following way (Handcock and Morris, 1998, 1999):

$$\text{MRP} = \frac{1}{2} (\text{LRP} + \text{URP}). \quad (10)$$

To test the hypothesis of no change with respect to the reference distribution, i.e. that the three indices have a statistically significant difference from zero, we use the asymptotic distribution of the estimates under the non-parametric null hypothesis that the reference and comparison distributions are identical. Under this hypothesis, the distribution of the group-level estimates of the MRP is asymptotically normal with a mean equal to 0 and a variance equal to (Morris et al., 1994, p. 218):

$$\text{Var}(\text{MRP}) = \frac{1}{3} \left(\frac{1}{m} + \frac{1}{n} \right). \quad (11)$$

Distributional approximations for the LRP and URP are similar. The variance in both cases is approximately (Handcock and Morris, 1999, p. 170):

$$\frac{5}{3} \left(\frac{1}{m} + \frac{1}{n} \right). \quad (12)$$

Therefore, given a chosen significance level, the p -value for testing the null hypothesis $H_0: \text{RP}=0$ against the alternative that one of the three indices is different from zero can be calculated as:

$$p\text{-value} = 1 - \Phi \left(\frac{|\text{RP}|}{\sqrt{\text{Var}(\text{RP})}} \right), \quad (13)$$

where $\Phi(\cdot)$ is the standard normal distribution function and RP denotes the median, lower or upper polarization index.

In practice, the normal approximations will be very good for sample sizes of 50 or more. As the sample sizes in our study (m and n) are typically on the order of thousands, the distributional approximations involved are excellent.

2.3 Blinder-Oaxaca type decomposition of location and shape differences

In this section we present a novel method for analyzing the effects of covariates on the observed distributional changes due to both the location and shape shifts. Novel because in the original relative distribution framework, the method proposed to measure the impact of polarization drivers does not provide intuitive results and it is of limited use for policy making purposes. By contrast, our method that combines the relative distribution approach and the regression based decompositions, can produce an easily interpretable set of results.

In the relative distribution setting, the exploration of the distributional impacts of changes in covariates requires that the overall relative density is adjusted for these changes using the technique described in Handcock and Morris (1999, ch. 7). This technique partials out the impact of changes in the distribution of the covariates – the “composition effect” – and the modifications in the conditional distributions of household consumption expenditure given the covariate levels – the “residual effect”. Conceptually, this parallels the traditional regression-based decomposition that separates changes in covariates (the X 's) from changes in the “returns” to the covariates (the regression coefficients, or β 's). However, the covariate adjustment technique proposed by Handcock and Morris does not provide a simple and intuitively accessible way of dividing up the changes exclusively due to a location shift or shape differences into the contribution of changes in the distribution of each single covariate and that of the changing “returns” to the covariates; also, differently from what happens in the classical regression decomposition approach, its drawback is making it difficult to summarize the contributions above into a single value as, for example, the estimated coefficients obtained by the regression procedure would make it possible to quantify.

The framework we propose integrates the spirit of the relative distribution approach and recent developments from the regression-based decomposition literature. This can be regarded as an extension of the covariate adjustment technique developed by Handcock and Morris and can be used to quantify the impact of an arbitrary number of covariates on distributional differences due to both location and shape shifts, so as to identify the key drivers of these changes.

In detail, we decompose the component relative distributions that represent differences in location and shape by applying a procedure recently proposed by Firpo et al. (2009) for the decomposition of wage differentials. The method is based on running unconditional quantile regressions to estimate the impact of changing the distribution of explanatory variables along the entire distribution of the dependent variable and using the traditional Blinder (1973) and Oaxaca (1973) decomposition framework to decompose differentials at selected quantiles of the consumption distribution.

To estimate the unconditional quantile regression, we have first to derive the *re-centered influence function* (RIF) for the τ -th quantile of the dependent variable distribution – consumption, in our case – which can be shown as (Firpo et al., 2009; Essama-Nssah and Lambert, 2011; Fortin et al., 2011):

$$\text{RIF}(c; q_\tau, F_C) = \begin{cases} q_\tau + \frac{\tau}{f_C(q_\tau)}, & c > q_\tau, \\ q_\tau - \frac{1-\tau}{f_C(q_\tau)}, & c < q_\tau, \end{cases} \quad (14)$$

where q_τ is the sample quantile and $f_C(q_\tau)$ is the density of consumption C at the τ -th quantile. In practice, the RIF is estimated by replacing all unknown quantities by their observable counterparts. In the case of (14) unknown quantities are q_τ and $f_C(q_\tau)$, which are estimated by the sample τ -th quantile of C and a standard non-parametric kernel density estimator, respectively. Firpo et al. (2009) show that the unconditional quantile regression can be implemented by running a standard OLS regression of the estimated RIF on the covariates X :⁹

$$\text{E}[\text{RIF}(C; q_\tau, F_C) | X = x] = X\beta_\tau, \quad (15)$$

where the coefficient β_τ represents the approximate marginal effect of the explanatory variable X on the τ -th unconditional quantile of the household consumption distribution. Applying the law of iterated expectations to the above equation, we also have:

$$q_\tau = \text{E}_X[\text{E}[\text{RIF}(C; q_\tau, F_C) | X = x]] = \text{E}[X]\beta_\tau. \quad (16)$$

This yields an *unconditional quantile interpretation*, where β_τ can be interpreted as the effect of increasing the mean value of X on the unconditional quantile q_τ .¹⁰

⁹ This can be performed using the Stata's command `rifreg`, which is available for download at <http://faculty.arts.ubc.ca/nfortin/datahead.html>.

¹⁰ As discussed in more detail by Fortin et al. (2011), one important reason for the popularity of OLS regressions in economics is that they provide consistent estimates of the impact of an explanatory variable, X , on the population *unconditional* mean of an outcome variable, Y . This important property stems from the fact that the conditional mean, $\text{E}[Y | X = x]$, averages up to the unconditional mean, $\text{E}[Y]$, due to the law of iterated

Using unconditional quantile (RIF) regression, an aggregate decomposition for location and shape differences can then be implemented in a spirit similar to the Blinder-Oaxaca decomposition of mean differentials as follows:

$$\hat{\Delta}_\tau^t = \hat{c}_\tau^t - \hat{c}_\tau^0 = \hat{\Delta}_X^t + \hat{\Delta}_\beta^t + \hat{\Delta}_J^t, \quad (17)$$

where the total difference in consumption at the same quantile τ of the year t 's comparison and year 0's reference distributions, $\hat{\Delta}_\tau^t$, is decomposed into one part that is due to differences in observable characteristics (endowments) of the households, $\hat{\Delta}_X^t$, one part that is due to differences in returns (coefficients) to these characteristics, $\hat{\Delta}_\beta^t$, and a third part – for which no clear interpretation exists – that is due to interaction between endowments and coefficients, $\hat{\Delta}_J^t$. In particular, once the RIF regressions for the τ -th quantile of the comparison and reference consumption distributions have been run, the estimated coefficients can be used as in the standard Blinder-Oaxaca decomposition to perform a detailed decomposition into contributions attributable to each covariate. The aggregate decomposition can be generalized to the case of the detailed decomposition in the following way:¹¹

$$\hat{\Delta}_\tau^t = \underbrace{\sum_{k=1}^K (\bar{X}_k^t - \bar{X}_k^0) \hat{\beta}_{\tau,k}^0}_{\hat{\Delta}_X^t} + \underbrace{(\hat{\alpha}^t - \hat{\alpha}^0) + \sum_{k=1}^K (\hat{\beta}_{\tau,k}^t - \hat{\beta}_{\tau,k}^0) \bar{X}_k^0}_{\hat{\Delta}_\beta^t} + \underbrace{\sum_{k=1}^K (\bar{X}_k^t - \bar{X}_k^0) (\hat{\beta}_{\tau,k}^t - \hat{\beta}_{\tau,k}^0)}_{\hat{\Delta}_J^t}, \quad (18)$$

expectations. As a result, a linear model for conditional means, $E[Y|X=x] = X\beta$, implies that $E[Y] = E[X]\beta$, and OLS estimates of β also indicate what is the impact of X on the population average of Y . When the underlying question of economic and policy interest concerns other aspects of the distribution of Y , however, estimation methods that “go beyond the mean” have to be used. A convenient way of characterizing the distribution of Y is to compute its quantiles. A quantile regression model for the τ -th conditional quantile $q_\tau(X)$ postulates that $q_\tau(X) = X\beta_\tau$. By analogy with the case of the mean, β_τ can be interpreted as the effect of X on the τ -th conditional quantile of Y given X . Unlike conditional means, however, conditional quantiles do not average up to their unconditional population counterparts, i.e. $q_\tau(Y) \neq E_X[q_\tau(X)] = E[X]\beta_\tau$, where $q_\tau(Y)$ is the unconditional quantile. As a result, the estimated β_τ cannot be interpreted as the effect of increasing the mean value of X on q_τ . RIF regression offers instead a simple way of establishing a direct link between unconditional quantiles of the distribution of Y and household characteristics X because of (16), which says that the conditional expectation of (15) – the expected value of the RIF – is equal to the unconditional quantile of interest.

¹¹ Following Jones and Kelley (1984), we focus here on the so-called “threefold” decomposition, which uses the same reference distribution for both $\hat{\Delta}_X^t$ and $\hat{\Delta}_\beta^t$ but introduces the interaction term $\hat{\Delta}_J^t$. Equations (17) and (18)

can also be written by reversing the reference and comparison distribution designation for both $\hat{\Delta}_X^t$ and $\hat{\Delta}_\beta^t$, as

well as by allocating the interaction term to either $\hat{\Delta}_X^t$ or $\hat{\Delta}_\beta^t$ so as to implement a “twofold” decomposition.

However, while these various versions are used in the literature, using one or the other does not involve any specific estimation issue (Fortin et al., 2011). Hence, for the sake of exposition, we shall utilize the decomposition introduced in the text for the rest of our analysis.

where k represents the k -th covariate and $\hat{\alpha}$ and $\hat{\beta}_{\tau,k}$ are the estimated intercept and slope coefficients, respectively, of the RIF regression models for the comparison and reference samples.¹²

Specifically, since we use an additive median shift to identify and separate out changes due to location differences in the consumption distribution, the decompositions above are carried out using the medians ($\tau = 0.5$) of the location-adjusted and unadjusted reference populations, so that the total difference to be decomposed according to (17) and (18) is:

$$\hat{\Delta}_{0.5}^{0L} = \hat{c}_{0.5}^{0L} - \hat{c}_{0.5}^0 = \rho, \quad (19)$$

where ρ denotes the difference between the medians of the year t 's comparison and year 0's reference distributions (see footnote 6). As location-adjustment is performed by adding ρ to every household consumption expenditure of the original reference population to match its median with that of the comparison population, without altering the shape, the decomposition of the differential (19) can be operated once and its results assumed to hold simultaneously across the entire relative distribution representing changes exclusively due to a location shift. For what concerns the shape shift, the differentials to be decomposed are instead as follows:

$$\hat{\Delta}_{\tau}^t = \hat{c}_{\tau}^t - \hat{c}_{\tau}^{0L}, \quad \tau = 0.1, \dots, 0.9, \quad (20)$$

where the quantiles c_{τ} are estimated as deciles of the comparison and location-adjusted distributions – the latter having the median of the comparison sample but the shape of the reference one.

Notice that the differentials (20) represent *horizontal* distances, or decile gaps, between the distributions involved in the decomposition exercise, whereas the idea underlying the relative distribution framework typically focuses on *vertical* ratios, or relative proportions. Hence, the “declining middle class” scenario would suggest that negative differentials $\hat{\Delta}_{\tau}^t$ are to be expected for deciles below the median, whereas for those above the median the total differences given by (20) should be positive. Intuitively, this is because in this case the population shifts from the center of the consumption distribution to the upper and lower deciles, so that the cut-off points identifying the

¹² Notice that in order to decompose the total difference $\hat{\Delta}_{\tau}^t$ according to (18) it is also necessary to estimate two counterfactual consumption distributions, namely, the distribution that can be obtained by combining the distribution of characteristics of the comparison sample with the returns for households' observable characteristics of the reference sample, $\bar{X}^t \hat{\beta}_{\tau}^0$, and the distribution obtained by combining the distribution of characteristics of the reference sample with the returns for households' characteristics of the comparison sample, $\bar{X}^0 \hat{\beta}_{\tau}^t$, where \bar{X} represents the covariates mean. This can be done automatically within Stata by invoking Jann's (2008) `oaxaca8` command, which is the routine used in this study to perform empirical applications of Equation (18).

deciles below the median in the comparison distribution come before those of the reference distribution along the consumption scale, while cut-off points for deciles above the median come after.

3 Results

3.1 Changes in the Ghanaian consumption distribution

To introduce the results obtained from using the methods and data described in previous sections, in Figure 1(a) we present two probability density functions of the Ghanaian distribution of total consumption expenditure.¹³ The solid line is the distribution of household consumption in 1991/92, taken as the baseline throughout the analysis. The density drawn with the dotted line, which we will treat as the comparison, is the distribution in 2012/13.¹⁴ Examining these two distributions, we see that the reference or 1991/92 distribution has a slight right skewness, while the comparison distribution has a larger median and variance.

However, the graphical display above does not provide much information on the relative impact that location and shape changes had on the differences in the two distributions at every point of the expenditure scale. It also does not convey whether the upper and lower tails of the consumption distribution were growing at the same rate and for what reasons (i.e. location and/or shape driven). As already pointed out in Subsection 2.2, this is exactly what relative distribution methods are particularly good at pulling out of the data.

The relative density of total consumption expenditure of Ghanaian households between 1991/92 and 2012/13 is examined in Figure 1(b), showing the fraction of households in 2012/13 that fall into each decile of the 1991/92 distribution.¹⁵ The graph offers the immediate impression that the proportion of households in the upper deciles increased dramatically throughout the two decades, while the

¹³ To handle data sparseness, the two densities have been obtained by using an adaptive kernel estimator with a Silverman's plug-in estimate for the pilot bandwidth (see e.g. Van Kerm, 2003). The advantage of this estimator is that it does not over-smooth the distribution in zones of high expenditure concentration, while keeping the variability of the estimates low where data are scarce – as, for example, in the highest expenditure ranges.

¹⁴ Obviously, reversing the reference and comparison distribution designation will change the view provided by the relative distribution graph and the displays of the estimated effects of location and shape shifts, because these are defined in terms of the reference distribution scale. However, designating which distribution will serve as the reference is a decision that must be made by the analyst, and in our application the natural choice was suggested by time ordering. In addition, the relative polarization indices (measurements of the degree to which a comparison distribution is more polarized than a reference distribution, and defined in terms of the relative distribution of the comparison relative to the median-adjusted reference) are *symmetric*, meaning that they are effectively invariant to whether the 1991/92 or 2012/13 consumption distribution is chosen as the reference – in fact, swapping the comparison and reference distributions yields indices of the same magnitude and opposite sign (Handcock and Morris, 1999, pp. 71-72; Hao and Naiman, 2010, pp. 88-89). Thus, reversing the reference and comparison distributions designation will not alter our findings in a substantive way – if not for the fact that polarization would now be analyzed in the reverse direction of time.

¹⁵ Throughout, we rely on the R statistical package `reldist` (Handcock, 2015) to implement the relative distribution method.

proportion in the bottom and around the middle declined. Indeed, if we choose any decile between the first and the seventh in the 1991/92 distribution, the fraction of households in 2012/13 whose consumption rank corresponds to the chosen decile is less than the analogous fraction of households in 1991/92.

While the display of the relative distribution points to the dominant trend for the entire period, the dominant trend may be masking some of the more subtle changes. To see these, we decompose the relative density into location and shape effects according to Equation (5). Figure 1(c) presents the effect only due to the median shift, that is the pattern that the relative density would have displayed if there had been no change in distributional shape but only a location shift of the density. The effect of the median shift was quite large. This alone would have virtually eliminated the households in the first four deciles of the 1991/92 consumption distribution and placed a considerable fraction of them in the top end of the 2012/13 distribution. Note, however, that neither tail of the observed relative distribution is well reproduced by the median shift. For example, the top decile of Figure 1(c) is about 2.5, below the value of 3.6 observed in the actual data, and the bottom deciles of the same figure are also substantially lower than observed.

These (and other) differences are explained by the shape effect presented in Figure 1(d), which shows the relative density net of the median influence. Without the higher median, the greater dispersion of consumption expenditures would have led to relatively more low-consuming households in 2012/13, and this effect was mainly concentrated in the bottom decile. By contrast, at the top of the distribution the higher spread worked in the same direction of the location shift: operating by itself, it would have increased the share of households in the top decile of the 2012/13 consumption distribution by nearly 120%. In sum, once changes in real median expenditure are netted out, a U-shaped relative density is observed, indicating that polarization was hollowing out the middle of Ghanaian household consumption.

Relative distribution methods permit us to also analyze how re-distribution across households took place over the entire time period. For each wave of the GLSS between 1991/92 and 2012/13, Figure 2 shows the shape effect of the household consumption relative density using 1991/92 as the reference sample.¹⁶ Following the plot through each successive wave, one is offered with the immediate impression that the fraction of households at both the top and bottom tails of the Ghanaian consumption distribution increased consistently over the course of the last two decades, while the fraction in the middle declined. Polarization, or the “hollowing out of the middle”, has been therefore the consistent trend in distributional inequality for all the GLSS waves since 1991/92. Because this period was also characterized by a sizable shift in location, viewed together these results indicate that,

¹⁶ The relative distribution, and therefore its shape effect, is by definition flat in the reference year (Morris et al., 1994, p. 211).

in the course of the upswing in consumption expenditures, some households fell behind, while others shifted toward the top, joining the ranks of those whose consumption put them in the top decile in 1991/92.

To summarize these changes, we present in Figure 3 the set of relative polarization indices computed from the GLSS data using Equations(7)-(9).¹⁷ These indices track changes in the shape of the distribution only, and they code the direction as well as the magnitude of the change. The overall index (MRP) rises continuously and the rise is statistically significant from the outset, thus confirming the visual impression from Figure 1(d). Decomposing the MRP into the contributions from the lower and upper tails of the distribution, it also appears that “downgrading” dominated “upgrading” in the polarization upswing – the value of the LRP is indeed always greater than that of the URP.

3.2 Temporal decomposition

To get a more compact picture of the timing and nature of the polarization trend described above, we can break the 21-year period into 3 sub-periods – 1991-98, 1998-2005, and 2005-12 – and highlight the changes that took place *within* each of them. The top three panels of Figure 4 show the relative distribution for each sub-period. In contrast to the 21-year decile series, which takes 1991/92 as the reference distribution for all waves, each panel here takes the beginning year of the sub-period for the reference distribution and the end year for the comparison. The displays clearly point to the median up-shift in household consumption expenditure as the dominant trend for each sub-period. These are the images of a “rising tide that lifts all boats”, i.e. the effect of a location shift that was the most influential contributor to the overall pattern during all sub-periods. The differences due to the median shift – representing what the relative density would have looked like if there had been no change in distributional shape – are plotted in the middle row panels of Figure 4. As expected, the strongest effects were in the bottom deciles, confirming that more low-consuming households joined the ranks of those whose consumption levels put them in the top half of the reference distributions. However, once changes in location are netted out, there is also an indication of growing polarization that is not evident in the overall relative distributions. The differences explained by the shape changes are presented in the bottom row panels of Figure 4, where the median-adjusted relative distributions take an approximate U-shape. Strong growth occurred in the fraction of households at the top and bottom tails of the period-specific consumption distributions, while sizable declines occurred in the middle. This polarizing trend seems nearly symmetric for the years 2005 to 2012, while throughout the 1990s and up to the mid-2000s the growth in the lower tail of the distribution was noticeably stronger than in the upper tail.

¹⁷ Since the value of the three indices always equals 0 in the baseline year (Morris et al., 1994, p. 209), polarization summaries for 1991/92 were not included in the graphical display.

The relative polarization indices, shown in Table 3, capture these changes well. The MRP index is always positive and statistically significant (p -value = 0.00). Decomposing the MRP into the contributions to distributional change made by the segments of the distribution above and below the median, it appears that “downgrading” dominated “upgrading” in the polarization upswing over the course of the first two sub-periods: the value of the lower relative polarization index (LRP) is indeed greater than that of the upper relative polarization index (URP) – 0.26 vs. 0.17 and 0.27 vs. 0.11, respectively – which is consistent with the visual impression from the shape shifts above. The values of the indices in the 2005-12 period denote instead a nearly perfectly symmetric polarization in each tail.

In sum, while often less noticeable in any single period when compared to the large swings in median household consumption expenditure, the growth in polarization was a major contributor to the overall changes in the Ghanaian consumption distribution since the early 1990s. Behind these shape shifts, however, was probably a set of key drivers. The following section relies therefore on GLSS data to examine how the changes above have been associated with consumption growth and, thereby, identifies the main drivers behind the polarization upsurge.

3.3 The drivers of growing polarization in Ghanaian household consumption

The presentation of polarization results over the three sub-periods requires a considerable amount of space. For the sake of brevity, we chose to present only part of the results and made an effort to present the main findings in an abridged format. For example, we decided not to comment on the econometric results of the unconditional quantile regression, and to place the decomposition tables in the appendix, and regarding the polarization decomposition results, to focus our attention only on the top percentiles results (top two and bottom two).

Overall this is not a big limitation since, as shown in panels (a) and (b) of Figure 5, the inter-quantile analysis has detected a significant variation in the percentiles’ cut-offs (between deciles inequality, measured by interquartile ratios) primarily among these deciles and a very limited one among the rest of the distribution. Furthermore, the other component of polarization, the so-called “identification” (measured by deciles’ coefficient of variation, CV) tended to be more accentuated in these deciles rather in the central ones. Looking at sub-periods, it clearly emerges that, in 1991-98 and 2005-12, the between component was compensated by a high identification component, thus neutralizing the modification of inequality; differently, in the sub-period 2005-12 it appears both a sustained growth of between component and an important reduction of identification component (growth of CV) especially for what concern the 10-th and 90-th deciles.

Table 4 compares the counterfactual cut-off points (labelled with “c”) – the cut-offs of the reference distribution augmented with the location effect between the two sub-periods – with the cut-offs of the comparison distribution. In all three sub-periods, the cut-offs of the bottom percentiles of the comparison distribution are significantly lower than those of the reference, indicating, as we discussed in the previous subsection, lower relative polarization, whereas for the top percentiles the opposite holds: the comparison distribution cut-offs are higher than the reference ones, indicating upper relative polarization.

The Oaxaca-Blinder (OB) methodology (Oaxaca, 1973; Blinder, 1973), decomposes the difference between cut-offs into that part that is due to group differences in the magnitudes of the determinants (endowments effect) of consumption, on the one hand, and group differences in the effects of these determinants (coefficients effect), on the other. Coefficient and endowment variations are aggregated by groups of variables: primary, secondary and tertiary education are grouped into the education attainment group; private, public and self-employment of household head are grouped into employment category; the infrastructure index captures the access to basic services;¹⁸ urbanization and residence in regions other than Upper East (urban and regional dummies having as baseline Upper East); and household structure (household size and all other household characteristics). The interaction term and the constant are also included so that the sum of all decomposition elements adds up to the total differences between cut-offs. Below any decomposition graph, we present a table summarizing the main variable trends for upper and lower polarization.

Recalling previous section results regarding 1991-98 sub-period, the polarization increased as testified by the shifts leftward and rightward of the lower and upper cut-offs respectively. The polarization decomposition shows how the combined effect of household composition, infrastructure index and the constant increased the lower polarization while location effects and education tended to reduce the effect. On the upper deciles nearly the same variables played a pro-polarization role (Figure 6). Between 1991 and 1998 growth concentrated in urban areas and in few regions on the Coast or in the immediate inland (Ashanti region) among households with relatively higher levels of education and with access to a number of basic infrastructures. This group of households occupying the top two deciles of the distribution distances itself from the rest of other groups determining an increase in the upper polarization.

The 1998-2005 sub-period sees polarization growing. In this decade, Ghana experienced a boom in cocoa production and exports. The cocoa boom generated, in the western and coastal areas, a high demand for the workforce, but also for capital and infrastructure, and the skills of the workforce and the rise in revenues even at lower levels translated into a higher demand for capital, infrastructure and

¹⁸ The infrastructure index is obtained by combining four variables through principal component analysis: access to protected water, access to electricity, access to protected sanitation, and access to safe sources of cooking.

skills (Molini and Paci, 2015). These resources were relatively scarce, and the price effect and variation in returns was, thus, substantial. In these areas, the cocoa boom had a positive impact on poverty, but did not benefit everybody equally.

The drivers of polarization, both upper and lower, were very similar (Figure 7). Household characteristics, educational attainment and basic infrastructures all tended to have pro-inequality outcome and increased the tails size of the 2005 distribution, indeed more polarization. It is worth noting the particular importance of changes in the household structure in explaining the upper polarization. Top deciles were particularly benefitting from the demographic dividend stemming from smaller families and lower dependency ratios. The only set of variables that countered this increase were the location/urban ones. The cocoa boom and the relatively good performance of many rural areas in the Central and Coastal part of the country such as Ashanti, Volta, Eastern, Western and Central region (Molini and Paci, 2015) explains this positive distributional impact.

Finally, between 2005 and 2012, the upper polarization substantially stagnates. Compared to the previous sub-period, the distributional changes of this sub-period are driven by a positive variation in endowments and stagnation in the returns on covariates (see appendix). This seems to suggest that the high returns obtained in the previous period encouraged households to invest in assets and human capital. This clearly reduced their scarcity, but, at the same time, returns massively declined. The greater availability of people in the nonfarm sector who had low levels of educational attainment (typically primary school) determined a clear decline in their relative returns (Molini and Paci, 2015).

Differently from the previous period, urban and regional variables drive polarization (Figure 8). Households residing in Greater Accra and the urban areas of Ashanti region performed well and increased their relative economic advantage over the rest of the country. Interestingly, the drivers of upper polarization are very similar to those playing a role in the 1991-98 sub-period. In addition to the urban and regional variables, the infrastructure index, the employment variables and education had a strong impact on polarization. As for 1998-2005, the variations in household composition benefit the top percentiles and contribute significantly to the increase of polarization.

4 Concluding remarks

The topic of the increasing gap between the richer and poorer is gaining momentum thanks, in particular, to the large attention that has been obtained in recent research on world inequalities (see e.g. Stiglitz, 2012, 2014, Piketty, 2014, and Atkinson, 2015, inter alia). The overall idea that emerges is that in the last 20/30 years both developing and developed countries went through dramatic distributional changes that increased disparities.

The main contribution of our paper is proposing a tool that displays these changes but also identifies and quantifies the underlying drivers. We focus our attention on polarization defined as the combination of divergence from global and convergence on mean local incomes. The method developed blends two different frameworks of distributional analysis: relative distribution (Handcock and Morris, 1998 1999) and unconditional quantile regression (Firpo et al., 2009). The advantage over other methodologies is that it allows to single out the different drivers of polarization at different points of the consumption distribution.

Ghana, almost unique among SSA countries, offers the opportunity to analyze the last two decades' distributional changes, since four comparable household surveys are available. The country also presents interesting specificities. Since 1991, poverty had declined very fast, inequality has not increased dramatically and yet the country has seen a rapid surge in polarization. The results of our analysis suggest that the distributional changes hollowed out the middle of the Ghanaian household consumption distribution and increased the concentration of households around the highest and lowest deciles.

Results on drivers of polarization indicate that although there is some heterogeneity across the various sub-periods in particular in terms of magnitude, household characteristics, educational attainment and access to basic infrastructures all tended to increase over time the size of the upper and lower tails of the consumption distribution and as a consequence the degree of polarization. Urban rural and regional variables started to have a strong impact on polarization only in the last decade; households residing in Greater Accra and the urban areas of Ashanti region performed well and increased their relative economic advantage over the rest of the country.

From a policy perspective, the pro-polarization impact of variables that tend to change slowly over time is of particular concern. It is very unlikely that policy makers can find a quick fix to the problem and any intervention will produce results only in the long run. This implies that the country needs to start now to develop a strategy that, if not able to immediately reverse polarization, at least can mitigate its impact. The creation of a modern social protection system, the expansion in the access to basic services, the continued effort to expand primary and secondary education are all interventions that can pay off and help the country to maintain its social cohesion.

References

- Alvaredo, F., and Piketty, T. (2010), “The dynamics of income concentration in developed and developing countries: a view from the top”, in López-Calva, L. F., Lustig, N. (eds.), *Declining Inequality in Latin America: A Decade of Progress?*, pp. 72-99 (Baltimore: Brookings Institution Press).
- Aryeetey, E., and Baah-Boateng, W. (2015), *Understanding Ghana’s Growth Success Story and Job Creation Challenges* (Working Paper No. 140, World Institute for Development Economics Research, United Nations University, Helsinki, available at <https://www.wider.unu.edu/publication/understanding-ghana%E2%80%99s-growth-success-story-and-job-creation-challenges>).
- Atkinson, A. B. (2015), *Inequality: What Can Be Done?* (Cambridge MA: Harvard University Press).
- Blinder, A. (1973), “Wage discrimination: reduced forms and structural estimates”, *The Journal of Human Resources*, 8, 436-55.
- Chakravarty, S. R. (2009), *Inequality, Polarization and Poverty: Advances in Distributional Analysis* (New York: Springer-Verlag).
- (2015), *Inequality, Polarization and Conflict: An Analytical Study* (New York: Springer-Verlag).
- Clementi, F., Dabalén, A. L., Molini, V., and Schettino, F. (2014), *The Centre Cannot Hold: Patterns of Polarization in Nigeria* (Working Paper No. 149, World Institute for Development Economics Research, United Nations University, Helsinki, available at http://www.wider.unu.edu/publications/working-papers/2014/en_GB/wp2014-149/).
- (2015), “When the centre cannot hold: patterns of polarization in Nigeria”, *Review of Income and Wealth*, DOI: 10.1111/roiw.12212.
- Clementi, F., and Schettino, F. (2015), “Declining inequality in Brazil in the 2000s: what is hidden behind? ”, *Journal of International Development*, 27, 929-52.
- Cooke, E. F. A., Hague, S., and McKay, A. (2016), *The Ghana Inequality and Poverty Analysis: Using the Ghana Living Standards Survey*, mimeo.
- Deaton, A., and Zaidi, S. (2002), *Guidelines for Constructing Consumption Aggregates for Welfare Analysis* (LSMS Working Paper No. 135, World Bank, Washington DC, available at <https://openknowledge.worldbank.org/handle/10986/14101>).

- Deutsch, J., Fusco, A., and Silber, J. (2013), “The BIP trilogy (bipolarization, inequality and polarization): one saga but three different stories”, *Economics: The Open-Access, Open-Assessment E-Journal*, 7, 1-33.
- Duclos, J.-Y., Esteban, J.-M., and Ray, D. (2004), “Polarization: concepts, measurement, estimation”, *Econometrica*, 72, 1737-72.
- Essama-Nssah, B., and Lambert, P. (2011), *Influence Functions for Distributional Statistics* (ECINEQ Working Paper No. 236, Society for the Study of Economic Inequality, Verona University, Italy, available at <http://www.ecineq.org/milano/WP/ECINEQ2011-236.pdf>).
- Esteban, J.-M., and Ray, D. (1999), “Conflict and distribution”, *Journal of Economic Theory*, 87, 379-415.
- (2008), “Polarization, fractionalization and conflict”, *Journal of Peace Research*, 45, 163-82.
- (2011), “Linking conflict to inequality and polarization”, *The American Economic Review*, 101, 1345-74.
- Firpo, S., Fortin, M., and Lemieux, T. (2009), “Unconditional quantile regressions”, *Econometrica*, 77, 953-73.
- Fortin, M., Lemieux, T., and Firpo, S. (2011), “Decomposition methods in economics”, in Card, D., and Ashenfelter, O. (eds.), *Handbook of Labor Economics*, vol. 4, part A, pp. 1-102 (Amsterdam: North-Holland).
- Foster, J. E., and Wolfson, M. C. (1992), *Polarization and the Decline of the Middle Class: Canada and the US* (OPHI Working Paper No. 31, University of Oxford, Oxford, available at <http://www.ophi.org.uk/working-paper-number-31/>, now in *Journal of Economic Inequality*, 8, 247-73, 2010).
- Handcock, M. S. (2015), *Relative Distribution Methods* (1.6-4 edn.; Los Angeles, CA, project home page at <http://www.stat.ucla.edu/~handcock/RelDist>).
- Handcock, M. S., and Morris, M. (1998), “Relative distribution methods”, *Sociological Methodology*, 28, 53-97.
- (1999), *Relative Distribution Methods in the Social Sciences* (New York: Springer-Verlag).
- Hao, L., and Naiman, D. Q. (2010), *Assessing Inequality* (Thousand Oaks CA: SAGE Publications, Inc.).
- Haughton, J., and Khandker, S. R. (2009), *Handbook on Poverty and Inequality* (Washington DC: World Bank).

- Jann, B. (2008), "The Blinder-Oaxaca decomposition for linear regression models", *Stata Journal*, 8, 453-79.
- Jenkins, S. P., and Van Kerm, P. (2005), "Accounting for income distribution trends: a density function decomposition approach", *Journal of Economic Inequality*, 3, 43-61.
- Jones, F. L., and Kelley, J. (1984), "Decomposing differences between groups. A cautionary note on measuring discrimination", *Sociological Methods and Research*, 12, 323-43.
- Molini, V. and Paci, P. (2015). *Poverty Reduction in Ghana: Progress and Challenges* (World Bank, Washington DC, available at <https://openknowledge.worldbank.org/handle/10986/22732>).
- Morris, M., Bernhardt, A. D., and Handcock, M. S. (1994), "Economic inequality: new methods for new trends", *American Sociological Review*, 59, 205-19.
- Oaxaca, R. (1973), "Male-female wage differentials in urban labour markets", *International Economic Review*, 14, 693-709.
- Piketty, T. (2014), *Capital in the Twenty-First Century* (Cambridge MA: The Belknap Press of Harvard University).
- Silber, J., Deutsch, J., and Yalonetzky, G. (2014), "On bi-polarization and the middle class in Latin America: a look at the first decade of the twenty-first century", *Review of Income and Wealth*, 60, S332-S52.
- Stiglitz, J. E. (2012), *The Price of Inequality: How Today's Divided Society Endangers Our Future* (New York: W. W. Norton & Company).
- (2015), *The Great Divide: Unequal Societies and What We Can Do About Them* (New York: W. W. Norton & Company).
- Van Kerm, P. (2003), "Adaptive kernel density estimation", *The Stata Journal*, 3, 148-56.
- Wolfson, M. C. (1994), "When inequalities diverge", *The American Economic Review*, 84, 353-58.
- (1997), "Divergent inequalities: theory and empirical results", *Review of Income and Wealth*, 43, 401-21.

Tables

Table 1: Summary measures of Ghanaian household total consumption expenditure, 1991/92 to 2012/13.

	1991/92	1998/99	2005/06	2012/13
Observations	4,523	5,998	8,687	16,772
Mean	459.91	568.45	736.80	883.48
Median	352.66	438.04	559.44	655.60
Consumption shares				
Bottom 5	1.11	1.00	0.79	0.82
Bottom 10	2.71	2.42	2.08	2.13
Bottom 20	6.82	6.21	5.65	5.63
Top 20	44.78	44.47	46.59	46.94
Top 10	29.16	28.17	30.75	30.43
Top 5	18.52	17.41	19.95	19.17
Inequality measures				
Gini	0.38	0.38	0.41*	0.41
Theil	0.25	0.25	0.30*	0.29

* Denotes statistically significant change from the previous period at the 5 % level (p -value < 0.05).

Source: authors' own calculation using GLSS data sets.

Table 2: Inter-quantile consumption ratios by GLSS Wave, 1991/92 to 2012/13.

Wave	p10/p50	p25/p50	p75/p25	p75/p50	p90/p10	p90/p50
1991/92	0.46	0.66	2.37	1.56	5.23	2.42
1998/99	0.41	0.63	2.60	1.64	6.00	2.48
2005/06	0.39	0.61	2.63	1.62	6.36	2.46
2012/13	0.39	0.62	2.68	1.66	6.73	2.65

Source: authors' own calculation using GLSS data sets.

Table 3: Relative polarization indices by sub-periods, 1991/92 to 2012/13.

	Index	<i>p</i> -value
1998/99 to 1991/92		
MRP	0.22	0.00
LRP	0.26	0.00
URP	0.17	0.00
2005/06 to 1998/99		
MRP	0.19	0.00
LRP	0.27	0.00
URP	0.11	0.00
2012/13 to 2005/06		
MRP	0.14	0.00
LRP	0.14	0.00
URP	0.14	0.00

Source: authors' own calculation using GLSS data sets.

Table 4: Counterfactual Reference cut-offs vs. comparison cut-offs: by deciles and sub-periods.

Decile	1991c	1998	1998c	2005	2005c	2012
1-st	248.74	181.03	302.43	216.83	312.99	258.36
2-nd	296.69	246.72	368.12	304.00	400.17	357.47
8-th	704.60	803.14	924.54	1,011.40	1,107.56	1,242.97
9-th	940.64	1,084.86	1,206.26	1,377.14	1,473.31	1,738.20

Source: authors' own calculation using GLSS data sets.

Figures

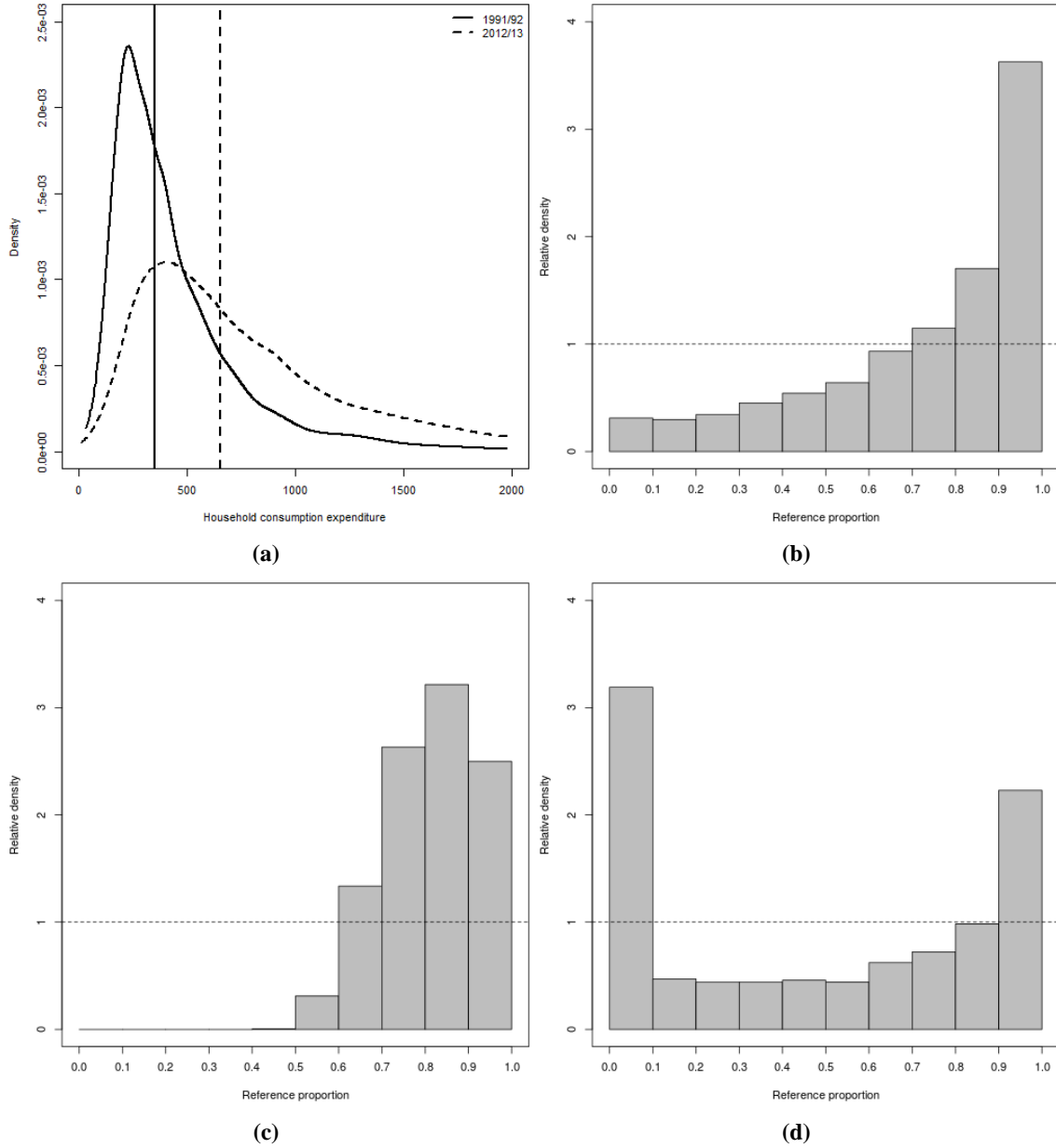


Figure 1: Changes in the Ghanaian household consumption distribution between 1991/92 and 2012/13. **(a)** Kernel distributions. Expenditures in the upper tiers of the densities have been truncated for better presentation of the graph, where the vertical lines denote the medians of the two survey waves. **(b)** Relative consumption distribution. **(c)** The effect of the median difference in consumption growth. **(d)** The median-adjusted relative consumption distribution (the effect of changes in distributional shape).

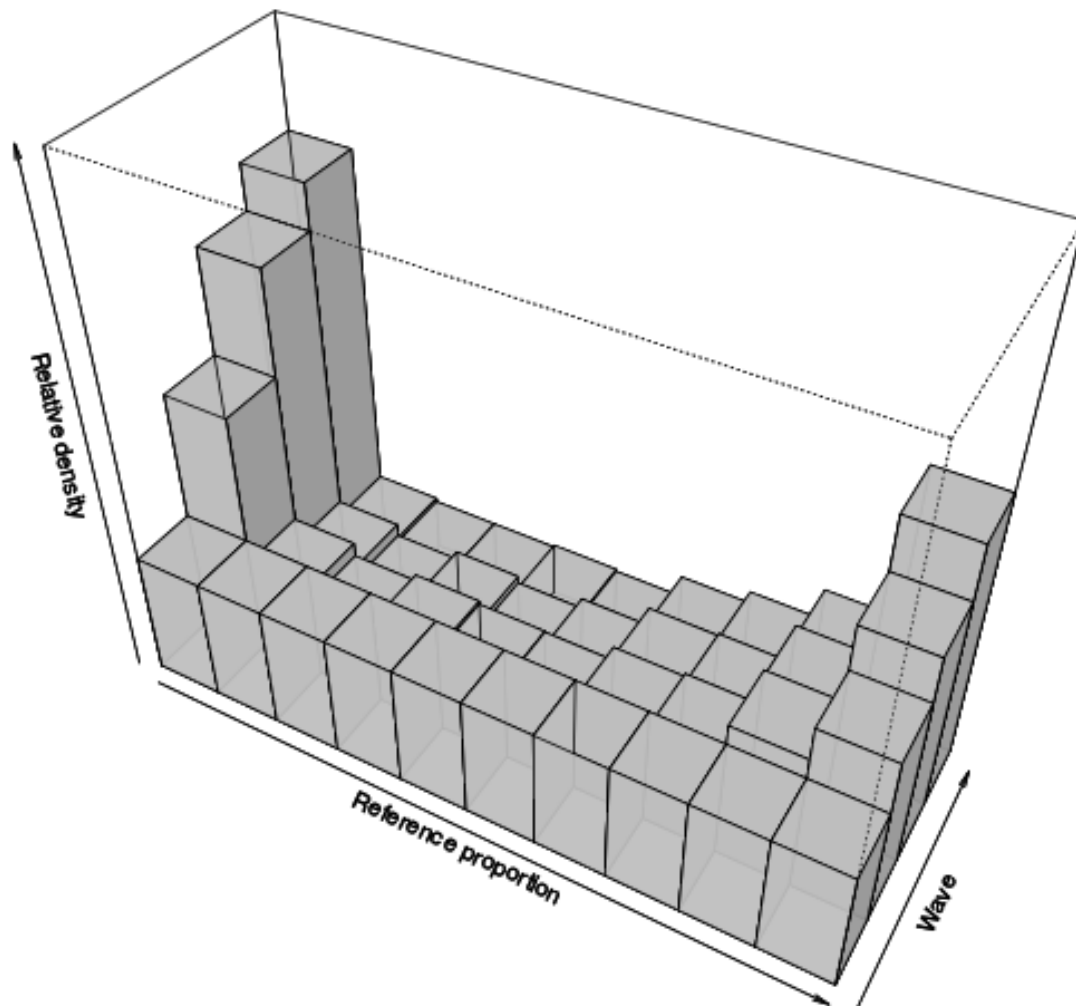


Figure 2: Median-adjusted relative consumption distribution series for Ghana, 1991/1992 to 2012/2013.

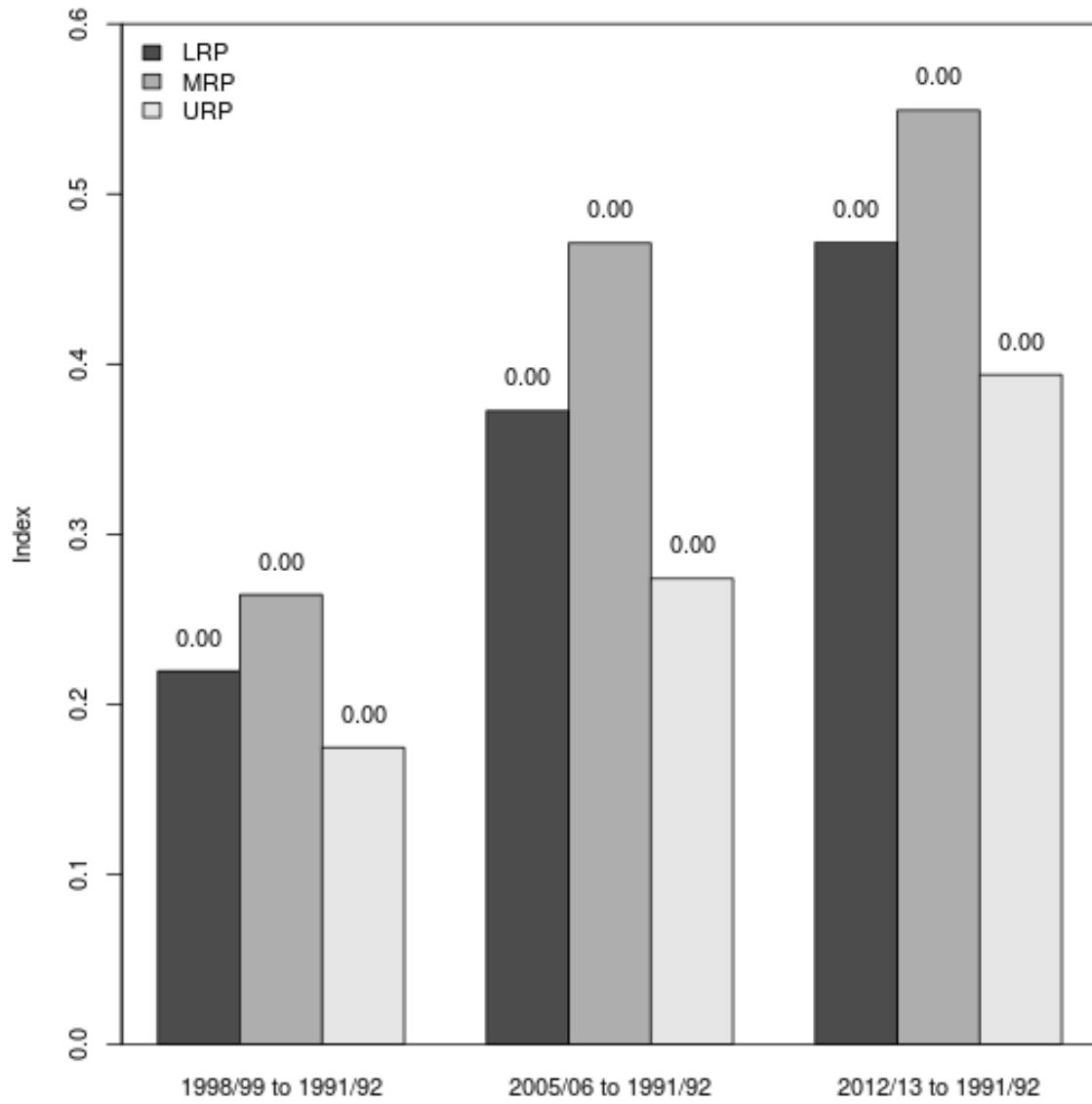


Figure 3: Relative polarization indices by wave. The number above each bar indicates the p -value for the null hypothesis that the index equals 0.

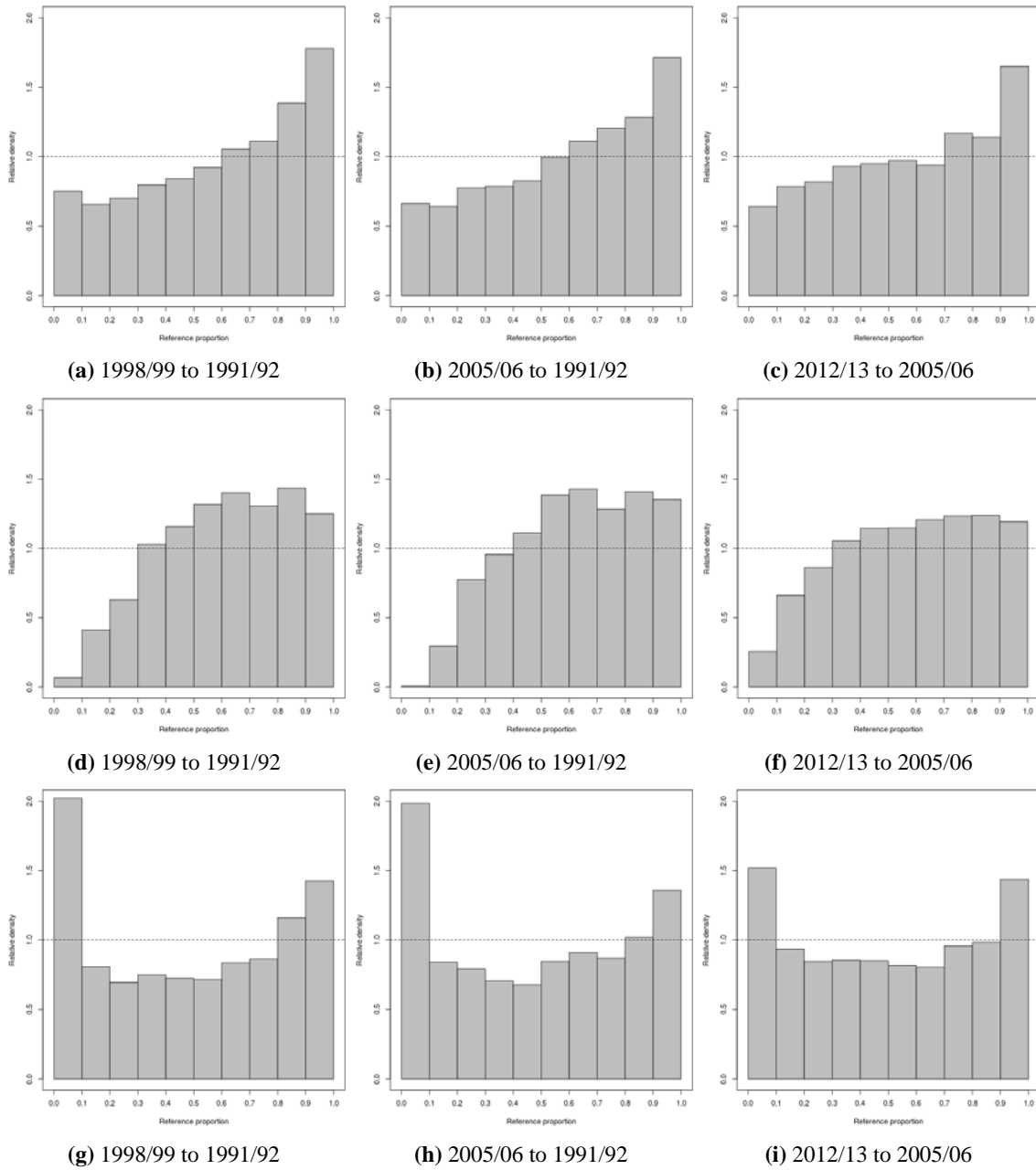


Figure 4: Location and shape decomposition of the relative consumption distribution for Ghana by sub-periods. The top row shows the overall change by sub-period, the middle shows the effect of the median shift (the shape-adjusted relative distribution), and the bottom shows the effect of the shape changes (the median-adjusted relative distribution).

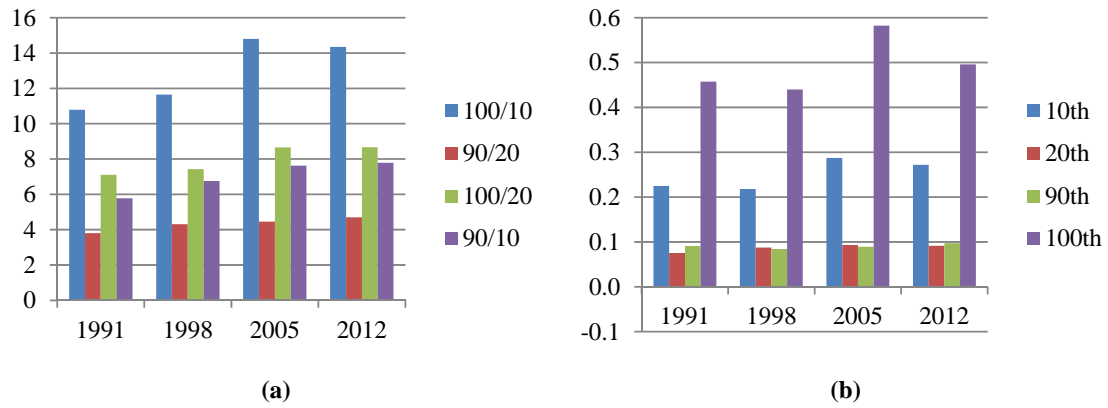
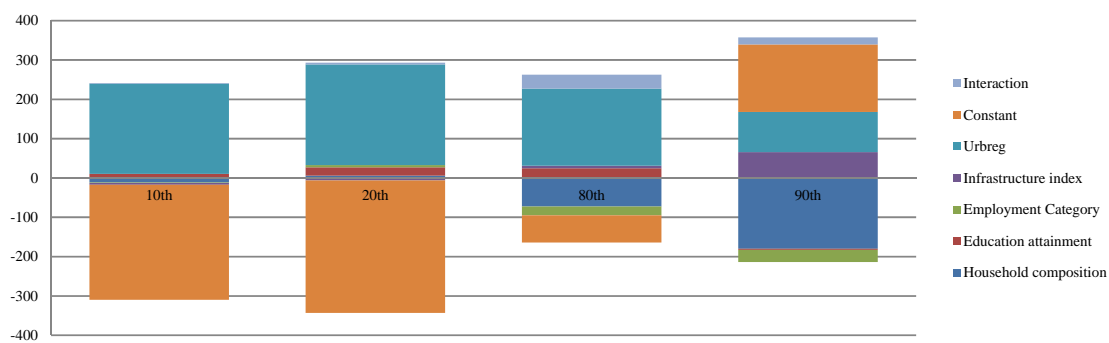
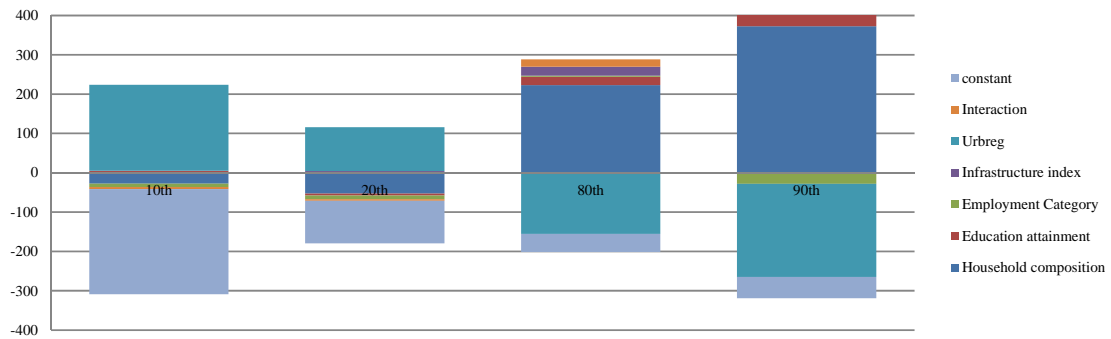


Figure 5: (a) Inter-decile ratio by year, using counterfactual distributions; (b) coefficient of variation, by year and decile.



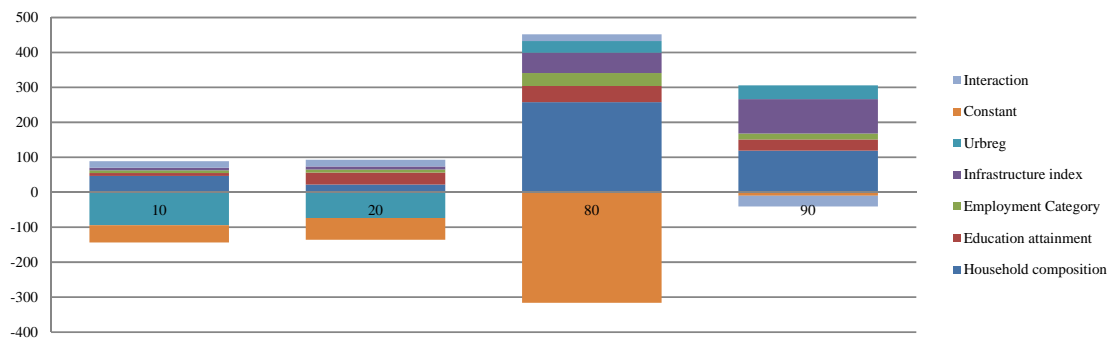
Total	Lower polarization	Upper polarization
Urbreg	---	+++
Infrastructure	+	+
Education	--	++
Employment	-	-
Household	+	----
Constant	++++	++

Figure 6: Blinder-Oaxaca type decompositions, 1991-98.



Total	Lower polarization	Upper polarization
Urbreg	---	---
Infrastructure	+	+
Education	++	++
Employment	+	-
Household	++	++++
Constant	+	+

Figure 7: Blinder-Oaxaca type decompositions, 1998-2005.



Total	Lower polarization	Upper polarization
Urbreg	++++	++
Infrastructure	-	+++
Education	-	++
Employment	-	++
Household	--	++++
Constant	++	---

Figure 8: Blinder-Oaxaca type decompositions, 2005-12.

Appendix

Table A.1: Location effect RIF-regression results.

		1991	1998	2005	2012
Number of obs		4,523	5,998	8,687	16,772
F(25, 4497)		61.57	144.54	166.58	273.29
Prob > F		0.00	0.00	0.00	0.00
R-squared		0.27	0.29	0.32	0.32
Adj R-squared		0.26	0.29	0.32	0.31
Root MSE		241.33	322.80	413.75	500.16
		<i>Coef.</i> <i>P>z</i>	<i>Coef.</i> <i>P>z</i>	<i>Coef.</i> <i>P>z</i>	<i>Coef.</i> <i>P>z</i>
Demographic Features	Household size	-15.08 0.00	-21.83 0.00	-29.53 0.00	-33.16 0.00
	Share of Children	-32.44 0.27	-79.78 0.02	-20.62 0.60	-32.67 0.41
	Share of Care-Dependent Persons	9.47 0.76	-0.58 0.99	83.26 0.03	-16.22 0.70
	Household Head Age	-0.40 0.36	-0.57 0.24	-0.95 0.07	-1.18 0.03
	Sex of Household Head	-16.34 0.16	7.83 0.54	56.80 0.00	60.86 0.00
	Share of Adult Males	139.49 0.00	122.12 0.00	110.30 0.00	213.38 0.00
	Share of Adult Females	207.10 0.00	265.36 0.00	327.24 0.00	418.16 0.00
Education Features	Up to Primary School	9.42 0.55	21.63 0.18	41.40 0.03	41.15 0.02
	Up to Secondary School	32.80 0.01	55.43 0.00	84.64 0.00	111.18 0.00
	Higher than Secondary School	100.94 0.00	129.84 0.00	232.60 0.00	302.11 0.00
Socioeconomic Features	Private Workers	41.62 0.04	48.09 0.02	93.72 0.00	57.21 0.01
	Public Workers	56.47 0.00	53.63 0.01	100.33 0.00	58.92 0.04
	Non Agricultural Self Employed	57.42 0.00	44.95 0.00	117.33 0.00	132.62 0.00
	Agricultural Self Employed	36.81 0.04	-7.08 0.65	6.11 0.75	10.02 0.61
Other	Assets	40.53 0.00	78.13 0.00	78.43 0.00	117.91 0.00
	Western	-34.48 0.21	339.59 0.00	267.86 0.00	211.91 0.00
	Central	65.31 0.02	152.87 0.00	261.49 0.00	146.10 0.00
	Greater Accra	-3.50 0.90	349.13 0.00	131.97 0.00	323.70 0.00
	Volta	17.54 0.54	206.22 0.00	165.38 0.00	160.94 0.00
	Eastern	9.02 0.76	229.43 0.00	299.66 0.00	169.09 0.00
	Ashanti	42.01 0.12	253.64 0.00	223.86 0.00	186.47 0.00
	Brong Ahafo	-29.55 0.28	248.58 0.00	173.82 0.00	187.04 0.00
	Northern	7.51 0.79	146.83 0.00	162.93 0.00	59.73 0.00
	Upper East	-106.72 0.00	28.75 0.18	-5.44 0.82	89.90 0.00
	Urban Area Residence	102.49 0.00	27.64 0.05	143.21 0.00	92.73 0.00
Constant	338.30 0.00	219.03 0.00	266.81 0.00	295.86 0.00	

Table A.2: Location effect OB results.

		1998-91		2005-1998		2012-05	
Median predicted (1)		438.18		559.53		655.62	
Median predicted (2)		352.69		438.18		559.53	
Difference		<i>Coef.</i>	<i>P>z</i>	<i>Coef.</i>	<i>P>z</i>	<i>Coef.</i>	<i>P>z</i>
		85.49	0.00	121.36	0.00	96.09	0.00
		<i>Endowments</i>					
Demographic Features	Household size	7.26	0.00	-1.53	0.17	3.92	0.00
	Share of Children	-0.17	0.43	0.85	0.05	0.14	0.63
	Share of Care-Dependent Persons	0.10	0.77	0.00	0.99	0.16	0.39
	Household Head Age	-0.30	0.41	0.12	0.56	-0.30	0.25
	Sex of Household Head	0.51	0.21	1.25	0.54	-0.97	0.03
	Share of Adult Males	2.10	0.00	2.36	0.00	0.25	0.46
	Share of Adult Females	4.43	0.00	1.69	0.04	3.10	0.00
Education Features	Up to Primary School	0.64	0.55	-0.67	0.20	2.32	0.04
	Up to Secondary School	0.39	0.29	1.18	0.04	0.92	0.13
	Higher than Secondary School	2.30	0.00	0.78	0.13	3.05	0.00
Socioeconomic Features	Private Workers	0.57	0.11	2.18	0.03	2.89	0.00
	Public Workers	-2.22	0.00	-0.83	0.05	-0.71	0.08
	Non Agricultural Self Employed	1.83	0.00	-1.52	0.01	6.74	0.00
	Agricultural Self Employed	2.36	0.04	0.31	0.65	-0.07	0.77
Other	Assets (see note)	11.05	0.00	12.20	0.00	23.96	0.00
	Western	-0.37	0.36	-3.06	0.08	-2.33	0.03
	Central	0.79	0.16	-4.27	0.00	0.26	0.80
	Greater Accra	-0.11	0.91	-3.38	0.10	3.21	0.00
	Volta	1.06	0.54	-14.02	0.00	1.99	0.00
	Eastern	-0.34	0.76	6.69	0.00	-8.74	0.00
	Ashanti	0.85	0.20	-2.60	0.11	6.22	0.00
	Brong Ahafo	1.29	0.29	4.72	0.00	1.09	0.12
	Northern	-0.23	0.80	8.28	0.00	-3.27	0.00
	Upper East	3.53	0.00	0.76	0.19	0.04	0.84
	Urban Area Residence	2.69	0.01	0.65	0.12	17.60	0.00
<i>Total</i>	40.02	0.00	12.16	0.01	61.45	0.00	

Table A.2: Continued.

		<i>Coefficients</i>					
Demographic Features	Household size	-42.27	0.07	-44.53	0.05	-21.26	0.39
	Share of Children	-9.85	0.28	12.62	0.25	-2.44	0.83
	Share of Care-Dependent Persons	-0.40	0.82	4.20	0.10	-4.66	0.08
	Household Head Age	-7.63	0.80	-17.61	0.60	-10.72	0.76
	Sex of Household Head	15.43	0.16	29.74	0.03	3.12	0.88
	Share of Adult Males	-3.67	0.67	-2.67	0.80	25.30	0.04
	Share of Adult Females	14.82	0.20	17.06	0.23	25.64	0.10
Education Features	Up to Primary School	1.39	0.59	3.60	0.44	-0.04	0.99
	Up to Secondary School	9.06	0.24	12.04	0.20	11.50	0.28
	Higher than Secondary School	0.80	0.47	5.18	0.01	3.92	0.09
Socioeconomic Features	Private Workers	0.40	0.83	3.45	0.15	-4.41	0.25
	Public Workers	-0.37	0.91	4.30	0.17	-3.17	0.29
	Non Agricultural Self Employed	-2.33	0.53	15.84	0.00	2.83	0.56
	Agricultural Self Employed	-4.36	0.06	2.16	0.59	0.47	0.89
Other	Assets (see note)	-15.35	0.00	-0.04	0.98	0.81	0.12
	Western	37.21	0.00	-7.90	0.04	-5.65	0.12
	Central	9.07	0.02	12.56	0.00	-10.12	0.00
	Greater Accra	41.40	0.00	-32.32	0.00	26.68	0.00
	Volta	15.50	0.00	-5.83	0.24	-0.33	0.90
	Eastern	31.40	0.00	7.38	0.04	-17.52	0.00
	Ashanti	33.44	0.00	-5.31	0.34	-6.28	0.25
	Brong Ahafo	32.38	0.00	-5.43	0.03	1.21	0.70
	Northern	13.18	0.00	1.03	0.62	-12.43	0.00
	Upper East	7.34	0.00	-0.72	0.30	4.54	0.00
	Urban Area Residence	-24.47	0.00	40.81	0.00	-19.02	0.03
	Constant	-119.27	0.02	47.78	0.38	29.04	0.62
<i>Total</i>		32.83	0.00	97.41	0.00	17.01	0.04
		<i>Interaction</i>					
<i>Total</i>		12.64	0.01	11.79	0.02	17.63	0.00

Table A.3: Shape effect RIF-regression results.

	1991				1998				2005				2012				
	10th	20th	80th	90th	10th	20th	80th	90th	10th	20th	80th	90th	10th	20th	80th	90th	
Number of obs	4523	4523	4523	4523	5998	5998	5998	5998	8687	8687	8687	8687	16772	16772	16772	16772	
R-squared	0.127	0.1734	0.2317	0.1763	0.1861	0.247	0.2613	0.1856	0.2751	0.2835	0.2675	0.1993	0.183	0.2469	0.2508	0.1817	
Adj R-squared	0.1221	0.1688	0.2274	0.1717	0.1827	0.2439	0.2582	0.1821	0.273	0.2814	0.2654	0.197	0.1818	0.2458	0.2497	0.1805	
Root MSE	166.26	171.9	537.4	1044.6	210.52	222.58	679.77	1115.3	264.97	283.64	839.63	1665.2	318.58	332.91	1247	1952.8	
	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	Coef.	P>z	
Demographic Features	Household size	-5.455	0.006	-6.928	0.000	-32.832	0.000	-51.281	0.000	-8.061	0.000	-10.717	0.000	-48.615	0.000	-64.869	0.000
	Share of Children	-21.170	0.320	-43.926	0.048	-111.941	0.064	-104.837	0.351	3.000	0.887	-23.086	0.318	-128.827	0.046	-103.057	0.292
	Share of Care-Dependent Persons	-8.818	0.678	-3.716	0.866	46.025	0.526	47.323	0.743	-6.929	0.746	-3.083	0.891	90.496	0.212	232.394	0.064
	Household Head Age	0.267	0.433	0.040	0.905	-1.449	0.088	-2.502	0.090	0.064	0.849	0.004	0.991	-2.302	0.009	-5.539	0.000
	Sex of Household Head	-9.927	0.223	-20.890	0.014	-12.108	0.600	8.489	0.843	3.588	0.677	6.933	0.456	40.865	0.097	73.518	0.050
	Share of Adult Males	38.180	0.045	60.945	0.002	295.954	0.000	614.298	0.000	4.293	0.838	21.861	0.325	265.513	0.000	514.394	0.000
Share of Adult Females	51.753	0.010	59.880	0.007	527.357	0.000	1061.252	0.000	64.299	0.005	105.117	0.000	563.185	0.000	913.254	0.000	
Education Features	Up to Primary School	-1.718	0.888	6.207	0.603	-4.850	0.863	46.613	0.352	22.156	0.055	39.210	0.001	-0.712	0.978	-14.756	0.692
	Up to Secondary School	20.687	0.015	10.693	0.231	70.770	0.003	126.750	0.002	36.533	0.000	49.809	0.000	110.820	0.000	95.150	0.013
	Higher than Secondary School	31.035	0.031	46.767	0.002	79.310	0.323	226.447	0.165	49.212	0.000	55.819	0.000	283.759	0.000	461.548	0.000
Socioeconomic Features	Private Workers	16.356	0.195	29.349	0.017	-6.958	0.876	-3.960	0.961	16.356	0.086	32.899	0.008	15.531	0.748	-23.545	0.753
	Public Workers	23.548	0.029	33.154	0.003	109.050	0.002	105.537	0.119	40.818	0.000	66.810	0.000	-3.288	0.941	-51.651	0.447
	Non Agricultural Self Employed	29.897	0.001	25.997	0.006	125.672	0.000	117.759	0.030	21.088	0.029	37.049	0.000	87.867	0.002	89.237	0.038
	Agricultural Self Employed	10.539	0.385	13.894	0.250	56.689	0.091	43.785	0.459	-16.849	0.157	-4.316	0.714	6.674	0.829	-11.287	0.787
Other	Assets (see note)	0.153	0.969	7.849	0.058	109.418	0.000	220.129	0.000	12.241	0.015	28.248	0.000	166.629	0.000	205.553	0.000
	Western	36.647	0.135	21.765	0.343	-150.440	0.006	-151.534	0.072	320.415	0.000	370.078	0.000	291.636	0.000	265.460	0.000
	Central	53.456	0.024	57.058	0.011	7.668	0.896	72.621	0.422	286.159	0.000	277.257	0.000	31.524	0.311	-21.939	0.600
	Greater Accra	39.498	0.091	26.341	0.230	-31.771	0.612	19.447	0.851	295.482	0.000	334.505	0.000	481.309	0.000	626.103	0.000
	Volta	61.673	0.008	43.929	0.047	-65.669	0.255	-58.228	0.513	246.610	0.000	267.853	0.000	148.898	0.000	192.025	0.000
	Eastern	49.943	0.040	34.317	0.139	-25.750	0.669	-1.156	0.990	301.459	0.000	336.623	0.000	111.418	0.000	41.520	0.315
	Ashanti	55.452	0.017	49.261	0.022	9.256	0.866	126.459	0.137	290.869	0.000	316.334	0.000	227.152	0.000	206.419	0.000
	Brong Ahafo	43.935	0.070	27.833	0.226	-40.178	0.475	51.060	0.558	298.072	0.000	327.446	0.000	217.315	0.000	180.736	0.000
	Northern	-39.280	0.145	-24.473	0.315	7.602	0.855	103.440	0.240	168.506	0.000	186.967	0.000	183.467	0.000	199.252	0.000
	Upper East	-80.541	0.030	-105.691	0.001	-94.546	0.114	5.736	0.952	46.488	0.309	75.230	0.017	73.709	0.034	103.582	0.031
	Urban Area Residence	39.315	0.000	63.669	0.000	183.278	0.000	237.926	0.000	23.203	0.008	17.726	0.078	56.254	0.039	26.511	0.523
	Constant	196.253	0.000	273.832	0.000	755.886	0.000	875.930	0.000	-95.854	0.004	-63.309	0.025	686.530	0.000	1047.534	0.000
										-242.144	0.000	-50.372	0.171	761.965	0.000	1114.916	0.000
														-194.757	0.000	-16.084	0.594
																541.794	0.000
															1201.148	0.000	

