UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

Dipartimento di Scienze economiche e metodi matematici

SOUTHERN EUROPE RESEARCH IN ECONOMIC STUDIES

Inequality of Opportunity in Sub-Saharan Africa

Paolo Brunori, Flaviana Palmisano and Vito Peragine

SERIES WORKING PAPERS N. 08/2015

Inequality of Opportunity in Sub-Saharan Africa*

Paolo Brunori†
Flaviana Palmisano‡
Vito Peragine§

December 2015

Abstract

In the last decades inequality of opportunity has been extensively studied by economists, on the assumption that, in addition to being normatively undesirable, it can be related to low potential for growth. In this paper we evaluate inequality of opportunity in 11 Sub-Saharan Africa countries. According to our results, the portion of total inequality which can be attributed to exogenous circumstances is between 30% and 40% for the generality of countries considered. We also find a positive association between total consumption inequality and inequality of opportunity and we study the different sources of unequal opportunities. Finally, we address a number of methodological issues that typically arise when measuring inequality of opportunity with imperfect data, which is the typical case in developing countries.

Keywords: Consumption inequality, Equality of opportunity, Sub-Saharan Africa.

JEL codes: D63, E24, O15, O40.

---

*This work was supported by the World Bank in the framework of the "Poverty in Africa: Revisiting the Facts" report. We are grateful to Francisco Ferreira, Kathleen Beegle, Isis Gaddis, and Camila Galindo Pardo for their comments and support. The paper was presented at the 6th ECINEQ Meeting, July 2015, University of Luxembourg and at the Second SITES/IDEAS Meeting, October 2015, University of Florence, we are grateful to all the participants for useful comments.All errors remain our own.

†University of Bari. E-mail: paolo.brunori@uniba.it.

‡University of Luxembourg. E-mail: flaviana.palmisano@gmail.com

§University of Bari. E-mail: vitorocco.peragine@uniba.it

1 Introduction

Sub-Saharan Africa (SSA) countries are especially known for their high levels of economic inequality and poverty (see, for instance, Ellis, 2012; Moradi and Baten, 2005; Thorbecke, 2013). However, the specific features of these inequalities remain largely understudied. Yet the understanding of the different sources of inequality is a necessary step toward the implementation of policies that may foster a sustained and 'shared' growth in these countries. There is in fact a rooted consensus on the argument that not all inequalities are the same: in particular, it has been convincingly argued (see World Bank, 2006; Ferreira et al., 2014; Marrero and Rodriguez, 2013) that the degree of the inequality caused by differences at birth (such as gender, ethnicity, or parental background) or, more generally, by factors beyond the individual control may be related to low growth, more so than other effort-based inequalities. The idea is that when exogenous circumstances play a strong role in determining individual outcome, there is a sub-optimal allocation of resources and lower potential for growth. To put it differently, the existence of inequality traps, which systematically exclude some groups of the population from participation in the economic activity, is harmful to growth because they discourage effort and investment by individuals, provoke a loss of productive potential, and contribute to social and institutional instability. The arguments above suggest that analysing the specific 'horizontal' dimensions of inequality is particularly important in both developing and underdeveloped countries.

One way to assess these kinds of inequalities is to implement the Equality of Opportunity (EOp) framework (see Roemer, 1998; Fleurbaey, 2008), which provides a model to distinguish between that part of inequality caused by exogenous circumstances outside the individual responsibility, considered to be objectionable and therefore deserving a compensatory intervention, and the part of inequality generated by individual choices and effort, which is on the contrary considered to be fair and not to be eliminated. The EOp theory has spurred a huge amount of theoretical and empirical works focussing on the measurement of inequality of opportunity (see the recent surveys by Ferreira and Peragine, 2015; Ramos and Van de gaer, 2015; Roemer and Trannoy, 2015). However, most of the literature has been concerned with inequality of opportunity (IOp) in Western developed countries, with only a small set of studies dedicated to developing countries.[1] One reason for this is that measuring opportunity inequality is not an easy task: its informational requirements are quite high if compared to the standard measurement of income or consumption inequality.[2] Therefore, these are more commonly met in surveys and databases that refer to wealthier countries. Hence, as argued above, such analysis would be particularly needed in developing countries.

This paper is a contribution in this direction. Specifically, it is the first attempt to evaluate inequality of opportunity in a large set of SSA countries by using 13 different surveys that contain information about individual circumstances and outcomes.

Our contribution to the literature is twofold. First, we contribute to the understanding of economic inequality in 11 SSA countries (i) by showing the portion of consumption inequality, which can be attributed to inequality of opportunities, and

[1]In particular, only two contributions exist in the literature, namely Cogneau and Mesplé-Somps (2008) and Piraino (2015) that propose an analysis of inequality of opportunity for African countries.

[2]See on this Hassine (2011).

(ii) by identifying the most disadvantaged groups of the population in each country. This analysis can help to understand the social and economic mechanisms that generate inequalities and can help in identifying priorities in anti-poverty policies in different countries. Second, this paper offers a methodological contribution to the literature on the measurement of inequality of opportunity by addressing a number of methodological issues that typically arise in the realisation of this task in the presence of imperfect data, which is the typical case in developing countries.

Our analysis is made possible through the availability of large-sample surveys built upon a common methodology and providing information on socio-economic background of adult individuals. We use a set of 13 surveys that were implemented during a period ranging from 2000 to 2013 and covering the following countries: Comoros, Democratic Republic of Congo, Ghana, Guinea, Madagascar, Malawi, Niger, Nigeria (two waves), Rwanda, Tanzania (two waves), and Uganda (two waves).

Our estimates are shown on a sub-sample of the original data, adult household members with observable relevant characteristics, nevertheless they uncover a dramatic picture. Total consumption inequality is remarkable in all the countries, although quite variable across them: the Gini index ranges from 0.55 for Comoros to 0.31 for Niger, but in general the Gini index is around 0.4 in all countries considered. The entire region of SSA is confirmed as one of the most unequal regions in the world. Moreover, for the three countries for which two waves are available (Nigeria, Tanzania, and Uganda), the results show an increase in inequality in recent years. As far as inequality of opportunity is concerned, our estimates witness that the impact of exogenous circumstances is noticeable in every country, although this impact is quite variable across them: the portion of total inequality which can be attributed to (the observable) exogenous circumstances is between 30% and 40% for the generality of countries considered. This is a striking result, particularly if one considers that the computed measures are only lower bound estimates of the inequality of opportunity level in each country. We also look at the association between total consumption inequality and inequality of opportunity: although some re-rankings do exist, the data show a positive relationship between the two kinds of inequalities.

The sources of unequal opportunities also differ across countries. For example in Comoros and Niger birthplace play a strongest role in determining IOp, while in Congo is clearly ethnicity to be the dominating circumstance.

The ranking of countries in terms of inequality of opportunity is robust with respect to the inequality measure used, but our estimates are sensitive with respect to the estimation approach and to the choice of the exogenous circumstances. In this paper we address this issue by exploring two different estimation approaches (parametric and non-parametric) and by proposing an 'adjusted' inequality of opportunity measure, which takes into account the differences between countries in the number of the circumstances variables. This methodology should make the cross-country comparison more reliable.

Our results differ substantially from the only previous contribution that has focussed on inequality of opportunity in SSA:[3] Cogneau and Mesplé-Somps (2008) analysed five SSA countries (Ivory Coast, Ghana, Guinea, Madagascar, and Uganda) by using data collected between 1985 and 1994. They use a very coarse set of cir-

[3]See also Piraino (2015) for a study of IOp in South Africa.

cumstances (parental background) and, in fact, their results show a much lower level of inequality of opportunity: with some variation between countries, their estimates show that the portion of inequality attributed to exogenous circumstances is between 10% and 20%. Unlike Cogneau and Mesplé-Somps (2008), we extend the analysis to a larger set of countries and a bigger set of circumstances for each country; moreover, we provide a more data-extended and methodologically intensive analysis.

The paper is organised as follows. Section 2 briefly reviews the concept of opportunity inequality and discusses some measurement issues. Section 3 describes the data, and the non-parametric and parametric analyses of inequality of opportunity for the periods and countries considered are presented in Sections 4 and 5, respectively. Section 6 provides a summary of the current findings and concludes with suggestions for further research on IOp in SSA countries.

2 Methodology

2.1 A model of equality of opportunity

The canonical model of EOp assumes that the outcome of an individual, y, is entirely determined by two classes of variables: circumstances and effort (see Roemer, 1998; Van de gaer, 1993; Peragine, 2002). For simplicity, we refer here to the individual outcome as 'income', but any other interpretation of outcome, such as consumption, would in principle be possible. Circumstances are denoted by c and belong to a finite set Ω: examples are gender, age, ethnicity, region of birth, or parental background. These are factors beyond an individual's control but nonetheless exogenously affect income. Effort is denoted by e and belongs to the set Θ, and it may be treated either as a continuous or a discrete variable. This is a factor that endogenously affects the individual income since it is the result of one's own choices. The different forms of luck that may affect the individual income can be classified either as circumstances or as responsibility characteristics. Individual income can then be expressed as follows:

$$y = g(c, e) \tag{1}$$

The production function $g : \Omega \times \Theta \rightarrow \mathbb{R}_+$ is assumed to be monotonic in both arguments, while circumstances and effort are assumed to be orthogonal.[4] This is a reduced form model in which neither the opportunities themselves nor the individual decision process to exert a given level of effort are explicitly modeled. The model builds on the argument that (non-observable) individual opportunities can be inferred by observing joint distributions of circumstances, effort, and income, which fully characterise a population of individuals. For simplicity, let us treat effort, as well as each element of the vector of circumstances, as discrete variables. This would allow the population to be partitioned in two ways: into types in which all individuals share the same circumstances and into tranches in which everyone shares the same degree of effort.

Roughly speaking, the source of unfairness in this model is given by the effect that circumstance variables (which lie beyond individual responsibility) have on individual outcomes. However, there are different ways to measure such effect. This

[4]This assumption is motivated by the theoretical argument that it would be hardly sustainable to hold people responsible for the factor e in a situation in which it were dependent on exogenous characteristics.

measurement exercise can be thought of as a two-step procedure: first, the actual distribution is transformed into a counterfactual distribution that reflects only and fully the unfair inequality, while all the fair inequality is removed. In the second step, a measure of inequality is applied to this counterfactual distribution. The construction of the counterfactual distribution should reflect two distinct and independent principles: the reward principle, which is concerned with the apportion of outcome to effort and, in some of its formulations, requires to respect the outcome inequalities due to effort; the compensation principle, according to which all outcome inequalities due to exogenous circumstances are unfair and should be compensated for by society. In particular, the existing literature has developed two main versions of the compensation principle and two consequent approaches to the measurement of opportunity inequality, namely the ex-ante and the ex-post approach.

According to the ex-ante approach, there is equality of opportunity if the set of opportunities is the same for all individuals, regardless of their circumstances. Hence in the ex-ante version, the compensation principle is formulated with respect to individual opportunity sets: it requires reducing the inequality between these opportunity sets. In the model introduced above, the income distribution of a given type is interpreted as the opportunity set of all individuals with same set of circumstances. Hence, the focus is on the inequality between-types: the counterfactual distribution should eliminate the inequality within the types (reward) and reflect the inequality between the types (ex-ante compensation). Let us underline here a dual interpretation of the types in the EOp model: on one hand, the type is a component of a model that, starting from a multivariate distribution of income and circumstances, allows us to obtain a distribution of (the value of) opportunity sets enjoyed by each individual in the population. On the other hand, given the nature of the circumstances typically observed and used in empirical applications, the partition into types may be of interest *per se*: they can often identify well-defined socio-economic groups, possibly deserving special attention by policymakers.

Alternatively, according to the ex-post approach, there is equality of opportunity if and only if all those who exert the same effort end up with the same outcome. The compensation principle, in the ex-post version, is thus defined with respect to individuals with the same effort but different outcomes. This means that opportunity inequality within this approach is measured as inequality within the tranches. Hence, the corresponding counterfactual distribution should reflect the inequality within the tranches (ex-post compensation) but should eliminate the inequality between the tranches (reward).

Different measures, which are either consistent with the ex-ante or the ex-post approaches, have been proposed in the literature (see Ferreira and Peragine, 2015; Ramos and Van de gaer, 2015): they express different and sometimes conflicting views on equality of opportunity and in fact the rankings they generate may be different. In addition, their informational requirements are quite different: while for the ex-ante approach one needs to observe the individual outcome and the set of circumstances, for the ex-post approach a measure of individual effort is required. Therefore, in addition to normative considerations, the choice of which methodology to adopt should also reflect data availability. In our case, the database we use does not contain a satisfactory measure of effort. For this reason, we focus on the ex-ante approach and, among the various measures coherent with such approach, we use the between-types inequality measure, which was proposed, among others, by Peragine (2002), Checchi and Peragine (2010), and Ferreira and Gignoux (2011). It relies on a counterfactual distribution, which is obtained by replacing each individual's income by the average income of the type an individual belongs to, independently from the level of effort exerted.[5] This smoothing transformation, intended to remove all inequality within types, can be performed by using either a parametric or a non-parametric method. These are discussed in the following section.

2.2 Parametric and non-parametric approaches

Given a distribution of income Y of size N, with n types indexed by $i = 1, ..., n$, for each type i the population size will be denoted by m_i , its population share by q_i, and its mean income by $\mu_i(y)$. According to the between-types inequality measure, the counterfactual distribution Y_s is obtained by replacing each individual income with the value of the opportunity set of that individual, that is, the mean income of the type to which the individual belongs. Hence, by ordering the types on the basis of their mean such that $\mu_1(y) \leq ... \leq \mu_j(y)... \leq \mu_n(y)$, the counterfactual distribution corresponding to Y is defined as $Y_s = (\mu_1(y)\mathbf{1}_1, ..., \mu_i(y)\mathbf{1}_i, ..., \mu_n(y)\mathbf{1}_n)$, where $\mathbf{1}_i$ is the unit vector of size m_i. For a given measure of inequality $I : R_+^N \rightarrow R_+$, the part of inequality due to initial circumstances will be given by $I(Y_s)$ or in relative terms

[5]The use of the average of the type for the smoothing transformation is justified, from a normative point of view, in light of the utilitarian reward principle, according to which society should express full neutrality with respect to inequalities due to effort. See Ferreira and Peragine (2015) for a discussion of the different formulations of the reward principle proposed in the literature and Lefranc et al. (2009) and Peragine and Serlenga (2008) for empirical analyses based on different versions of the reward principle.

by:

$$\text{IOp} = \frac{I(Y_s)}{I(Y)} \tag{2}$$

Equation (2) measures the portion of overall inequality that can be attributed to unequal opportunities. In most empirical analyses $I(Y)$ is represented by the mean logarithmic deviation (MLD), because it is perfectly decomposable in between- and within-types inequality. However, the MLD has some undesirable properties: in particular, it tends to be more sensitive to extreme values and is not bounded above; therefore, when inequality is measured on a distribution of type means, from which extreme values have been removed by the smoothing operation, it tends to be very underestimated by the MLD. For these reasons, in this paper we follow Aaberge et al. (2011) and use the Gini index,[6] which has well known desirable characteristics, although it is not perfectly decomposable in between- and within-types inequality whenever the type income distributions overlap.[7] Therefore, in general:

$$\text{Gini}(Y) = \text{Gini}(Y_{within}) + \text{Gini}(Y_s) + \text{K} \tag{3}$$

Where K is a residual greater than zero when there is overlapping between the types distributions.

It deserves to be noted that, given a set of selected circumstances defined on the basis of normative grounds and observability constraints, any within-type variation in individual outcome is attributed to personal effort. However, the vector of observed

[6]In the Appendix, for a robustness check we also compute the mean logarithmic deviation.

[7]See Lambert and Aronson (1993) for an insightful discussion on the Gini decomposition.

circumstances is likely to be a sub-vector of the theoretical ('true') vector of all possible circumstances that determine a person's outcome. Hence, as in any other empirical analysis of this kind, we face the issue of omitted circumstance variables. This problem is often addressed by the argument that the IOp estimates should be interpreted as lower-bound estimators of true inequality of opportunity, that is, the inequality that would be captured by observing the full vector of circumstances. It can be shown, in fact, that increasing the number of observed circumstances increases IOp (see Ferreira and Gignoux, 2011; Luongo, 2011).

However, this interpretation renders IOp estimates barely comparable across studies, particularly when comparing, for instance, the IOp of a country with a large number of observable circumstances to the IOp of another country with only few observable circumstances. Moreover, the error made in comparing these two quantities might not be random but correlated with data quality. Elbers et al. (2008) discuss this issue in a more general setting concerning any estimate of between-group inequality. They claim that, when decomposing total inequality into a between and a within component, the estimate of between-group inequality might be artificially too low because it compares between-group inequalities with the inequality measured in a counterfactual population in which each individual is a group. To overcome this problem they propose an adjusted measure of between-group inequality, which is equivalent to the actual between-group inequality normalized by the maximum possible between-group inequality that could be reached in the population, given the number of groups. The latter is defined as the extent of between-group inequality in a counterfactual distribution (Y_a) obtained by ranking outcomes from the lowest to the richest and then partitioning the distribution in such a way that the groups have same population share as the actual group. Hence, adjusted IOp (Adj-IOp) can be expressed as follows:

$$\text{Adj-IOp} = \frac{I(Y_s)}{I(Y_a)} \tag{4}$$

Although the problem they are looking at does not exactly correspond to our problem of partial observability, their solution can be usefully applied to this context. This adjusted measure is appealing as it accounts for the number of types and their relative weights. Adj-IOp solves, at least in part, the problem of comparing IOp estimates based on different number of observable characteristics. Therefore, in the following we propose estimates of both IOp and Adj-IOp.

The non-parametric approach discussed so far is data-intensive: as the partition into types becomes finer, the population size of each type decreases, bringing about a decline in the precision of the estimates of the type mean, consequently giving rise to a bias in the estimation of IOp. In countries such as those considered in this paper, where data limitation on circumstances might seriously hamper the analysis, an alternative, parametric approach to the estimation of inequality of opportunity that economises on data requirement could be explored. It is based on the assumption that a simple linear relationship characterises equation (1), given that circumstances are exogenous by definition and they may also influence effort (see Bourguignon et al., 2007; Ferreira and Gignoux, 2011; Ferreira et al., 2011). Therefore, equation (1) could be re-expressed in reduced form as: $y = \phi(c, \epsilon)$, and a linearised version of this equation would lead to:

$$y = \beta c + \epsilon \tag{5}$$

The estimated $\hat{\beta}$ coefficient of the ordinary least square (OLS) estimation of equation (5) will incorporate both the direct effect of circumstances on outcome and its indirect effect through effort. Clearly, this will be true only if the $\hat{\beta}s$ estimated with OLS are unbiased estimates of the real effect of circumstances. Inequality of opportunity will then be obtained by applying an index of inequality to the distribution of the predicted values $\hat{y}$ from the OLS estimation of equation (5), that is: [8]

$$\text{IOp} = I(\hat{y}) \tag{6}$$

Relative inequality of opportunity will be equal to:

$$\text{IOp} = \frac{I(\hat{y})}{I(y)} \tag{7}$$

It is worth noticing that the parametric approach is fully consistent with the ex-ante utilitarian assumption used in the non-parametric modeling of IOp. Here, the only difference is that the expected outcome, given observable circumstances, is obtained using the predicted values from a linearised OLS model. Assuming a linear effect of circumstances we no longer need to construct types in order to predict this outcome, and we can exploit all information contained in the variables describing circumstances, that is, all values assumed by each circumstance.

The literature has recognised that also the parametric approach has some limita-

[8]See Ferreira and Gignoux (2011).

tions, however. First, it indirectly imposes a precise functional form linking circumstances and outcome. Moreover, the OLS estimation of equation (5) requires that one controls for a number of dummy variables. In fact, the set of circumstances that is generally used in empirical analyses typically includes parental education, parental occupation, area of birth, and ethnicity. Such variables are not cardinal, and each one needs to be transformed into a number of dummy variables equal to the number of values it assumes in order to make equation (5) operational.

When no cardinal circumstance is observable, estimating equation (5) through an OLS regression brings to the estimation of a shift in the regression intercept associated to each category of every circumstance, for instance, having white collar parents or being a first generation immigrant. This implies a severe restriction in the construction of the counterfactual distribution, because it imposes a fixed effect for each circumstance. For example, it could be the case that being a first generation immigrant has a completely different meaning depending on whether one's parents are university professors or construction workers. On the other hand, in a parametric approach this effect is defined to be the same. To take into account the interaction between circumstances, one needs to interact dummies. However, once all dummies have interacted, one intercept is estimated for each type, and our OLS estimate becomes equivalent to the non-parametric approach.

Thus, from one side the motivations for the use of a parametric approach appear to be clear when cardinal measures of circumstances are available (such as parental income). However, they are less convincing when all circumstances can be only modeled through dummies.

From the other side, by assuming a fixed effect of each circumstance, the output of the parametric approach can be easily used to estimate the partial effect of each circumstance on outcome: for example, following Wendelspiess and Soloaga (2014), we could implement a Shapley-Shorrocks decomposition based on the average marginal effect of each circumstance over all their possible permutations. This method leads to a path-independent identification of the contributions of each circumstance.

The discussion above suggests the adoption of both the parametric and non-parametric approaches. In fact, as underlined by Ferreira and Gignoux (2011), the two approaches may be considered complementary.

3 Data

Our analysis is based on the following surveys:

- *Enquête Intégrale auprè des Ménages* (EIM) for Comoros, carried out by the Statistical Office of the Ministry of Land Planning and Settlement;

- Enquête sur la Consommation des Ménages (ECM) for Congo D. R. (year 2010), carried out by National Institute of Statistics (Ministry of Planning);

- *Ghana Living Standards Survey* (GLSS) for Ghana (year 2013), carried out by the Ghana Statistical Service - National Data Archive (GSS);

- *Enquête Integré de Base pour l'Evaluation de la Pauvreté* (EIBEP) for Guinea (year 2003), carried out by the National Directorate of Statistics (Ministry of Economics and Finance)

- *Enquête Périodique auprè des Ménages* (EPM) for Madagascar (year 2005), carried out by the National Institute of Statistics (INSTAT);

- *Third Integrated Household Survey* (IHS3) for Malawi (year 2010), carried out by the National Statistical Office of Malawi;

- *National Survey on Household Living Conditions and Agriculture* (ECVM) for Niger (year 2011-12), carried out by the National Institute of Statistics of Niger;

- *General Household Survey* (GHS) for Nigeria (years 2010-11 and 2012-13), carried out by the National Bureau of Statistics of Nigeria;

- *Enquête Intégrale sur les Conditions de Vie des Ménages* (EICV) for Rwanda (year 2000), carried out by National Institute of Statistics of Rwanda (NISR);

- *National Panel Survey* (NPS) for Tanzania (years 2009-10 and 2010-22), carried out by the National Bureau of Statistics of Tanzania;

- *Uganda National Panel Survey* (UNPS) for Uganda (years 2009-10 and 2010-11), carried out by the Uganda Bureau of Statistics.

They are all representative at a national level and cover both urban and rural areas. Table 1 lists the surveys used, the year they refer to, their original sample size, and a link to the documentation. Our analysis is based on a sub-sample of the original data, obtained by considering only individuals aged 15 years or more for whom information about circumstances beyond individual control are available. The outcome considered is per capita consumption, which encompasses consumption for both food and non-food goods, that is, we assume a proportional intra-household

Table 1: Data sources

country	*survey*	*year*	*sample size*	documentation
Comoros	EIM	2004	18,373	IHSN
Congo D. R.	ECM	2010	110,529	IHSN
Ghana	GLSS	2013	39,826	GSS
Guinea	EIBEP	2003	25,319	INSG
Madagascar	EPM	2005	30,271	INSTAT
Malawi	IHS3	2010	30,137	World Bank
Niger	ECVM	2011-12	12,118	World Bank
Nigeria	GHS	2010-11	14,916	World Bank
Nigeria	GHS	2012-13	14,560	World Bank
Rwanda	EICV	2000	17,69	INSR
Tanzania	NPS	2009-10	9,175	World Bank
Tanzania	NPS	2010-11	11,394	World Bank
Uganda	UNPS	2009-10	8,268	World Bank
Uganda	UNPS	2010-11	7,509	World Bank

distribution of consumption and zero economies of scales in consumption. Although we use different surveys, the results are comparable across countries since the consumption variable has been adjusted for inflation and translated into 2011 purchasing power parity (PPP) international dollars (World Bank, 2015).

A fundamental step in the measurement of inequality of opportunity is the identification of the vector of observable circumstances. This is a normative choice, subject to the constraint of data availability. Our data contain information on a small set of basic circumstances, but nonetheless of prominent importance. For each country, in fact, we can observe a subset of the following: ethnicity, parental education and occupation, birthplace (see Table 2 for details).

As for the specific circumstances, parental education and occupation are widely used as circumstances in the empirical literature on IOp that has dealt with developed countries. The importance of the socio-economic origin is emphasised also by the sociological literature on social stratification and social mobility, which focusses on occupation-based social classes. A vast amount of evidence has been produced on the effect of socio-economic background on children's outcomes during adulthood. This literature is however traditionally Western-centric and has rarely concentrated on SSA countries. Nevertheless, there is also evidence supporting the argument that parental education and occupation act as circumstances on individual outcome in the specific SSA context. For instance, it has been shown that, in these countries, the nutritional status of a child is strongly correlated to parental occupation with obvious, although indirect, consequences on his outcome in the future (Madise et al., 1999). Parental education, instead, has been shown to be an important factor in determining whether or not a child is currently attending school; whereas, school improvements in parental education have been shown to increase the schooling of children, which, in addition to improving their health and reducing the status of extreme poverty, has direct effects on the outcome prospects of these children (see, among others, Glick and Sahn, 2000; Lloyd and Blanc, 1996; Lassibille and Tan, 2005; Schultz, 2004).

Ethnicity and birthplace are variables of paramount importance in SSA, historically characterised by civil and ethnic conflicts, which arrest or even reverse the growth and development process of the this specific part of the African continent. Even today, SSA countries face impressive challenges to peace and stability and have fallen prey to continuous armed ethnic conflicts. Between 1946 and 2002, not less than 1.37 million battle-related deaths occurred in 47 civil wars in SSA (Lacina and

Table 2: Circumstances observed by country

country	circumstances			
	birthplace	parental education	parental occupation	ethnicity
Comoros	√	√	√	
Congo DR	√	√	√	
Ghana	√	√		
Guinea	√	√	√	
Madagascar	√	√		√
Malawi	√	√		√
Niger	√			√
Nigeria		√	√	
Rwanda	√	√	√	
Tanzania	√	√		
Uganda	√			√

Note: Ethnicity for Democratic Republic of Congo is observable but the documentation to decode it is missing thus rendering it impossible to construct the partition in types necessary for the non-parametric estimates of IOp. In Malawi, mother tongue is used as a proxy for ethnicity.

Source: Surveys listed in Table 1.

Gleditsch, 2005). In 2011, for instance, SSA has had 91 instances of this type of conflict, compared to the 89 of 2010 (see Brautigam and Knack, 2014; De Ree and Nillesen, 2009). Moreover, previous studies have shown that high levels of ethnic diversity are strongly linked to high informal market premiums, poor financial development, low provision of infrastructure, and low levels of education. Ethnicity has a strong influence on inequality in Africa where ethnic fractionalisation has given rise to a political economy of unequal subsidies and discrimination (Easterly and Levine, 1997; Milanovic, 2003). The area is also characterised by regional disparities in access to opportunities. Hence, it appears natural to treat ethnicity and birthplace as circumstances in the context of our analysis.

It is important to note that cross-country comparisons of IOp must be interpreted while bearing in mind that the subset of circumstances used may vary across countries, as different surveys usually collect different information on circumstances.[9]

In order to provide meaningful non-parametric estimates of IOp, the circumstances observed for each country require some additional treatment. While the parametric approach, assuming a linear effect of circumstances on outcome, can exploit all the information contained in the variables that describe circumstances, the non-parametric approach is forced to aggregate some of this information. Thus, to estimate the mean of each type with a sufficient degree of confidence, the sample size of each type should not be too small. Circumstances are therefore aggregated to reduce the number of types and to increase their size. Tables 7 to 16 in the Appendix contain the details of the partition in types used for the non-parametric estimates in each country. These tables represents the so-called 'opportunity profile' (Ferreira and Gignoux, 2011), a country specific list of types, their rank, and the value of their opportunity set. These profiles are interesting *per se*, as they identify the most deprived groups in each society.

4 Results: the non-parametric approach

4.1 Consumption inequality and opportunity inequality

Table 3 reports, for each country and wave, the estimates of total inequality, inequality of opportunity, and inequality of opportunity ratio (all computed by using

[9] All individuals with missing information on the circumstances are dropped from the analysis.

Table 3: Inequality and IOp, non-parametric estimates

country	*sample*	*consumption per capita*	*types*	*inequality Gini*	*IOp Gini*	*IOp % Gini*	*max between-groups Gini*	*Adj-IOp % Gini*
Comoros	5,936	2,975	36	0.5532	0.1657	29.95	0.5489	30.18
Ghana	42,519	1,838	24	0.4143	0.1457	35.16	0.3960	36.80
Guinea	24,866	1,000	32	0.4275	0.1594	37.28	0.4257	37.44
Madagascar	28,951	415	30	0.3701	0.1026	27.71	0.3680	27.87
Malawi	30,137	855	64	0.4739	0.2071	43.71	0.4734	43.75
Niger	11,774	1,071	48	0.3106	0.1022	32.91	0.3087	33.11
Nigeria 2010-11	14,916	1,298	20	0.3885	0.1459	37.54	0.3792	38.47
Nigeria 2012-13	14,560	1,601	20	0.3897	0.1429	36.66	0.3795	37.65
Rwanda	14,112	641	24	0.4436	0.1149	25.90	0.4385	26.20
Tanzania 2009-10	9,119	1,133	52	0.3935	0.1687	42.88	0.3930	42.93
Tanzania 2010-11	11,391	1,112	52	0.3966	0.1609	40.57	0.3961	40.62
Uganda 2009-10	8,194	1,157	24	0.4523	0.1785	39.46	0.4470	39.93
Uganda 2010-11	7,454	1,039	24	0.4748	0.1885	39.71	0.4691	40.19

Note: Per capita consumption is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

the Gini index). Moreover, the first three columns contain information about the sample size, the average per capita consumption, and the number of types in which each country is partitioned. Note also that we do not report any estimate for Congo DR due to the already mentioned impossibility of aggregating ethnic groups to obtain types.

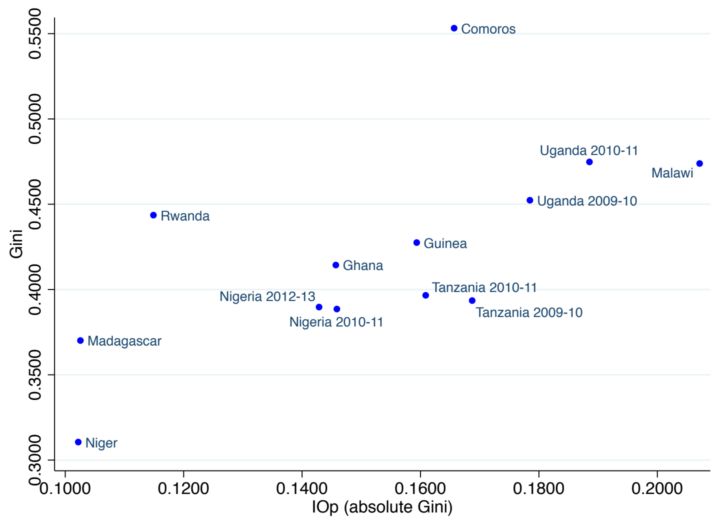
Total inequality is remarkable in all the countries, although quite variable across them: the Gini index ranges from 0.55 for Comoros to 0.31 for Niger, but in general the Gini is around 0.40. The entire region of SSA is confirmed as one of the most unequal regions in the world. For the three countries for which observations for more than one year are available (Nigeria, Tanzania, and Uganda) the results bear witness to an increase in inequality: hence, the recent dynamics, where available, show a regressive pattern. The ranking of countries according to their level of inequality seems to be robust to the choice of the inequality measure (whether the Gini or the MLD index, the latter reported in Appendix III table 16), there is in fact only one instance of re-ranking occurring between Tanzania and Niger.

As far as the inequality of opportunity is concerned, the estimates show an equally dramatic albeit different picture. The share of inequality that can be attributed to different exogenous factors is extremely high and variable across all countries: it ranges between 26% for Rwanda and 44% for Malawi, and is more generally between 30% and 40% for the other SSA countries. In other words, according to the observed circumstances, more than one third of the observed inequalities in consumption can be attributed to exogenous factors, that is, to inequality of opportunity. This is a striking result, particularly if one considers that the computed measures are only lower bound estimates of the inequality of opportunity level in each country.

It is also interesting to look at the association between total consumption inequality and opportunity inequality as depicted in Figure 1. This figure could be interpreted as a generalisation of the so-called "Great Gatsby" curve (Corak, 2013), showing a negative relationship between income inequality and social mobility. Our results show that, although countries with higher consumption inequality are also characterised by a higher level (portion) of inequality of opportunity, there is also considerable re-ranking between countries taking place in passing from total inequality to IOp. Notable here is the case of Comoros, which has the highest total inequality but it has the second to lowest IOp of all countries examined here.

In sum, our estimates allow to divide the 10 SSA countries under analysis into three main groups. The first group is represented by the three countries with highest

Figure 1: Total inequality and inequality of opportunity



Source: Surveys listed in Table 1.

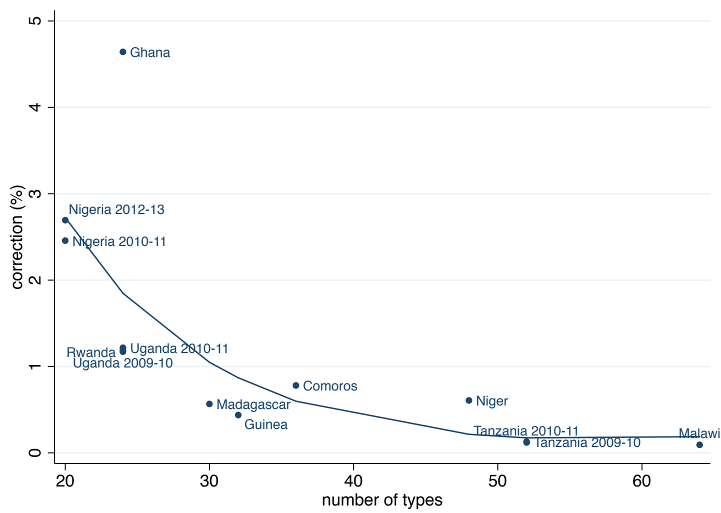
share of IOp, namely Tanzania, Malawi, and Uganda; Malawi and Uganda have also highest level of total inequality. The second group is represented by the three countries with lower share of inequality of opportunity, namely Rwanda, Madagascar, and Comoros, that nevertheless exhibit a comparatively high level of consumption inequality; the third represented by all the other countries having relatively middle shares of IOp (i.e., Ghana, Guinea, Niger, and Nigeria).

4.2 'Adjusted' inequality of opportunity

The last two columns of the third part of Table 3 report, respectively, the adjusted IOp according to the Gini index and it share on total consumption inequality. As discussed above, the normalisation of inequality with respect to the number of types is particularly relevant in the present context, as we are comparing IOp in countries whose specific consumption distribution is partitioned into a very different number of types: from a minimum of 20 in Nigeria to a maximum of 64 in Malawi.

Figure 2 plots the difference between IOp and Adj-IOp as a percentage of IOp against the number of types. Figure 2 also shows a clear pattern for this correction (approximated with a fractional polynomial curve), approaching zero as the number of types increases.

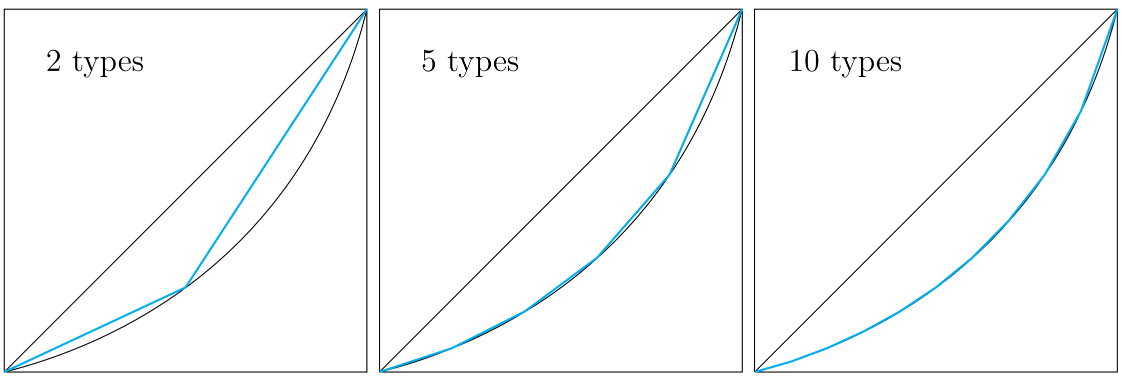
Figure 2: Adj-IOp correction and number of types



Source: Surveys listed in Table 1.

The figure makes it clear that the adjustment procedure does not add particularly relevant information in our context. The correction is never above 5% and it is smaller than 2% for countries with a number of types above 40. Hence, the higher the number of types the lower the impact of the adjustment, and this result is rather general. To grasp this drawback consider Figure 3, plotting the difference between total Gini, twice the area between the black Lorenz curve and the diagonal, and the maximum between-group Gini, twice the area between the blue broken line, for three hypothetical group partitions: one group, five groups, ten groups. The difference between the two possible denominators of IOp will depend on the shape of the original Lorenz curve; the example clarifies that this difference approaches zero very quickly as the number of types increases. Therefore, the adjustment proposed by Elbers et al. (2008) loses relevance whenever the number of types is in the order of tens.

Figure 3: Lorenz curve and maximum between type inequality Lorenz curve



Note: Lorenz curves for the maximum between-group inequality (light blue) are drawn assuming a population partitioned into equally sized types.

5 The parametric approach

Table 4 reports, for each country and wave considered, the results of the parametric estimates of IOp. The first part of the table contains information about sample size, mean per capita consumption, and number of regressors (all dummies) used to assess the share of total inequality explained by circumstances. The number of regressors is given by the number of observable circumstances multiplied by the number of values that each circumstance can take.[10] The second part of the table contains the estimates of total inequality, IOp in absolute terms and as share of total inequality, using the Gini coefficient.[11]

Table 4: Inequality and IOp, parametric estimates

country	*sample*	*consumption per capita*	*number of regressors*	*total inequality Gini*	*IOp Gini*	*IOp (%) Gini*
Comoros	5,936	2,975	91	0.5532	0.2305	41.67
Congo DR	39,578	1,535	402	0.3634	0.1739	47.84
Ghana	42,519	1,838	125	0.4143	0.2304	55.61
Guinea	24,866	1,000	96	0.4275	0.1504	35.18
Madagascar	28,951	415	445	0.3701	0.2080	56.22
Malawi	30,137	855	71	0.4739	0.2637	55.64
Niger	11,774	1,071	50	0.3106	0.1249	40.22
Nigeria 2010-11	14,916	1,298	40	0.3885	0.1640	42.21
Nigeria 2012-13	14,560	1,601	40	0.3897	0.1661	42.62
Rwanda	14,112	641	76	0.4436	0.1851	41.74
Tanzania 2009-10	9,119	1,133	41	0.3935	0.1911	48.57
Tanzania 2010-11	11,391	1,112	40	0.3966	0.1820	45.90
Uganda 2009-10	8,194	1,157	100	0.4523	0.2159	47.74
Uganda 2010-11	7,454	1,039	102	0.4748	0.2497	52.58

Note: Per capita consumption is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

[10]The analytical results of the OLS regression for each country are available from the authors upon request.

[11]See Appendix III for a parametric estimate of IOp using MLD.

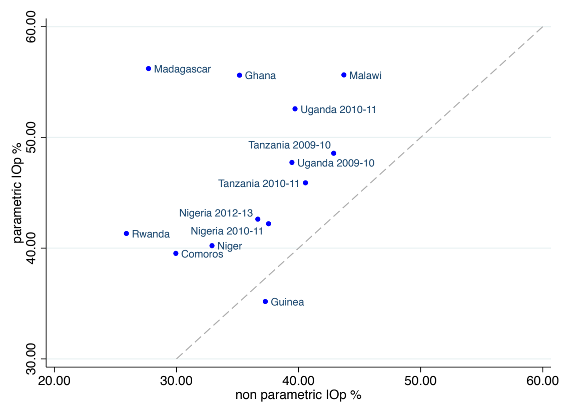
In general, non-parametric estimates tend to be lower than their parametric version; however, this must not be necessarily the case. Recall that parametric and non-parametric approaches differ in two aspects: the former imposes a linear relationship between circumstances and outcome, the latter aggregates some information contained in variables beyond individual control. Setting aside the problem of partial observability, both constraints imply that IOp is a downward bias estimate of the real IOp under very general conditions. Imposing linearity reduces the variability that can be explained by circumstances in all cases except when y is a linear function of c. Similarly, ignoring some of the circumstances' variability decreases the ability of these variables to explain total inequality, unless the inequality between the groups aggregated is zero.

Thus, when the bias implied by the assumption of linearity is smaller than the bias introduced aggregating circumstances, the parametric IOp is larger than the non-parametric IOp. However, there can be cases in which the linearity assumption implies a larger distortion than the aggregation of circumstances: in this case the non-parametric IOp will be larger.

Figure 4 shows the discrepancy between the two approaches. Parametric estimates are reported on the vertical axis and the non-parametric estimates on the horizontal axis. The first clear feature that stands out is that, with the exception of Guinea, parametric estimates are always larger than non-parametric ones. As for the ranking, with the considerable exception of Guinea, Ghana, and Madagascar, there is a clear positive relationship between the rankings generated by the two approaches.

The discrepancy between the two approaches seems to be driven by the very high number of regressors used to estimate equation (5) and the rather low number of types used to construct the counterfactual distribution for the non-parametric estimates. An extreme case is that of Madagascar, in which the number of regressors is the highest, 462, while the number of types is 30, one of the lowest. Moreover, Madagascar jumps from being one of the least unequal countries when IOp is parametrically estimated to being one of the most unequal when IOp is non-parametrically estimated. Such a difference should be expected whenever the number of regressors (which by definition increases the total variability explained) is much larger than the number of types.

Figure 4: IOp parametric and non-parametric estimates

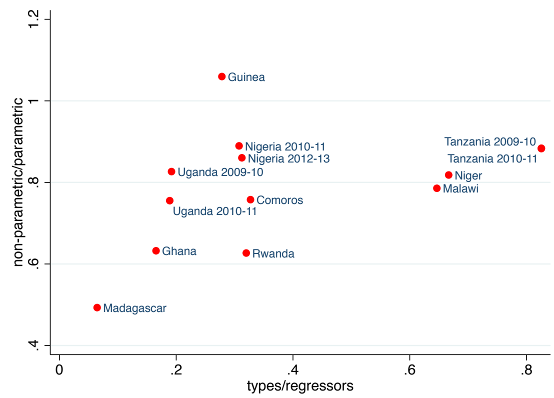


Source: Surveys listed in Table 1.

However, it must be noted that the high number of regressors in Madagascar is mainly due to the high number of possible birthplace, that is, 397 dummies, far more than the six provinces in which Madagascar was divided at the time of the survey (now 22 regions), and also more than three times the 111 districts of the country. Birthplace in this survey are cities (commune urbaines). Not surprisingly the coefficients for the dummies of such a detailed subdivision of the territory are generally not statistically significant. It seems, therefore, unreasonable to include all the possible birthplaces among the controls of the OLS estimation of equation (5), since the estimates of their effect on circumstances would not be reliable. A viable solution consists of aggregating birthplaces into districts or provinces. Indeed this is exactly what we do with the non-parametric approach: we trade-off some of our regressors' variability with statistical significance. Therefore, in cases like that of Madagascar, with few observable qualitative characteristics that can take a large number of values, it would be more advisable to follow a non-parametric approach, which has the additional quality of not imposing linearity, rather than a parametric one.

This issue is examined for all countries in Figure 5 where we determine whether the difference between parametric and non-parametric estimates is really due to the difference between the number of types and the number of regressors. The vertical axis reports the ratio between the two estimates (non-parametric over parametric), and the horizontal axis reports the ratio between the number of types and the number of regressors. Indeed, the positive correlation between the two ratios suggests that the number of regressors does play a role in making parametric estimates. Obviously, the correlation is far from perfect, and Guinea is an interesting case. Although for this country we have 113 regressors and 32 types, the parametric estimate of IOp is smaller than the non-parametric one. The case of Guinea provides an example of how assuming a linear effect of circumstances on outcome actually provokes a downward bias of our IOp estimates, which is larger than the bias induced by aggregating circumstances when using the non-parametric approach.

Figure 5: Number of types and number of regressors



Source: Surveys listed in Table 1.

The literature has traditionally judged the assumption of linearity to be less important in determining the magnitude of IOp than the issue related to the number of circumstances. However, the case of Guinea clearly highlights that there are cases in which the opposite can happen. Table 5, a simplified version of the opportunity profile presented in Table 9 in Appendix II, clarifies this point. The effect of parental occupation on children outcome depends on area of birth: on average, in Guinea, having a father employed in agriculture is associated with low consumption. By contrast, being born in the region of Labe to parents working in the agricultural sector implies that one belongs to the type with the best outcome prospects. The effect of birthplace and parental occupation on consumption are clearly not linear. This is not just a statistical feature, but it has a clear economic meaning: the Labe re-

Table 5: Non-linear impact of circumstances: the case of Guinea

birthplace	parental occupation	per capita consumption
rest of Guinea	agriculture	843.10
rest of Guinea	other	1,117.60
Labe	other	1,272.88
Labe	agriculture	1,805.57

Note: This example is obtained by aggregating data in Table 9.

gion has recently experienced prolonged periods of substantial surplus in agricultural production and is one of several comprising the centre of national and international agricultural trade flows (FEWS, 2013). Therefore, an individual who was born into a farming family in Labe has the best possible condition in terms of economic opportunities. It is clear that for the specific case of Guinea, the parametric procedure neglects the interaction between parental occupation and area of birth.

In sum, among the main reasons for the possible inconsistency between the parametric and non-parametric approaches, we find that the small number of observable characteristics and the possible high number of values they can assume do play an important role. In fact, our results demonstrate that a high number of regressors tends to make parametric estimates higher than the non-parametric estimates. However, the assumption of a linear effect of circumstances on outcome, implicit in the parametric approach, can provoke a downward bias of IOp.

5.1 Detecting the contribution of the specific circumstances

In this section we analyse the degree of association between each circumstance and the level of individual consumption, in order to shed some light on the relative importance of the different circumstances in determining inequality of opportunity. We are aware that this analysis does not identify the causal effect of each circumstance to IOp, unobservable determinants of the individual outcome are likely to be correlated with the observable circumstances preventing a causal identification (see Ferreira and Gignoux (2011) for a discussion). Nevertheless, we believe that the description of the different degrees of association may help to provide an interpretative framework for our estimates of IOp across the SSA countries considered. In doing so we are going beyond the distinction between inequality of opportunity and other inequalities; we are decomposing IOp by source. The same level of IOp can have rather distinct meaning depending on the relative importance of different circumstances. A country with a predominant role for ethnic inequalities may appear very different from a country in which the main channel of transmission of wealth is parental education or occupation.

In this analysis, an important aspect to consider is how circumstances are coded. As we already discussed ordinal and categorical variables are recoded by a set of dummy variables (as many as the number of values the circumstance assumes). Now, *ceteris paribus*, the share of total IOp explained by a circumstance is an increasing function of the number of dummies used to measure it. If, like in Madagascar, birth location assumes 397 possible values, this circumstance is likely to explain more variability than the circumstance 'ethnicity', captured in the same country by 24 dummies. Therefore, when looking at the partial effect of a circumstance, one should consider both the share of total IOp "explained" by that circumstance and the number of variables used to describe it.

Table 6 contains the share of IOp explained by each circumstance (left) and number of variables used to explain each circumstance (right). The partial IOp of each circumstance is obtained by applying the Shapley value decomposition of the between-type Gini coefficient.[12]

Birth location appears to be the most important factor explaining IOp in Niger and in Comoros. At the time of the survey - 2004 - Comoros islands were in a period of relative stability after Colonel Azali Assoumani was elected President in 2002 and agreement about the federal institutional framework was reached. However, one must bear in mind that these islands have been historically characterised by political instability originated from conflicts between islands. On the one hand the political power is traditionally concentrated in Grand Comore (Ngazidja), on the other Anjouan - the second largest island - has the largest port of the archipelago which has brought economic development and power to the island (IMF, 2006). Birth region has a very small role in Malawi there 42 of the 71 regressors explain only 29% of total IOp.

A point is worth emphasising here: birthplace is a relatively easy piece of information to collect and often refers to a very detailed partition of the national territory. Of course, this variable correlates with a number of other variables such as ethnicity, mother tongue, and family wealth. Therefore, in the interpretation of this decomposition, one should bear in mind that any inequality due to omitted circumstances that correlates with observable circumstances is captured by the latter. This could be the case of Rwanda, where ethnicity is not observed and inequality between 18

[12]For this analysis we rely on the algorithm proposed by Araar and Duclos (2009).

regions of birth explains 29% of total IOp.

Ethnicity is observed in Congo, Ghana, Madagascar, Malawi, and Niger. Congo is one of the country in the world with the highest level of ethnic diversity (Goren, 2014). Its 374 recorded ethnic groups explain a large share of total IOp (68%). In Uganda, ethnicity consistently explains nearly 50% of total IOp. Although inequality between ethnic groups is much smaller in Madagascar - only 17% of total IOp - it is explained by inequality between only 24 ethnic groups.

Parental occupation and education explain half of total IOp in Nigeria. The role of parental education appears also important in Guinea (47%) and Rwanda (36%). This is consistent with what observed in the opportunity profiles of the two countries (see Appendix II): in both cases individuals in types with parents employed in the agricultural sector crowd the lowest position of the ranking. Parental occupation has an important role also in Ghana (35%) and Congo (20%). In the case of Congo 20% of total IOp is explained by only 9 dummies describing parental occupation.

Parental education is the most frequently observed circumstance (it is missing only for Uganda and Niger). Its role is heterogeneous with a share ranging from 57% in Tanzania to 5% in Comoros. Finally, in Malawi parental education explains a high share of IOp (52%); a share that is explained by only 14 of the 71 dummies.

Table 6: Shapley value decomposition of IOp

	Share of IOp (%)				regressors				
country	*Birth location*	*Ethnicity*	*Parental occ.*	*Parental edu.*	*Birth location*	*Ethnicity*	*Parental occ.*	*Parental edu.*	*Total*
Comoros 2004	84.19		10.76	5.06	55	0	23	12	90
Congo 2012		67.88	20.07	12.05	0	374	9	17	400
Ghana 2013		40.75	34.89	24.36	0	63	18	42	123
Guinea 2003	38.22		46.68	15.10	44	0	18	35	97
Madagascar 2005	59.16	17.28		23.56	397	24	0	24	445
Malawi 2010	29.34	18.19		52.47	42	15	0	14	71
Niger 2011-2012	87.36	12.64			40	10	0	0	50
Nigeria 2011-2012			53.86	46.14	0	0	22	18	40
Nigeria 2012-2013			52.64	47.36	0	0	22	18	40
Rwanda 2000	29.03		36.14	34.83	18	0	20	38	76
Tanzania 2009-2010	49.39			50.61	0	0	27	13	40
Tanzania 2010-2011	43.46			56.54	0	0	27	13	40
Uganda 2009-2010	52.44	47.56			57	43	0	0	100
Uganda 2010-2011	50.57	49.43			57	43	0	0	100

The share of regressors is the number of regressors that describe the circumstances divided by the total number of regressors *Source: Authors' calculation based on surveys listed in Table 1.*

6 Conclusion

Inequality in SSA countries is generated by many factors. The area of birth, ethnicity or the education level of the parents are, for instance, among the most important factors. Inequality of opportunity, that is, the extent to which these kinds of factors determine the outcome of individuals in adulthood, contribute to increase overall inequality, and violate principles of fairness. Although the empirical literature on IOp measurement has proliferated in the last decades, there are very few contributions that focus on inequality of opportunity in SSA countries. The lack of estimates for this part of the world is mainly due to the lack of reliable data on individual outcome and circumstances.

This paper has utilised 13 reliable household consumption surveys to assess IOp in 11 SSA countries. All information about exogenous factors provided by these surveys have been used. These encompass information on region of birth, parental education and occupation, and ethnicity. We have complemented the analysis by estimating the partial effect of each circumstance in determining IOp and the adjusted measure of IOp proposed by Elbers et al. (2008).

Overall, inequality of opportunity is very high in every country in this analysis, although this is quite variable across them, and countries with higher total inequality do not always show higher IOp. With respect to the ranking of countries, instead, while our results are robust to the choice of the inequality measure, they appear to be less robust to the choice of the estimation method.

From a methodological point of view, our analysis shows that some of the tools proposed in the literature for the measurement of inequality of opportunity in Western countries need to be handled with caution when analysing IOp in SSA. In all countries analysed, circumstances beyond individual control such as ethnicity, birthplace, and parental background interact in determining individual opportunity in a much more complex way than what we typically observe in Western societies. Thus, as a focus of future research, this complexity should be examined with country-specific and more data-intensive studies to further elucidate the best possible methods for determining IOp in non-Western countries.

Appendix I. Circumstances treatment for the non-parametric estimates

For Comoros, the circumstances considered are birthplace, parental education, and father and mother occupation. Birthplace (originally categorised into 37 villages) is recoded into three categories: born in Grande Comore, born in Anjouan (Grande Comore and Anjouan being the largest and most important islands), others (born in other smaller islands or outside Comoros). Parental education is coded into two categories: both parents having no education, at least one of the two having an elementary degree or higher. Father occupation is coded into three categories: employed in the agricultural sector, housekeeper, or other. The same coding is applied to mother occupation (see Table 7 in Appendix II for details).

In the Democratic Republic of Congo, the observable individual characteristics include father education, father occupation, and ethnicity. Father education is coded in number of years of education completed. Father occupation is recorded in ten categories ranging from family helper to senior executive. The survey records 399 possible ethnic groups. Unfortunately, these groups cannot be decoded using the available documentation, so although we can identify individuals sharing the same ethnicity, we cannot group them into a homogeneous macro-group as we do for other countries. This lack of information renders it impossible to construct types of sufficient size to allow inferences about their average per capita consumption (32 ethnic groups contain only one respondent and 131 less than 10). However, ethnicity can still be used as a dummy to explain inequality following the parametric approach.

We therefore present only parametric IOp estimates for Congo.

In the case of Ghana, we use information on three circumstances: birthplace, parental education, and ethnicity. In order to have a proper partition of the distribution, all three circumstances are recoded. In particular, birthplace (originally represented by 17 regions of birth) distinguishes between individuals born in the northern, central, and southern parts of the country. Parental education is coded as for Comoros. Finally, we divide the 64 ethnic groups present in the original data into four categories, according to their linguistic similarities. To this aim, we refer to the three main linguistic groups of the country: KWA, Gur, and Mande. Obviously, the fourth category encompasses the remaining ethnic groups (see Table 8 in Appendix II for details).

Birthplace, parental education, and parental occupation are the circumstances used for Guinea. Birthplace (originally represented by 34 villages of birth) is partitioned into seven categories depending on the region of birth: Kankan, Nzerekore, Faranah, Kindia, Labe, Mamou, and Boke. The last category refers to those individuals born outside Guinea. Parental education is coded as for Comoros, whereas parental occupation is partitioned into two categories: the first encompassing those individuals whose parents are employed in the agricultural sector, and the second encompassing all the individuals who have at least one parent that is employed in a sector other than the agricultural one (see Table 9 in Appendix II for details).

The data of Madagascar provides birthplace, parental education, and ethnicity as endogenous characteristics. Birthplace (originally represented by about 400 villages) is based on the 6 administrative provinces of birth: Antananarivo, Fianarantsoa, Mahajanga, Toamasina, Toliara, and Antsiranana. Note, however, that we aggregate Mahajanga and Toamasina (on the basis of geographic distance), ending up with five categories. Parental education is coded as for Comoros. The 25 ethnic groups present in the original dataset are grouped into three main categories on the basis of their main geographic location: the Costal, the Highlander, others (see Table 10 in Appendix II for details).

Observations for Malawi, in contrast, are only available for two circumstances: birthplace and parental education. Birthplace encompasses 31 categories, one for each district, plus one category grouping those individuals born outside Malawi. Again, parental education is coded as for Comoros (see Table 11 in Appendix II for details). Note that although information on mother tongue (used as a proxy of ethnicity) is available for Malawi, this information is only used for the parametric estimates because of the problem generated by the smallest size of types when the partition also accounts for ethnicity.

For Niger, the set of circumstances is represented by birthplace and ethnicity. Birthplaces (originally indicated as one of the 40 departments of birth) is coded into nine categories: the seven regions of the country (Agadez, Diffa, Dosso, Maradi, Tahoua, Tillaberi, and Zinder), the capital (Niamey), and others (individuals born outside Niger). The ethnic groups represented in the survey are Arab, Djema, Haoussa, Kanouri-Manga, Peul, Touareg, Toukou, foreigners, and a residual made of other ethnic groups. In order to have types with a sufficiently large population to allow for inference, in our analysis we group together Arab, Toukou, foreigners, and others (see Table 12 in Appendix II for details).

In the case of Nigeria, the circumstances considered are parental education, father occupation, and mother occupation. Concerning the first circumstance of education, we define the following categories: individuals with parents who have no education, individuals with at least one parent who has some primary education (not completed), individuals with at least one parent who completed primary education, individuals with at least one parent who has some secondary education (not completed), and individuals with at least one parent who completed secondary education or has a higher degree. Concerning the second circumstance of father occupation, we define the following categories: individuals whose father is employed in agriculture or not working and individuals whose father is employed in a different sector. The same coding is used for mother occupation. The partitions are made of 20 types with a sample size ranging between 99 and 4,941 (see Table 13 in Appendix II for details).

The circumstances available from Rwanda's data are birthplace, parental education, and parental occupation. Birthplace is characterised by six categories, each of them representing one of the five administrative regions of the country, with the last encompassing people born outside Rwanda. For parental education we follow the coding used for Comoros. The third circumstance, parental occupation, is coded into two categories: the first groups individuals with parents who both work in the agricultural or fishery sector, and the second groups individuals with at least one parent who does not work in the agricultural or fishery sector (see Table 14 in Appendix II for details).

As for Tanzania, we observe two circumstances: birthplace and parental education. Birthplace is categorised into 25 administrative regions: Dodoma, Arusha, Kilimanjaro, Tanga, Morogoro, Pwani, Dar es Salaam, Lindi, Mtwara, Ruvuma, Iringa, Mbeya, Singida, Tabora, Rukwa, Kigoma, Shinyanga, Kagera, Mwanza, Mara, Manyara, Kaskazini Unguja, Kusini Unguja, Mjini Magharibi, Kaskazini Pemba, and Kusini Pemba. Parental education is classified in one of two categories: both parents having a degree of education below elementary level or at least one parent having an elementary degree or higher (see Table 15 in Appendix II for details).

Birthplace and ethnicity are the information available for Uganda. Although the survey contains a large set of circumstances, such as parental education, parental occupation, area of birth, and ethnicity, we are forced to choose only two of them because of the large number of missing information for the other variables. In the original dataset, birthplace is distinguished into 56 districts plus the capital city. For practical reasons, this circumstance is recoded into four groups according to the level of development of each district as measured by the Human Development Index (UNDP, 2014), that is: low development (HDI between 0.231-0.433), lower intermediate (0.434-0.470), upper intermediate (0.472-0.498), and high (above 0.500). We also recode the 68 ethnic groups present in the original data on the basis on their linguistic origin: Easter Lacustrine Bantu, Western Lacustrine Bantu, Eastern Nilotic, Western Nilotic, and Ethnic Minorities (see Table 16 in Appendix II for details).[13]

[13]This subdivision is based on information reported by UNDP (2014) and Wairama (2001).

Appendix II. Opportunity Profiles

Table 7: Opportunity profile: Comoros 2004

rank	*birth location*	*parental education*	*father occupation*	*mother occupation*	*n 2004*	*p.c. consumption 2004*
1	others	none	other	other	29	1,844.41
2	others	none	housekeeper	agriculture	208	1,986.20
3	others	none	agriculture	other	71	2,088.10
4	Anjouan	elementary or above	other	agriculture	8	2,092.21
5	Anjouan	elementary or above	other	other	21	2,178.00
6	Grande Comore	none	agriculture	other	94	2,240.15
7	others	elementary or above	agriculture	agriculture	14	2,286.77
8	Anjouan	elementary or above	agriculture	other	26	2,294.45
9	Grande Comore	none	housekeeper	other	381	2,321.49
10	Anjouan	none	other	other	89	2,350.16
11	Grande Comore	none	agriculture	agriculture	643	2,351.03
12	Grande Comore	elementary or above	other	other	53	2,380.17
13	Grande Comore	elementary or above	housekeeper	agriculture	59	2,467.56
14	Grande Comore	none	housekeeper	agriculture	957	2,524.50
15	others	elementary or above	other	agriculture	2	2,546.14
16	Anjouan	none	housekeeper	other	268	2,562.99
17	Grande Comore	elementary or above	agriculture	other	22	2,577.82
18	others	none	housekeeper	other	125	2,589.02
19	Grande Comore	none	other	agriculture	173	2,608.87
20	others	elementary or above	agriculture	other	12	2,630.26
21	others	none	agriculture	agriculture	244	2,634.96
22	Anjouan	elementary or above	agriculture	agriculture	28	2,735.11
23	Anjouan	elementary or above	housekeeper	agriculture	10	2,857.03
24	Anjouan	none	agriculture	other	259	2,988.62
25	Grande Comore	elementary or above	other	agriculture	13	3,016.21
26	Anjouan	none	other	agriculture	75	3,022.97
27	Grande Comore	none	other	other	142	3,029.03
28	others	none	other	agriculture	16	3,059.25
29	Grande Comore	elementary or above	housekeeper	other	92	3,145.42
30	others	elementary or above	housekeeper	agriculture	16	3,219.43
31	Anjouan	none	housekeeper	agriculture	439	3,243.63
32	others	elementary or above	housekeeper	other	99	3,547.67
33	Anjouan	none	agriculture	agriculture	1,082	3,837.00
34	others	elementary or above	other	other	27	3,940.04
35	Anjouan	elementary or above	housekeeper	other	106	4,468.70
36	Grande Comore	elementary or above	agriculture	agriculture	33	36,616.50

Note: n 2004 is the sample size of each type in 2004; p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 8: Opportunity profile: Ghana 2013

rank	*ethnicity*	*birth location*	*parental education*	*n 2013*	*p.c. consumption 2013*
1	Kwa	north	none	793	917.40
2	others	north	none	282	923.73
3	Gur	north	none	11,519	1,103.89
4	Mande	north	none	244	1,285.88
5	Gur	centre	none	1,079	1,328.15
6	Gur	south	none	844	1,536.28
7	Gur	north	elementary or above	1,722	1,550.71
8	Mande	centre	none	78	1,561.71
9	Kwa	north	elementary or above	188	1,567.31
10	others	centre	none	147	1,570.04
11	Mande	south	none	67	1,683.66
12	others	centre	elementary or above	79	1,688.43
13	Mande	centre	elementary or above	36	1,731.18
14	Gur	centre	elementary or above	354	1,753.10
15	Kwa	south	none	7,852	1,792.07
16	Kwa	centre	none	2,962	1,907.49
17	Mande	north	elementary or above	23	1,980.65
18	Gur	south	elementary or above	363	2,154.66
19	Kwa	centre	elementary or above	3,715	2,181.44
20	others	north	elementary or above	34	2,257.19
21	others	south	none	185	2,330.63
22	Kwa	south	elementary or above	9,799	2,370.11
23	Mande	south	elementary or above	39	2,554.33
24	others	south	elementary or above	115	2,565.80

Note: n 2013 is the sample size of each type in 2013; p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 9: Opportunity profile: Guinea 2003

rank	*birth location*	*parental education*	*parenta occupation*	*n 2003*	*p.c. consumption 2003*
1	Kankan	none	agriculture	1,608	634.83
2	outside Guinea	elementary or above	agriculture	6	648.25
3	Nzerekore	elementary or above	agriculture	186	724.56
4	Faranah	none	agriculture	1,375	737.11
5	Nzerekore	none	agriculture	2,288	755.25
6	Boke	elementary or above	agriculture	65	772.03
7	outside Guinea	none	agriculture	155	843.57
8	Kankan	elementary or above	agriculture	83	846.85
9	Kankan	none	other	1,532	864.71
10	Mamou	elementary or above	agriculture	21	892.79
11	Mamou	none	agriculture	1,424	935.20
12	Faranah	elementary or above	agriculture	80	1,002.34
13	Boke	none	agriculture	1,321	1,020.08
14	Kindia	none	agriculture	1,740	1,024.43
15	Boke	none	other	973	1,029.60
16	Nzerekore	none	other	1,204	1,032.19
17	Faranah	none	other	930	1,081.10
18	Kindia	none	other	2,413	1,082.47
19	Kindia	elementary or above	agriculture	98	1,134.86
20	Kankan	elementary or above	other	278	1,211.21
21	Nzerekore	elementary or above	other	542	1,213.00
22	Labe	none	other	1,270	1,224.88
23	Mamou	none	other	1,131	1,235.99
24	Boke	elementary or above	other	329	1,289.69
25	outside Guinea	none	other	361	1,345.27
26	Labe	none	agriculture	1,456	1,356.63
27	Kindia	elementary or above	other	1,049	1,394.03
28	Mamou	elementary or above	other	221	1,422.85
29	Labe	elementary or above	other	258	1,590.13
30	outside Guinea	elementary or above	other	137	1,660.16
31	Faranah	elementary or above	other	277	1,699.66
32	Labe	elementary or above	agriculture	55	8,370.83

Note: n 2003 is the sample size of each type in 2003; p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 10: Opportunity profile: Madagascar 2005

rank	*ethnicity*	*birth location*	*parental education*	*n 2005*	*p.c. consumption 2005*
1	Coastal	Fianarantsoa	none	1,970	288.3178
2	others	Antananarivo	none	17	307.1195
3	Coastal	Mahajanga-Toamasina	none	4,094	324.658
4	Highlanders	Fianarantsoa	none	1,307	331.2441
5	Coastal	Toliara	none	2,137	344.4594
6	others	Toliara	none	35	391.0437
7	Coastal	Mahajanga-Toamasina	elementary or above	2,977	406.876
8	Coastal	Fianarantsoa	elementary or above	1,143	406.9271
9	Highlanders	Toliara	none	296	408.4453
10	Coastal	Antsiranana	none	1,121	414.0577
11	Highlanders	Fianarantsoa	elementary or above	2,232	421.0522
12	Highlanders	Antananarivo	none	1,796	429.2367
13	Highlanders	Antsiranana	none	66	431.6902
14	Coastal	Antananarivo	none	65	440.788
15	Highlanders	Mahajanga-Toamasina	none	1,124	457.916
16	others	Toliara	elementary or above	51	463.7084
17	others	Fianarantsoa	elementary or above	198	468.6341
18	others	Antananarivo	elementary or above	17	471.0351
19	Coastal	Toliara	elementary or above	1,118	471.8447
20	Highlanders	Toliara	elementary or above	662	484.0782
21	Highlanders	Mahajanga-Toamasina	elementary or above	1,766	494.0706
22	Coastal	Antsiranana	elementary or above	1,041	506.3073
23	Highlanders	Antananarivo	elementary or above	3,079	509.4836
24	others	Mahajanga-Toamasina	none	90	526.955
25	others	Fianarantsoa	none	199	534.6658
26	Highlanders	Antsiranana	elementary or above	103	564.9221
27	Coastal	Antananarivo	elementary or above	84	658.7123
28	others	Mahajanga-Toamasina	elementary or above	57	771.5343
29	others	Antsiranana	none	57	782.28
30	others	Antsiranana	elementary or above	49	1755.31

Note: n 2005 is the sample size of each type in 2005; p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 11: Opportunity profile: Malawi 2010

rank	*birth location*	*parental education*	*n 2010*	*p.c. consumption 2010*
1	310. Chikwawa	none	754	436.94
2	311. Nsanje	none	795	452.60
3	313. Neno	elementary or above	207	517.73
4	302. Machinga	none	807	522.29
5	301. Mangochi	none	861	524.74
6	313. Neno	none	483	535.28
7	310. Chikwawa	elementary or above	167	575.37
8	207. Mchinji	none	622	608.09
9	206. Lilongwe	none	1,332	649.88
10	306. Mwanza	none	501	649.95
11	312. Balaka	none	505	654.28
12	105. Mzimba	none	851	658.78
13	101. Chitipa	none	701	671.01
14	102. Karonga	none	652	674.53
15	309. Phalombe	none	727	676.20
16	208. Dedza	none	908	683.15
17	other	none	550	688.53
18	303. Zomba	none	677	707.21
19	205. Salima	none	671	710.82
20	204. Dowa	none	780	731.48
21	103. Nkhatabay	none	473	736.68
22	308. Mulanje	none	944	737.53
23	312. Balaka	elementary or above	235	751.51
24	306. Mwanza	elementary or above	215	780.78
25	305. Blanytyre	none	583	794.66
26	209. Ntcheu	none	725	796.74
27	301. Mangochi	elementary or above	167	798.94
28	201. Kasungu	none	553	811.98
29	203. Ntchisi	none	702	816.00
30	104. Rumphi	none	447	819.09
31	311. Nsanje	elementary or above	141	821.42
32	307. Thyolo	none	950	840.31
33	204. Dowa	elementary or above	311	855.84
34	202. Nkhota kota	none	534	867.61
35	303. Zomba	elementary or above	324	868.73
36	103. Nkhatabay	elementary or above	401	880.57
37	206. Lilongwe	elementary or above	406	883.34
38	207. Mchinji	elementary or above	243	895.77
39	302. Machinga	elementary or above	149	897.28
40	304. Chiradzulu	none	690	919.62
41	104. Rumphi	elementary or above	427	940.01
42	201. Kasungu	elementary or above	311	942.09
43	107. Mzuzu City	none	119	959.04
44	309. Phalombe	elementary or above	172	961.65
45	101. Chitipa	elementary or above	328	963.58
46	105. Mzimba	elementary or above	423	993.17
47	308. Mulanje	elementary or above	310	1,018.33
48	210. Lilongwe City	none	296	1,040.87
49	208. Dedza	elementary or above	244	1,079.29
50	304. Chiradzulu	elementary or above	298	1,187.05
51	209. Ntcheu	elementary or above	351	1,197.40
52	305. Blanytyre	elementary or above	317	1,209.50
53	102. Karonga	elementary or above	355	1,213.68
54	202. Nkhota kota	elementary or above	222	1,281.76
55	307. Thyolo	elementary or above	223	1,289.17
56	203. Ntchisi	elementary or above	245	1,305.45
57	205. Salima	elementary or above	261	1,376.61
58	314. Zomba City	none	313	1,382.68
59	315. Blantyre City	none	262	1,545.74
60	107. Mzuzu City	elementary or above	271	1,700.73
61	210. Lilongwe City	elementary or above	516	1,754.84
62	314. Zomba City	elementary or above	414	2,192.29
63	315. Blantyre City	elementary or above	493	2,724.07
64	other	elementary or above	222	3,021.64

Table 12: Opportunity profile: Niger 2011-2012

rank	*birth location*	*ethnicity*	*n 11-12*	*p.c. consumption 11-12*
1	Maradi	Peul	51	609.50
2	Maradi	Kanouri-Manga	16	747.85
3	Tillaberi	Touareg	251	753.47
4	Maradi	Touareg	113	826.32
5	Tahoua	Peul	29	893.59
6	Zinder	Touareg	176	927.62
7	Tillaberi	Djema	1,262	930.65
8	Maradi	Haoussa	1,215	945.87
9	Tahoua	Touareg	284	952.41
10	Dosso	Djema	996	969.80
11	Dosso	Kanouri-Manga	6	972.14
12	Zinder	Peul	52	975.32
13	Zinder	Kanouri-Manga	325	975.60
14	Zinder	other	90	977.14
15	Dosso	Peul	74	1,001.90
16	Dosso	Haoussa	565	1,026.85
17	Maradi	other	7	1,033.58
18	Zinder	Haoussa	870	1,048.96
19	Tahoua	Haoussa	1,108	1,066.47
20	Diffa	Peul	198	1,121.57
21	Diffa	Kanouri-Manga	516	1,127.07
22	Tillaberi	Haoussa	198	1,219.05
23	Diffa	other	170	1,224.04
24	Diffa	Touareg	10	1,325.61
25	Tahoua	Kanouri-Manga	7	1,358.41
26	Agadez	Peul	55	1,371.44
27	Tillaberi	other	14	1,432.79
28	Diffa	Djema	7	1,467.50
29	Tillaberi	Kanouri-Manga	6	1,510.25
30	Tillaberi	Peul	89	1,549.28
31	Agadez	Touareg	914	1,552.45
32	Agadez	Kanouri-Manga	5	1,637.37
33	Dosso	Touareg	10	1,686.33
34	Diffa	Haoussa	18	1,697.75
35	Niamey	Touareg	60	1,707.21
36	Dosso	other	6	1,717.18
37	Agadez	other	22	1,769.21
38	Niamey	Haoussa	489	1,914.64
39	Agadez	Haoussa	83	1,946.15
40	Niamey	Djema	954	2,015.78
41	Agadez	Djema	58	2,046.35
42	Niamey	Peul	165	2,179.48
43	Tahoua	Djema	26	2,329.03
44	Niamey	other	112	2,514.84
45	Zinder	Djema	33	2,531.27
46	Maradi	Djema	17	2,602.06
47	Tahoua	other	7	3,156.40
48	Niamey	Kanouri-Manga	35	3,160.12

Note: n 2011-2012 is the sample size of each type in 2011-2012; p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 13: Opportunity profile: Nigeria 2010-2011 and 2012-2013

rank	*parental education*	*father sector*	*mother sector*	*n 2010-11*	*eq. consumption 2010-11*	*n 2012-13*	*eq. consumption 2012-2013*
1	none	agriculture/not working	agriculture/not working	4,941	995.59	4,923	1,304.95
2	primary incomplete	agriculture/not working	agriculture/not working	1,478	1,022.60	1,460	1,155.32
3	secondary incomplete	agriculture/not working	agriculture/not working	357	1,030.36	416	1,389.08
4	primary incomplete	agriculture/not working	other sectors	1,048	1,033.73	1,076	1,120.54
5	primary incomplete	other sectors	other sectors	357	1,184.65	301	1,473.13
6	primary complete	agriculture/not working	agriculture/not working	818	1,233.36	837	1,599.52
7	none	agriculture/not working	other sectors	1,471	1,235.93	1,314	1,653.99
8	none	other sectors	agriculture/not working	241	1,244.19	240	1,645.72
9	primary incomplete	other sectors	agriculture/not working	150	1,270.40	141	1,538.64
10	primary complete	other sectors	agriculture/not working	370	1,457.19	364	1,849.59
11	none	other sectors	other sectors	463	1,477.95	411	1,858.29
12	secondary incomplete	other sectors	agriculture/not working	248	1,485.66	261	1,772.21
13	primary complete	agriculture/not working	other sectors	459	1,526.24	442	1,779.26
14	secondary incomplete	agriculture/not working	other sectors	342	1,545.63	303	1,841.72
15	secondary complete or above	agriculture/not working	agriculture/not working	99	1,620.88	123	1,779.18
16	secondary complete or above	other sectors	agriculture/not working	189	1,628.39	254	2,339.49
17	primary complete	other sectors	other sectors	682	1,784.08	642	2,061.40
18	secondary incomplete	other sectors	other sectors	694	1,822.75	676	2,084.43
19	secondary complete or above	agriculture/not working	other sectors	216	2,006.31	175	2,602.63
20	secondary complete or above	other sectors	other sectors	590	2,350.06	591	2,782.42

Note: n 2010-2011 (n 2012-2013) is the sample size of each type in 2010-2011 (2012-2013); p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 14: Opportunity profile: Rwanda 2000

rank	*birth location*	*parental education*	*parenta occupation*	*n 2000*	*p.c. consumption 2000*
1	Kigali	none	agriculture	856	499.89
2	North	none	agriculture	1,969	503.94
3	North	elementary or above	agriculture	684	516.34
4	Kigali	elementary or above	agriculture	345	530.25
5	South	none	agriculture	2,744	542.20
6	West	none	agriculture	2,626	555.12
7	West	elementary or above	agriculture	996	577.33
8	South	elementary or above	agriculture	1,320	593.17
9	East	elementary or above	agriculture	531	615.80
10	East	none	agriculture	1,004	635.71
11	South	none	other	28	831.01
12	outside Rwanda	none	agriculture	325	907.77
13	East	none	other	3	938.79
14	outside Rwanda	elementary or above	agriculture	311	1,028.17
15	Kigali	none	other	8	1,295.43
16	West	elementary or above	other	65	1,521.86
17	outside Rwanda	none	other	17	1,639.40
18	South	elementary or above	other	78	1,678.08
19	East	elementary or above	other	12	1,712.88
20	North	none	other	3	1,777.56
21	North	elementary or above	other	33	1,796.77
22	West	none	other	16	1,840.53
23	Kigali	elementary or above	other	51	1,993.87
24	outside Rwanda	elementary or above	other	87	4,163.61

Note: n 2000 is the sample size of each type in 2000; p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 15: Opportunity profile: Tanzania 2009-2010 and 2010-2011

rank	*birth location*	*parental education*	*n 09-10*	*p.c. consumption 09-10*	*n 10-11*	*p.c. consumption 10-11*
1	Kaskazini Pemba	below elementary	198	629.43	218	762.69
2	Rukwa	below elementary	103	693.35	104	642.80
3	Kigoma	below elementary	188	706.21	214	663.65
4	Dodoma	below elementary	189	784.72	211	731.62
5	Kusini Pemba	below elementary	204	791.48	239	905.25
6	Kigoma	elementary or above	135	792.55	220	925.11
7	Mwanza	below elementary	230	809.92	332	827.18
8	Tabora	below elementary	234	824.96	278	884.48
9	Shinyanga	below elementary	365	827.26	462	784.89
10	Singida	below elementary	122	837.61	142	817.32
11	Kaskazini Unguja	below elementary	181	856.39	198	940.17
12	Ruvuma	below elementary	199	859.33	222	828.95
13	Manyara	below elementary	114	866.17	132	696.51
14	Mbeya	below elementary	239	903.06	249	977.03
15	Rukwa	elementary or above	100	917.05	133	894.88
16	Lindi	below elementary	289	918.80	327	945.96
17	Mtwara	below elementary	312	921.24	358	936.81
18	Tanga	below elementary	204	929.60	234	790.71
19	Ruvuma	elementary or above	178	930.56	231	795.08
20	Mara	below elementary	97	930.74	100	1,011.30
21	Kusini Pemba	elementary or above	133	939.32	155	1,080.71
22	Kaskazini Unguja	elementary or above	78	955.11	104	1,155.49
23	Kusini Unguja	elementary or above	54	958.46	82	1,169.14
24	Kaskazini Pemba	elementary or above	105	983.33	129	1,005.68
25	Morogoro	below elementary	187	992.61	207	995.97
26	Kusini Unguja	below elementary	46	996.15	65	1,213.29
27	Tabora	elementary or above	92	1,022.51	174	1,142.56
28	Kagera	below elementary	157	1,047.17	195	1,050.00
29	Shinyanga	elementary or above	186	1,111.98	312	1,068.13
30	Dodoma	elementary or above	96	1,131.33	112	1,144.34
31	Manyara	elementary or above	62	1,142.32	65	846.97
32	Lindi	elementary or above	135	1,152.99	202	1,268.26
33	Iringa	below elementary	214	1,158.59	218	1,039.38
34	Arusha	below elementary	130	1,215.97	168	1,069.28
35	Mbeya	elementary or above	158	1,218.00	250	1,274.27
36	Mjini/Magharibi Unguja	below elementary	159	1,271.92	178	1,236.54
37	Mtwara	elementary or above	213	1,272.47	304	1,316.40
38	Mara	elementary or above	107	1,293.36	155	1,136.23
39	Mwanza	elementary or above	186	1,294.52	328	1,209.00
40	Kagera	elementary or above	194	1,307.17	280	1,347.72
41	Mjini/Magharibi Unguja	elementary or above	289	1,322.01	341	1,505.85
42	Pwani	below elementary	191	1,332.65	227	1,250.90
43	Singida	elementary or above	102	1,340.07	137	1,324.37
44	Morogoro	elementary or above	211	1,409.79	280	1,408.47
45	Iringa	elementary or above	190	1,502.40	244	1,436.83
46	Tanga	elementary or above	227	1,511.35	300	1,242.77
47	Pwani	elementary or above	127	1,524.16	143	1,639.94
48	Kilimanjaro	below elementary	210	1,546.54	219	1,259.20
49	Dar es salaam	below elementary	146	1,609.63	167	1,606.85
50	Arusha	elementary or above	108	1,616.80	140	1,433.71
51	Kilimanjaro	elementary or above	276	1,938.00	336	1,950.82
52	Dar es salaam	elementary or above	469	2,455.81	570	2,248.75

Note: n 2009-2010 (2010-2011) is the sample size of each type in 2009-2010 (2010-2011); p.c. consumption is per capita consumption and it is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 16: Opportunity profile: Uganda 2009-2010 and 2010-2011

rank	*HDI class*	*ethnic group*	*n 09-10*	*eq. consumption 09-10*	*n 10-11*	*eq. consumption 10-11*
1	low	Ethnic minorities	46	565.01	25	643.57
2	intermediate	Western Nilotic	739	599.26	758	687.87
3	low	Western Nilotic	671	655.57	630	737.31
4	low	Central Sudanic	138	744.23	146	717.18
5	low	Eastern Nilotic	499	765.74	502	668.88
6	high	Ethnic minorities	535	796.01	432	634.67
7	intermediate	Western lacustrine Bantu	646	957.96	552	1,052.61
8	intermediate	Ethnic minorities	237	993.25	173	772.92
9	high	Eastern Nilotic	89	1,055.22	75	874.06
10	low	Eastern lacustrine Bantu	198	1,064.23	191	1,563.04
11	low	Western lacustrine Bantu	146	1,070.37	120	1,063.68
12	intermediate	Eastern Nilotic	284	1,112.25	266	902.45
13	intermediate	Eastern lacustrine Bantu	880	1,166.45	901	1,009.49
14	high	Western lacustrine Bantu	799	1,257.25	692	1,105.65
15	high	Western Nilotic	56	1,296.31	54	837.56
16	intermediate	Central Sudanic	296	1,301.76	287	945.67
17	high	Eastern lacustrine Bantu	1,920	1,646.84	1,633	1,639.90
18	high	Central Sudanic	15	2,499.20	17	1,935.08

Note: n 2009-2010 (2010-2011) is the sample size of each type in 2009-2010 (2010-2011); p.c. consumption is per capita consumption and it is expressed in 2011 PPP $ *Source: Authors' calculation based on surveys listed in Table 1.*

Appendix III. MLD estimates

Table 17: Non-parametric estimates (MLD)

country	*sample*	*consumption*	*types*	*inequality*	*IOp*	*IOp %*	max MLD	Adj. IOp %
Comoros	5,936	2,975	36	0.5358	0.0669	12.49	0.5225	12.81
Ghana	42,519	1,838	24	0.2949	0.0392	13.29	0.2741	14.30
Guinea	24,866	1,000	32	0.3121	0.0510	16.36	0.3071	16.62
Madagascar	28,951	415	30	0.2294	0.0179	7.82	0.2253	7.96
Malawi	30,137	855	64	0.3806	0.0744	19.54	0.3791	19.61
Niger	11,774	1,071	48	0.1562	0.0245	15.67	0.1549	15.80
Nigeria 2010-11	14,916	1,298	20	0.2623	0.0347	13.25	0.2376	14.62
Nigeria 2012-13	14,560	1,601	20	0.2603	0.0321	12.34	0.2367	13.57
Rwanda	14,112	641	24	0.3357	0.0425	12.66	0.3215	13.22
Tanzania 2009-10	9,119	1,133	52	0.2547	0.0448	17.59	0.2537	17.66
Tanzania 2010-11	11,391	1,112	52	0.2598	0.0410	15.79	0.2588	15.85
Uganda 2009-10	8,194	1,157	24	0.3450	0.0529	15.34	0.3310	15.99
Uganda 2010-11	7,454	1,039	24	0.3836	0.0558	14.55	0.3672	15.20

Note: Per capita consumption is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

Table 18: Parametric estimates (MLD)

country	*sample*	*consumption*	*types*	*inequality*	*IOp*	*IOp %*
Comoros	5,936	2,975	91	0.5358	0.0875	16.34
Congo DR	39,578	1,535	402	0.2236	0.0494	22.10
Ghana	42,519	1,838	125	0.2949	0.0858	29.10
Guinea	24,866	1,000	96	0.3121	0.0352	11.27
Madagascar	28,951	415	445	0.2294	0.0719	31.35
Malawi	30,137	855	71	0.3806	0.1268	33.32
Niger	11,774	1,071	50	0.1562	0.0259	16.58
Nigeria 2010-11	14,916	1,298	40	0.2623	0.0425	16.22
Nigeria 2012-13	14,560	1,601	40	0.2603	0.0457	17.57
Rwanda	14,112	641	76	0.3357	0.0696	20.73
Tanzania 2009-10	9,119	1,133	41	0.2547	0.0590	23.16
Tanzania 2010-11	11,391	1,112	40	0.2598	0.0538	20.71
Uganda 2009-10	8,194	1,157	100	0.3450	0.0771	22.35
Uganda 2010-11	7,454	1,039	102	0.3836	0.1062	27.69

Note: Per capita consumption is expressed in 2011 PPP $.

Source: Authors' calculation based on surveys listed in Table 1.

References

Aaberge, R., Mogstad, M., Peragine, V. (2011). Measuring long-term inequality of opportunity, *Journal of Public Economics*, 95(3), 193-204.

Araar A. and Duclos J. Y. (2009), An algorithm for computing the Shapley Value, Mimeo, PEP and CIRPEE, Universite Laval. available at: `http://dad.ecn.ulaval.ca/pdf_files/shap_dec_aj.pdf`.

Bourguignon, F., Ferreira, F.H.G., Menéndez, M. (2007). Inequality of opportunity in Brazil, *Review of Income and Wealth*, 53(4), 585-18.

Brautigam, D.A, Knack, S. (2014). Foreign aid, institutions, and governance in Sub-Saharan Africa, *Economic Development and Cultural Change*, 52(2), 255-285.

Checchi, D., Peragine, V. (2010). Inequality of opportunity in Italy, *Journal of Economic Inequality*, 8(4), 429-50.

Cogneau, D., Mesplé-Somps, S. (2008). Inequality of opportunity for income in five countries of Africa, in J. Bishop and B. Zheng (eds.) *Research on Economic Inequality*, Volume 16, Emerald Group Publishing Limited, 99-128.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility, *The Journal of Economic Perspectives*, 79-102.

De Ree, J., Nillesen, E. (2009). Aiding violence or peace? The impact of foreign aid on the risk of civil conflict in sub-Saharan Africa, *Journal of Development Economics*, 88(2), 301-313.

Easterly, W., Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions, *The Quarterly Journal of Economics*, 1203-1250.

Elbers, C., Lanjouw, P., Mistiaen, J., Ozler, B. (2008). Re-interpreting sub-group inequality decompositions, *The Journal of Economic Inequality*, 6(3), 1569-1721.

Ellis, F. (2012). 'We are all poor here': economic difference, social divisiveness and targeting cash transfers in sub-saharan Africa, *Journal of Development Studies*, 48(2), 201-214.

FEWS (2013). GUINEA production and market flow map report, FEWS NET April 2013, available at: http://www.fews.net/sites/default/files/documents/reports/Guinea-marketflow-final-en.pdf.

Ferreira, F.H.G., Gignoux, J. (2011). The measurement of inequality of opportunity: theory and an application to Latin America, *Review of Income and Wealth*, 57(4), 622-657.

Ferreira, F.H.G., Peragine, V. (2015). Equality of opportunity: theory and evidence, World Bank Policy Research Working Paper, (7217).

Ferreira, F.H.G., Lakner, C., Lugo, M., Ozler, B. (2014). Inequality of opportunity and economic growth : a cross-country analysis, Policy Research Working Paper WPS6915, World Bank.

Ferreira, F.H.G., Gignoux, J., Aran, M. (2011). Measuring inequality of opportunity with imperfect data: the case of Turkey, *The Journal of Economic Inequality*, 9(4), 651-680.

Fleurbaey, M. (2008). *Fairness, responsibility and welfare*, Oxford University Press.

Glick, P., Sahn, D.E. (2000). Schooling of girls and boys in a West African country: the effects of parental education, income, and household structure, *Economics of education review*, 19(1), 63-87.

Hassine, N.B. (2011). Inequality of opportunity in Egypt, *World Bank Economic Review*, 26(2), 265-295.

Goren E. (2014). How ethnic diversity affects economic growth, *World Development*, 59, 275-297

Lacina, B., Gleditsch, N.P. (2005). Monitoring trends in global combat: a new dataset of battle deaths, *European Journal of Population*, 21 (2-3), 145-166.

Lambert, P.J., Aronson, J.R. (1993). Inequality decomposition analysis and the Gini coefficient revisited, *The Economic Journal*, 103(420), 1221-1227.

Lassibille, G., Tan, J. P. (2005). The returns to education in Rwanda, *Journal of African Economies*, 14(1), 92-116.

Lefranc, A., Pistolesi, N., Trannoy, A. (2009). Equality of opportunity and luck: definitions and testable conditions, with an application to income in France, *Journal of Public Economics*, 93(11), 1189-1207.

Lloyd, C. B., Blanc, A. K. (1996). Children's schooling in Sub-Saharan Africa: the role of fathers, mothers, and others, *Population and development review*, 265-298.

Luongo, P. (2011). The implication of partial observability of circumstances on the measurement of IOp, *Research on Economic Inequality* 19(2), 23-49.

Madise, N. J., Matthews, Z., Margetts, B. (1999). Heterogeneity of child nutritional status between households: A comparison of six Sub-Saharan African countries, *Population studies*, 53(3), 331-343.

Marrero, G. A., Rodríguez, J. G. (2013). Inequality of opportunity and growth, *Journal of Development Economics*, 104, 107-122.

Milanovic, B. (2003). Is inequality in Africa really different?, World Bank Policy Research Working Paper, (3169).

Moradi, A., Baten, J. (2005). Inequality in Sub-Saharan Africa: new data and new insights from anthropometric estimates, *World Development*, 33(8), 1233-1265.

Peragine, V. (2002). Opportunity egalitarianism and income inequality: the rank-dependent approach, *Mathematical Social Sciences*, 44, 45-64.

Peragine, V., Serlenga, L. (2008). Higher education and equality of opportunity in Italy, In J. Bishop, and B. Zheng, eds., *Research in Economic Inequality*, 16.

Piraino, P. (2015). Intergenerational earnings mobility and equality of opportunity in South Africa, *World Development*, 67, 396-405.

Ramos, X., Van de gaer, D. (2012). Empirical approaches to inequality of opportunity: principles, measures and evidences, Working Paper 259. ECINEQ.

Roemer, J.E. (1998). *Equality of opportunity.* Harvard University Press, Cambridge, MA.

Roemer J.E., Trannoy, A. (2015). 'Equality of opportunity', in *Handbook of Income Distribution*, Anthony B. Atkinson and Francois Bourguignon (eds), vol. 2, Elsevier.

Schultz, P. (2004). Evidence of returns to schooling in Africa from household surveys: monitoring and restructuring the market for education, *Journal of African Economies*, 13(2), 95-148.

Thorbecke, E. (2013). The interrelationship linking growth, inequality and poverty in Sub-Saharan Africa, *Journal of African Economies*, 22(suppl 1), i15-i48.

UNDP (2014). Human Development Report 2014 sustaining human progress: reducing vulnerabilities and building resilience, United Nations Development Programme, New York.

Van de gaer, D. (1993). *Equality of opportunity and investment in human capital*, Ph.D. Dissertation, Katholieke Universiteit Leuven, Belgium.

Wairama B. (2001). Uganda: The marginalization of minorities. minority rights group international report.

Wendelspiess, W., Soloaga, I. (2014). Iop: estimating ex-ante inequality of opportunity, *Stata Journal*, 14(4), 830-846.

World Bank (2015). Purchasing power parities and the real size of world economies, 2015 International Bank for Reconstruction and Development The World Bank.

World Bank (2006). World Development Report 2006: equity and development, Washington, DC:World Bank.