Policy Research Working Paper 10723

New Evidence on Inequality of Opportunity in Sub-Saharan Africa

More Unequal Than We Thought

Aziz Atamanov
P. Facundo Cuevas
Jeremy Lebow
Daniel Gerszon Mahler

WORLD BANK GROUP
Poverty and Equity Global Practice
March 2024

Policy Research Working Paper 10723

Abstract

Unequal access to economic opportunity for individuals with different innate characteristics, such as ethnicity or parents' socioeconomic status, is often seen as both morally undesirable and bad for economic growth. This paper estimates inequality of opportunity, or the share of inequality explained by birth characteristics, across 18 countries in Sub-Saharan Africa. For many countries, this is the first time inequality of opportunity is measured. The paper uses nationally representative household survey data harmonized to allow for cross-country comparisons. Using consumption per capita as the outcome, the findings show that inequality of opportunity in Sub-Saharan Africa is stark and more pronounced than previously estimated. On average, inherited circumstances explain more than half of inequality in the region. Estimates range from 40 to 60 percent in most countries and reach 74 percent in South Africa. The findings show that birthplace, parents' education, and ethnicity tend to be the most significant contributors, but there is large variation in the importance of circumstances across countries. This represents the most comprehensive estimate of inequality of opportunity to date in the poorest and one of the most unequal regions in the world, and it underscores the pressing need for policy makers to intensify their efforts to address inequality of opportunity to foster societies that are more equitable and unlock the full potential for growth in the region.

This paper is a product of the Poverty and Equity Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp.

The authors may be contacted at jlebow@worldbank.org.

New Evidence on Inequality of Opportunity in Sub-Saharan Africa: More Unequal Than We Thought

Aziz Atamanov, P. Facundo Cuevas, Jeremy Lebow, Daniel Gerszon Mahler[1]

Originally published in the Policy Research Working Paper Series on *March 2024*.
This version is updated on *April 2024*.
To obtain the originally published version, please email prwp@worldbank.org.

JEL: D63; E24; O15; O40; O55

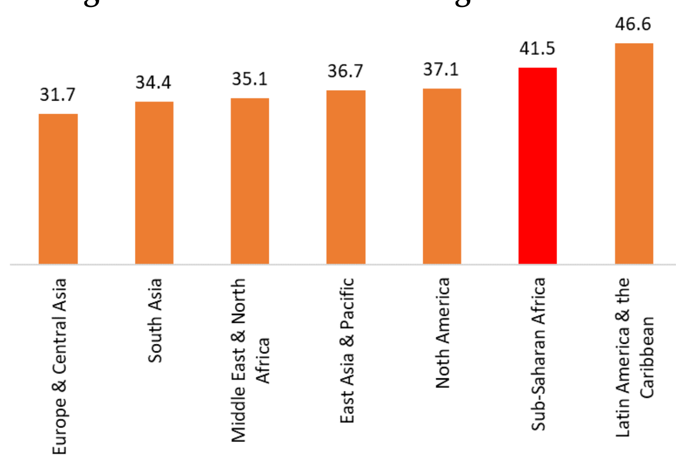
Keywords: Inequality of opportunity, consumption inequality, Sub-Saharan Africa, circumstances

[1] All authors: World Bank. Acknowledgements: The authors would like to thank Pierella Paci, Maurizio Bussolo, Ambar Narayan, Pedro Salas-Rojo, and Vito Peragine for their generous comments and suggestions. We are also thankful for overall guidance and support from Andrew Dabalen and Johan Mistiaen. This work also would not have been possible without the exceptional data harmonization work completed by Joe Green.

Section 1: Introduction

Sub-Saharan Africa (SSA) is not just the poorest region in the world, it is also one of the most unequal. By the latest data available, in 2019 SSA was home to around 14 percent of the world's population and 60 percent of the world's extreme poor, which is globally monitored at the International Poverty Line of $2.15 per person per day (in 2017 purchasing power parities, or PPP). At the same time, the average Gini coefficient is 41.5, second only to Latin America and the Caribbean, which has an average Gini of 46.6.[2] The magnitude and nature of income inequality in SSA may be a major impediment to growth, poverty reduction, and political stability in the region (Thornbecke 2023, Wu et al. 2024).

Figure 1: Average Gini coefficient across regions of the world



Unweighted average of Gini coefficients across countries in each region using latest data point available in the period 2011-2019. Source: World Bank Poverty and Inequality Platform (October 2023).

The literature emphasizes the contrasting role of two different types of inequalities. Inequality of opportunity (IOp) is the inequality resulting from factors determined at birth or during childhood and therefore outside a person's control. By contrast, inequality of effort (IE) reflects inequality arising from individual choices and decisions (Roemer 1993). The distinction matters not just from a moral point of view: evidence suggests that high IOp constrains economic growth because of resource misallocation as opportunities tend to favor those with certain inherited circumstances rather than those with more talent, creating inefficient allocations and preventing the economy from reaching its full potential. On the other hand, IE incentivizes human capital accumulation, investment,

[2] The Gini coefficient is a measure of income or wealth dispersion in which 0 represents perfect equality (all people have the same income or wealth) and 1 represents maximal inequality (one person has all the income or wealth while others have none).

and savings among talented individuals by directly reflecting reward to effort, thus stimulating growth (Ferreira 2005, Marrero et al. 2013). Empirically, the strong negative relationship between IOp and growth has been documented in various settings (Marrero et al. 2013, Bradbury et al. 2016; Carranza 2020).[3] Therefore, singling out the extent and sources of IOp and adequately addressing it is key to fostering much needed economic growth and poverty reduction, especially in SSA where poverty and inequality are high and economic growth is slowing (Wu et al. forthcoming).

The measurement of IOp revolves around the extent to which disparities in outcomes are due to circumstances beyond a person's control. This typically involves using income or consumption data as the outcome variable and birth characteristics as circumstances, such as region of birth or parental education and occupational background. Because of the heavy data requirements, empirical work has compiled estimates of IOp across many upper-income countries, while estimates in low- and middle-income settings, especially in SSA, are rare.

Until now, the main effort for the SSA region had been Brunori et al. (2019), which estimates IOp across 10 countries using data from 2000-2013. They find that IOp in SSA explains on average 47 percent of inequality in consumption with estimates ranging between 40-56 percent across countries. Though reflecting the data availability and methodological advancements available at the time, the analysis is limited to 10 countries and only a subset of variables is available in each country.[4] Because it is not feasible to observe all characteristics that are fixed at birth and outside an individual's control, IOp estimates are often seen as a lower bound estimate of true IOp (Ferreira et al. 2011). As often pointed out, when the set of observed circumstances is limited, the degree of underestimation is large and lower bound estimates of IOp end up being of limited use for policy (Kanbur & Wagstaff 2014).

---

[3] The negative relationship between IOp and growth has been documented using variation over time and across regions of the U.S. (Marrero et al. 2013, Bradbury et al. 2016) and in Europe (Carranza 2020). Evidence including developing countries is less conclusive and constrained by data limitations (Ferreira et al. 2018).

[4] In another noteworthy attempt, Ferreira et al. (2018) estimate IOp across 75 countries between 1980-2005, including 22 in Sub-Saharan Africa using data from the Demographic and Health Surveys (DHS). However, there are two large disadvantages to using the DHS to estimate IOp: welfare is measured using a wealth index because household consumption is not available, and parental information such as parents' education is unavailable. Their IOp estimates in SSA range from 1-39 percent, though excluding parental information likely leads to substantial downward bias. A World Bank report (2022a) implemented an IOp analysis in 5 countries in Southern Africa, though data limitations meant it was also unable to consider parental characteristics. This report found that IOp ranged from 15-26 percent, and increased to 46 percent in South Africa after race was included as a circumstance. Other studies examine drivers of inequality more generally – for example, Tetteh-Baah et al. (2024) use the DHS to calculate horizontal inequality in various outcomes across Africa based on location, ethnicity, gender, and religion – but they often do not include income or consumption as an outcome or incorporate information on parental characteristics.

In this paper, we provide the most comprehensive estimates of IOp in SSA to date. Our analysis spans 18 countries with data collected in 2012 or later.[5] We use consumption per capita as the outcome and we include the following circumstances: region of birth, urban-rural birthplace, parental education, parental industry, ethnicity, and religion. To estimate IOp, we use a recent methodological improvement based on machine-learning – conditional inference regression trees and forests – that effectively trades off positive and negative bias and has various other advantages, which are discussed further in Section 2. Our contribution relative to the existing literature on IOp in SSA is threefold. First, we expand the number of countries from 10 to 18, which increases the coverage of the SSA population from 39 to 59 percent. Second, we update estimates by a full decade by using and harmonizing the most recent available data. Third, we increase the circumstance variables included and their availability across countries, thus decreasing under-estimation and improving the comparability of IOp estimates across countries.

Taking a simple average across countries, we find that IOp explains 54 percent of consumption inequality. This is 15 percent higher than previous estimates in Brunori et al. (2019) and is shockingly high: it means that over half of inequality in SSA is due to IOp from inherited circumstances. We find that IOp is also 54 percent at the median and shows a wider range than previously thought. It reaches 74 percent in South Africa, followed by Niger (62 percent) and Burkina Faso (64 percent). By contrast, it is as low as 26 percent in Ethiopia, followed by Gabon (41 percent) and Malawi (45 percent). Put in context, our estimates suggest that IOp in SSA is broadly in the same range as IOp in South Asia and Latin America, and notably higher than in East Asia, Central Asia and high-income countries (World Bank, 2023; Bussolo et al., 2023).[6]

A breakdown of the importance of each circumstance reveals that birthplace, parental education, and ethnicity are most important, though there is variation in the importance of each circumstance across countries. South Africa stands out as the country in which race plays an outsized role in explaining consumption.

Given the variety of contexts and surveys, not all circumstances are available in every country. Therefore, we conduct two separate analyses: (1) estimate IOp using all available circumstances in each country, which produces the best possible country-specific estimate of IOp, and (2) estimate IOp using the circumstances that are available in all countries, which include birth region and parental education – two of the largest drivers

[5] Surveys in our study sample were collected between 2018-2019 for 11 countries and between 2012-2017 for the remaining 7 countries.

[6] A caveat is that it is difficult to compare IOp in SSA with other regions due to differences in available data and methodologies. An analysis using methods comparable to this paper confirms that IOp is much higher in SSA than in Europe, where it never exceeds 15 percent (Brunori et al., 2023).

of IOp which alone explain between 35-53 percent of inequality in consumption. The latter specification is more comparable across countries, and it reveals that inequality explained by these circumstances is generally higher in West Africa than in Eastern or Southern Africa (excluding South Africa which has very high IOp).

In light of the theory and evidence on IOp as an inhibitor of economic growth, our findings on the stark level of IOp in SSA underscore the pressing need to intensify efforts to invest in equalizing opportunities in the region, to create a more equitable society and to unlock its full growth potential.

The remainder of the paper is organized as follows: Section 2 presents the theoretical framework and estimation approach; Section 3 introduces the data used in the empirical work; Section 4 presents the results; Section 5 discusses and benchmarks the findings in the global context; and Section 6 concludes and discusses policy implications.

Section 2: Inequality of Opportunity – Theory and Estimation

In the canonical IOp framework, individual consumption, which is the measure of welfare we use in this paper, is seen as a function of two components: circumstances and effort (Roemer 1998, Van der Gaer 1993).[7]

$$C = f(c, e)$$

Thus, all individuals with the same circumstance and effort will have the same consumption, though the specific process through which circumstance and effort affect consumption is not modeled. Given a simple inequality measure of consumption, $I(C)$, the extent to which circumstances explain inequality is given by the Relative Inequality of Opportunity Index:

$$IOp = \frac{I(\hat{C})}{I(C)}$$

where $\hat{C}$ is a smoothed consumption measure in which each person is assigned the average consumption in their "type" or subgroup within which observable circumstances are shared. Thus, the numerator in IOp measures the counterfactual inequality in

[7] In this framework, the definition of effort is specifically the determinants of welfare that are orthogonal to circumstances. For example, if children with more educated parents also exert more effort in school, this is considered part of the effect of parental education.

consumption due only to circumstances, the denominator is total inequality, and the ratio is the share of inequality that is due only to circumstances.[8,9]

The primary inequality measure that we use in this paper, $I(.)$, is the Gini Index. While the Mean Log Deviation (MLD) is a common choice due to its decomposability properties, it has the disadvantage of being more sensitive to extreme values. The numerator of IOp by construction has no extreme values and will therefore be smaller as a share of overall inequality, pushing IOp estimates toward zero (Aaberge et al. 2011, Palmisano et al. 2022, Brunori et al. 2019).[10]

In practice, a typical approach to estimating $\hat{C}$, known as the parametric approach, is to use OLS to regress individual consumption on the circumstance variables and use the estimated coefficients to calculate predicted consumption (Ferreira et al. 2011). These coefficients capture both the direct effect of circumstances on consumption and the indirect effect via their influence on effort. For example, if individuals born to an ethnic group have lower high school completion rates on average, the model will consider the corresponding effect on consumption to be due to circumstances. Only differences in consumption that are orthogonal to ethnicity and all other circumstances will be attributed to effort. Furthermore, not being able to measure all birth circumstances means there will be omitted variables due to unobserved circumstances. This is why measures of IOp in the literature are typically seen as lower bounds on the true extent of IOp (Ferreira et al. 2011).

A limitation of this parametric approach is that it imposes an arbitrary functional form on the relationship between circumstances and consumption. For example, the role of parental education may vary across birth region or ethnicity, while the method described above imposes that it is the same for everyone, which will tend to create a downward bias in IOp. This approach can be improved and aligned with the original theory of Roemer (1993) by fully interacting all circumstance variables in the regression – identical to simply assigning individuals to the average consumption within their specific type.

[8] This is often referred to as "relative IOp", while "absolute IOp" is $I(\hat{C})$. We simply use IOp to refer to relative IOp.

[9] This approach is known as "ex-ante" IOp in the literature, in which equality of opportunity is achieved if opportunity is the same for all individuals regardless of circumstances. This differs from the "ex-post" approach, in which equality is achieved when there is no difference in consumption across individuals with the same *effort*. In this paper, we focus on ex-ante IOp. See Palmisano et al. (2022) for a summary of these two approaches.

[10] In Figure A1, we show that, as expected, IOp estimates are substantially smaller when using MLD instead of Gini, ranging from 8 percent in Ethiopia to 47 percent in South Africa, while the country IOp ranking remains similar.

This is known as the non-parametric approach. However, this is highly demanding of the data, bringing about an exponential increase in the number of types with each circumstance variable that is added, and in practice creating an upward bias (Brunori et al. 2018).

In this paper, we use an approach proposed by Brunori et al. (2023) that efficiently trades off these downward and upward biases by using a conditional inference regression tree to partition the sample into circumstance types. A regression tree is an algorithm that splits the data into non-overlapping nodes based on a set of covariates (in this case circumstances) to predict an outcome out-of-sample (in this case consumption). At each branch of the tree, the algorithm will find the circumstance split associated with the largest difference in consumption. This is done using a set of statistical tests to prevent over-fitting: First, the algorithm finds the circumstances that are statistically significantly related to consumption (specifically, with a p-value less than .01 after a Bonferroni correction for multiple hypothesis testing), and then selects the one with the smallest p-value. Next, for that circumstance, it tests all possible divisions to find the one that creates the largest difference in consumption between groups. If the circumstance is continuous or ordered, the algorithm chooses a splitting point. If it is categorical, the algorithm separates values into two groups. This process is repeated until no more splits are statistically significant, or the maximum allowed "depth" has been reached.

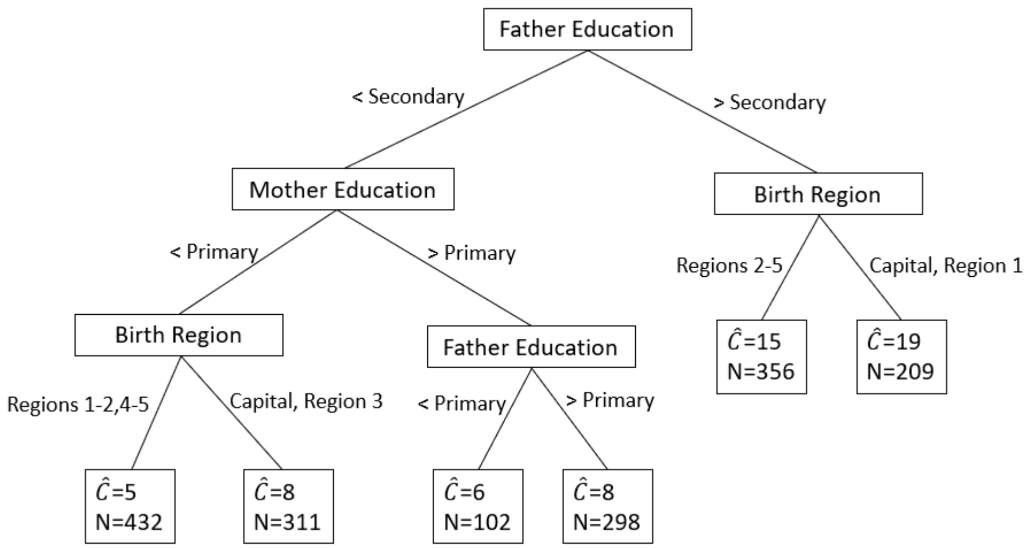
Consider the following fictitious example using three circumstances: father's education, mother's education, and region of birth. In the first node, the algorithm splits the sample by whether the father has completed secondary education or not. Among those whose fathers have completed secondary, the algorithm decides that the most important factor determining differences in consumption is region, with higher consumption among those born in the Capital or Region 1. At this point, the branch ends because there are no more circumstance variables significantly predicting consumption with a p-value smaller than .01 in this subsample. The left side of the tree demonstrates that circumstances used in previous nodes can reappear – among those whose fathers did not complete secondary and whose mothers completed primary, the most important factor is whether the father completed primary. The "leaves" at the end of the tree represent the 6 types and their average daily consumption $\hat{C}$. This example is presented using a depth of 3, while in our estimation the maximum depth is set to 10.[11]

A disadvantage of regression trees is that small changes to the data might change the splitting points detected and hence the level of IOp estimated. To circumvent this issue, we use random forests in our preferred specification. These are collections of hundreds

[11] The parameters used in the estimation are detailed in the Model Appendix. As a default, we used the parameters used by a forthcoming IOp database led by LSE's International Inequalities Institute.

of trees each using a subsample of observations and a subsample of circumstances at each node of the tree, and the predictions are averaged across them. Refer to Brunori et al. (2023) for more details on this procedure.

Figure 2: Example Regression Tree



In addition to efficiently trading off upward and downward bias, this method has a few important advantages. First, it reduces arbitrary specification choices by the researcher. The circumstances included are no longer arbitrary because the model will automatically select only those that are most predictive of consumption inequality. Likewise, the researcher no longer needs to specify a functional form because the model determines the interactions between variables and the splitting points of non-binary variables. Second, by letting the data speak, conditional inference trees reveal the unique way in which circumstances explain inequality in consumption in each individual country, and this can be observed by studying the results of the regression tree.

As a final step, we break down the importance that each circumstance plays in predicting consumption. This is achieved by randomly permuting the values of a given circumstance and observing how much the accuracy of predicted consumption decreases on average, giving us a measure of the relative importance of that circumstance (Brunori et al. 2023). It is important to note that these variables may be correlated with other unobserved factors, and thus this should not be interpreted as identifying the causal importance of each circumstance, but rather as the effect of the circumstance and all other variables that may be correlated with it.

Section 3: Data

Table 1 presents the 18 surveys included in our analysis. This list comprises 12 countries in West Africa including Nigeria, four countries in East Africa, Gabon in Central Africa, and South Africa. These countries account for 59 percent of the population of Sub-Saharan Africa in 2020 and were identified as those with a consumption survey between 2012-2019 with both region of birth and parental education available. As our outcome, we use per-capita household consumption rather than income given that consumption is a more direct measure of well-being and income measurement in surveys is subject to more error (Carletto et al. 2022). We include the following circumstance variables: birth region, birth urban-rural status, parental education, parental industry of work, religion, and ethnicity (race in the case of South Africa). Birth region and parent education are available in all countries, while other circumstances are available in most but not all countries. Father and mother industry and ethnicity are available in 12 countries, while religion and urban-rural birthplace are available in 15 countries (this is presented in Table A1 in the Appendix). Our study sample contains people aged 15 or over in each country.

Table 1: List of Countries and Surveys Included in Analysis

Country	Year	Survey
Benin	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Burkina Faso	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Côte d'Ivoire	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Ethiopia	2018-2019	Ethiopia Socioeconomic Survey
Gabon	2017	Enquête Gabonaise pour l'Évaluation de la Pauvreté
Gambia, The	2015-2016	Integrated Household Survey
Ghana	2016-2017	Ghana Living Standard Survey
Guinea-Bissau	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Liberia	2016	Household Income and Expenditure Survey
Malawi	2019-2020	Integrated Household Survey
Mali	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Niger	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Nigeria	2018-2019	Living Standards Survey
Senegal	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
South Africa	2017	National Income Dynamics Study
Tanzania	2017-2018	Household Budget Survey
Togo	2018-2019	Enquête Harmonisée sur le Conditions de Vie des Ménages
Uganda	2012-2013	Uganda National Panel Survey

We exclude gender from the list of circumstances because the outcome, consumption, is measured at the household level and does not capture within-household resource allocation. In other words, all individuals within the household are assigned the same consumption per capita, and this by construction constrains variability in characteristics like gender that vary heavily within the household.[12]

Parental education is grouped into four categories defined the same way across countries – None, Primary, Secondary, and Tertiary – and treated as ordered to avoid the algorithm grouping non-consecutive categories (to prevent, for example, a type grouping none and secondary separately from primary and tertiary). Parental industry is also grouped into four categories defined the same way across countries – Agriculture, Industry, Services, and Other. For birth region, we use the administrative division that yields between 5-15 regions in each country.[13] Being born in another country is treated as a separate category. In some cases, we assume the respondent's previous location is their birthplace. In the case of South Africa, ethnicity is not available, and we instead use race, which takes four categories. Refer to the Data Appendix for details on assumptions made during harmonization.

Given the varying number of response categories across countries for birth region, ethnicity, and religion, caution is required when comparing IOp across countries. For example, the number of ethnic groups varies from 11 (Guinea-Bissau) to 66 (Gabon). This is not a problem for interpretation of IOp to the extent that it reflects true differences in ethnic diversity across countries. However, it may instead reflect differences in the granularity of survey responses.[14] Likewise, the number of religions varies from 4-8. Religions are separated into subclassifications in some countries (for example, Christian is sometimes divided into Catholic, Protestant, Other Christian, etc.), and as with ethnicity, this could reflect differences in surveying as opposed to differences in the relevance of those subclassifications in each country.

In some cases, circumstances are missing for some observations. An advantage of conditional inference trees is that they avoid dropping observations with missing circumstance information. Instead, where information on circumstances is missing, the

---

[12] As a result, IOp changes very little when we include gender as a circumstance. We also exclude age, which like gender varies heavily within the household, but furthermore because age is a characteristic that varies over the lifecycle.

[13] The exceptions are Malawi, which only has 3 regions, and Tanzania, which has 31 regions.

[14] A possible solution becoming increasingly popular in the literature is to link ethnic groups to a linguistic tree to aggregate them into comparable levels of linguistically defined ethnic divisions. However, this requires having an ethnic mapping specific to the surveys we use – for example, an ethnic group like "Asante" will have different meanings in different surveys. To our knowledge, such a mapping has not been created for the specific surveys we use in our analysis.

algorithm searches for an alternate splitting point that mimics the original sample partition as closely as possible (Brunori et al. 2023). Notably, birthplace is missing for around 30 percent in Nigeria and Togo, and father's and mother's education are missing for 55 percent and 40 percent respectively in Uganda. In other cases, smaller numbers of observations are missing. All instances of missing data are listed in Table A2.

Section 4: Results

IOp Estimates

The IOp estimates are presented in Table 2. The estimates from the conditional inference tree and the random forest are similar, which is expected given the large sample in most of our surveys. We focus on the random forest as our preferred specification.

Table 2: Relative IOp Estimates

Country	Sample Size	Gini	Types	Rel. IOp (Tree)	Rel. IOp (Forest)
South Africa	27,042	0.66	24	0.73	0.74
Burkina Faso	24,545	0.49	29	0.64	0.64
Niger	16,811	0.39	16	0.60	0.62
Togo	15,084	0.43	28	0.60	0.60
Benin	22,084	0.39	38	0.59	0.60
Senegal	36,975	0.39	36	0.57	0.58
Ghana	36,225	0.44	39	0.55	0.57
Mali	20,255	0.37	17	0.57	0.57
Guinea-Bissau	24,947	0.35	33	0.54	0.54
Gambia, The	55,286	0.37	30	0.54	0.54
Liberia	19,506	0.36	23	0.49	0.54
Uganda	9,194	0.46	14	0.50	0.53
Nigeria	63,646	0.36	20	0.50	0.50
Côte d'Ivoire	34,423	0.38	32	0.46	0.49
Tanzania	25,534	0.42	20	0.44	0.48
Malawi	28,388	0.4	13	0.44	0.45
Gabon	17,709	0.38	20	0.39	0.41
Ethiopia	17,563	0.41	14	0.24	0.26
Average		**0.41**	**25**	**0.52**	**0.54**

Note: countries sorted from highest to lowest Rel. IOp (Forest). The Gini Index presented here is calculated in each survey using the analysis sample of people aged 15 or over.

We find that, taking a simple average across countries, IOp in SSA is 54 percent (last column of Table 2). This is 15 percent higher than previously estimated, and it reveals that most inequality in SSA is explained by inherited circumstance (Brunori et al. 2019). For most countries, IOp ranges from 40 to 60 percent and the median IOp is also 54 percent. The range is also wider than previously found at both the top and bottom of the distribution. IOp is highest in South Africa (74 percent), Niger (62 percent) and Burkina Faso (64 percent), while the countries showing lowest IOp are Ethiopia (26 percent), Gabon (41 percent) and Malawi (45 percent). After South Africa, the countries with the highest IOp are all in West Africa, ranging from 49 percent in Côte d'Ivoire to 64 percent in Burkina Faso.

Importance of each circumstance

Table 3 presents the estimates of the role played by each circumstance using the variable importance method described in Section 2 with the values normalized to sum to 1 in each row and the average across all countries presented in the final row. The numbers have a relative interpretation. For example, when the importance statistic for father's education in Burkina Faso is 0.09 and for mother's education is 0.04, it means that excluding father's education from the random forest would reduce the accuracy of predicted consumption by more than twice as much as would be the case if mother's education was excluded from the random forest. Taking a simple average across countries gives a picture of the circumstances most relevant for predicting consumption. Region of birth is the most important circumstance followed by ethnicity, though this is driven heavily by race in South Africa. The next circumstances in terms of importance are urban-rural birthplace, and father's and mother's education.

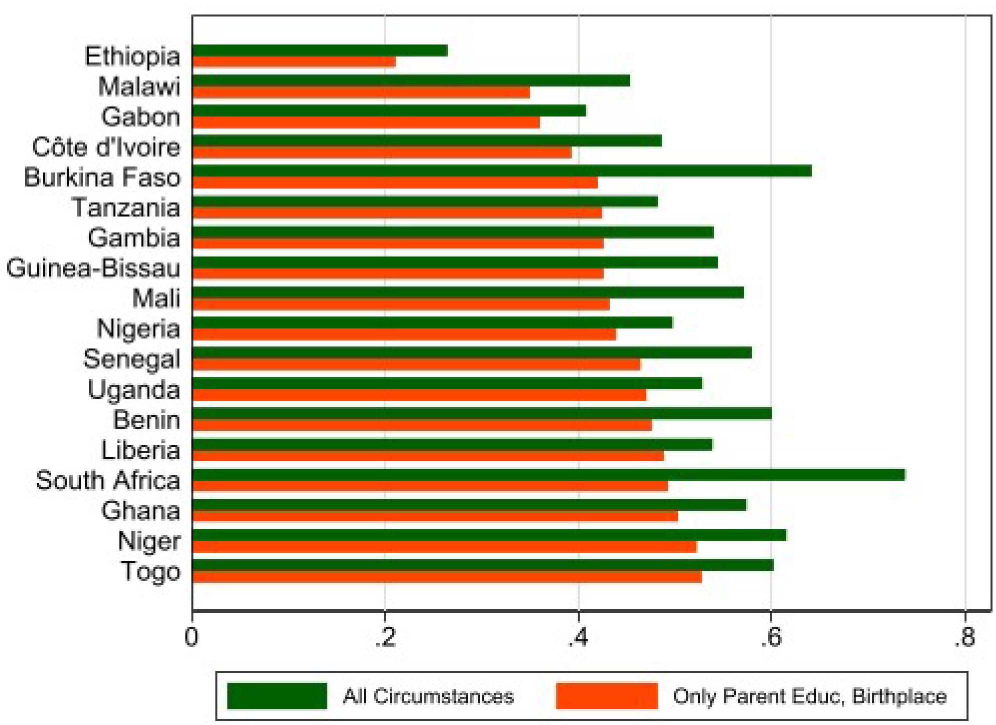
Table 3: Importance of Each Circumstance for Consumption

	Birth	Born	Father	Mother	Father	Mother	Religion	Ethnicity
South Africa	.03	.02	.02	.03			0	.90
Burkina Faso	.29	.34	.09	.04	.10	.03	.05	.06
Niger	.12	.37	.31	.09	.09	.02	0	
Togo	.12	.24	.17	.03	.08	.14	.08	.14
Benin	.26	.06	.17	.16	.06	.09	.04	.15
Senegal	.37	.15	.19	.06	.08	.02	.01	.11
Ghana	.33	.24	.12	.07			.05	.19
Mali	.29	.34	.20	.04	.12	.01	0	
Guinea-Bissau	.09	.25	.06	.15	.10	.08	.08	.19
Gambia, The	.26	.08	.04	.05	.05	.42	0	.10
Liberia	.19	.11	.16	.31				.24
Uganda	.15		.22	.33				.30
Nigeria	.48	.11	.11	.16			.14	
Côte d'Ivoire	.12	.22	.13	.09	.21	.06	.04	.14
Tanzania	.38	.17	.13	.25	.05	.02		
Malawi	.07	.25	.52	.16			0	
Gabon	.41		.11	.06	.04	.06	.16	.17
Ethiopia	.34		.19	.09	.12	.04	.23	
Average	**.24**	**.20**	**.16**	**.12**	**.09**	**.08**	**.06**	**.22**

There is some variation across countries in terms of which circumstances are more important drivers. Region of birth is most important in Nigeria where it is three times as important as any other circumstance. Father's education is most important in Malawi, while mother's education is most important in Liberia, Uganda, and Tanzania, and mother's industry in The Gambia. Race is more than ten-times as important as any other circumstance in South Africa and almost the most important circumstance in Guinea-Bissau, Liberia, and Uganda. Religion matters most in Ethiopia where it is the second-most important circumstance after birth region.

As discussed, cross-country comparisons are constrained by the fact that the availability of circumstance variables varies across countries. Therefore, we repeat the analysis using only region of birth and parental education as circumstances since these are measured in every country. We refer to this estimate as "comparable IOp" and results are presented in Figure 3. The estimates decrease moderately but remain high, ranging between 35-53 percent for all countries excluding Ethiopia. The ranking between countries is only loosely preserved. Notably, South Africa falls to having the 4th highest comparable IOp, and it decreases considerably for Burkina Faso as well. Excluding South Africa, the top 5 countries in terms of the share of inequality explained by these characteristics are all in West Africa. Ethiopia, Malawi, and Gabon remain the countries with the lowest IOp.

Figure 3: Share of Inequality Explained by Parental Education and Birthplace



The variable importance for comparable IOp, presented in Table 4, mirrors the results in Table 3, with birthplace being the most important followed by fathers' and mothers' education. These estimates are more comparable across countries because circumstances used across countries are the same. Region of birth now is at least twice as important as the other two circumstances Ghana, Guinea-Bissau, The Gambia, Nigeria, and Gabon. In The Gambia and Gabon, parental occupation, religion, and ethnicity were previously capturing a lot of this effect. Father's education continues to be most relevant in Malawi, and mother's education in Liberia, Uganda, and Tanzania.

Table 4: Variable Importance for Comparable IOp

Country	Region of Birth	Father Educ	Mother Educ
South Africa	.42	.34	.25
Burkina Faso	.54	.37	.10
Niger	.47	.41	.12
Togo	.61	.33	.07
Benin	.48	.32	.20
Senegal	.57	.33	.10
Ghana	.65	.20	.14
Mali	.55	.38	.07
Guinea-Bissau	.49	.23	.29
Gambia, The	.77	.13	.10
Liberia	.35	.26	.39
Uganda	.36	.29	.35
Nigeria	.64	.16	.20
Côte d'Ivoire	.45	.41	.14
Tanzania	.45	.17	.38
Malawi	.07	.64	.28
Gabon	.75	.17	.08
Ethiopia	.43	.38	.19
Average	**.50**	**.31**	**.19**

Comparable IOp uses only birth region and parental education as circumstances, which are available in all countries.

Regression trees

One of the most useful features of the methodology from the point of view of policy is that it produces opportunity trees that tell a story about IOp in the population. The trees provide a useful insight into the unique ways in which birth circumstances interact to create IOp in each country. They are also valuable to visualize the most disadvantaged groups in the population with great degree of granularity.[15] We include diagrams of all regression trees in the Tree Diagrams and Labels Appendix and discuss here a few examples for illustrative purposes.[16]

Using Senegal as an example, the circumstance that is most predictive of inequality in consumption is whether the person was born in Dakar or was foreign-born. Among those born in Dakar, the next important factor is whether the father has completed tertiary, while for those born outside Dakar, it is whether they were born in an urban or rural area. The tree reveals that the most prosperous type – with a daily per capita consumption of 16.3 USD in 2017 PPP – is someone born in Dakar and whose father completed tertiary. The least prosperous type – with a daily per capita consumption of 3.8 USD – is someone who was born outside Dakar, born in a rural area, and from a certain combination of regions (Kedougou, Kolda, Sedhiou, Tambacounda) and ethnic groups (Diola, Mandingue, Poular, and Other).[17]

In Table 5, we present the combination of circumstances that determine the most and least well-off groups in the five most populous countries in our study. In each case, the most advantageous type has a welfare that is orders of magnitude higher than the least-advantageous type. Whites with tertiary educated mothers have per capita consumption that is almost 13 times higher than Africans and Colored people with mothers with secondary education or less, born in the regions of Eastern Cape, KwaZulu-Natal or Mpumalanga. The first group makes up less than 2 percent of the population while the second group comprises 45 percent of the country.

[15] We caution against over-interpreting the partitions in the regression trees given that they are sometimes sensitive to small variations in the data and estimated model, which the random forest method addresses.
[16] For visualization purposes, the estimation for these diagrams uses a tree depth of 4 rather than 10, and ethnic groups that make up less than 3% of the population are grouped into an "other" category, thus reducing the influence of ethnicity in IOp but making it easier for the reader to identify major ethnic groups.
[17] The more detailed information in the regression tree will tend to mirror the results from the variable importance statistics presented in Table 3 – in Senegal, birth region was the most important circumstance, followed by father's education and urban-rural birthplace.

Table 5: Best and Worst Typology Description for Selected Countries

Country	Type	Description	Sample Share	Per-Cap Consumption
Nigeria	Best	Birth region = South South, South West, Other Country Father Education = Tertiary Mother Education = Tertiary	.02	7.6
	Worst	Birth region = North East, North West Religion = Islam, Traditional, Other Birth urban-rural = Rural	.29	2.5
Ethiopia	Best	Father education = Tertiary Birth Region = Addis, Amhara, Dire Dawa, Harari, Oromia, Other Country	.03	8
	Worst	Father education = None Birth Region = Afar, Gambela, Somali Father Industry = Agriculture Religion = Catholic, Protestant, Traditional	.03	2.6
Tanzania	Best	Birth region = Dar es Salaam, Arusha, Kilimanjaro, Other Country Mother education = Secondary or more	.01	11.9
	Worst	Birth region = Geita, Katavi, Kigoma, Lindi, Mwanza, Rukwa, Shinyanga, Simiyu, Songwe, Tabora Father education = Primary or less Mother education = None	.23	2.4
South Africa	Best	Race = White Mother Education = Tertiary	.01	80.7
	Worst	Race = African, Colored Mother education = Secondary or less Birth Region = Eastern Cape, KwaZulu-Natal, Mpumalanga	.45	6.3
Uganda	Best	Mother education = Secondary or more Father education = Tertiary	.02	14.2
	Worst	Mother education = Primary or less Ethnicity = Bagisu, Bakiga, Iteso, Langbi, Lugbara, Other Birth Region = Eastern, Northern	.37	2.8

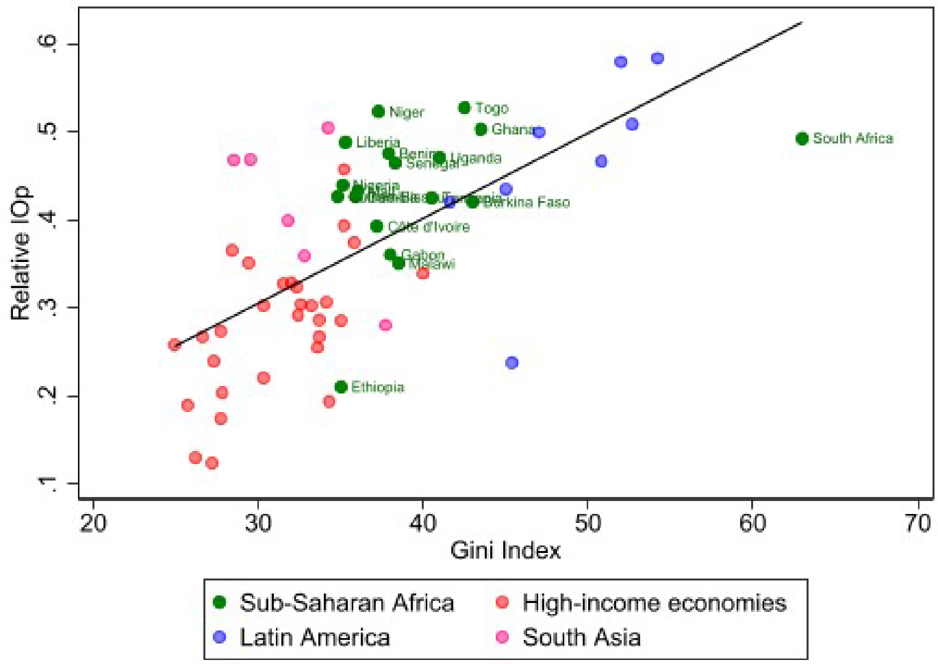
Per-capita consumption is measured in USD (2017 PPP)

Section 5: Comparison with Other Regions

In this paper, we have shown that IOp in SSA is higher than previously thought. In this section we place those findings in the global context. It is difficult to compare IOp in SSA with IOp other regions because of differences in methodology and circumstances used in existing global estimates. Nevertheless, Figure 4 shows that IOp in SSA is broadly in the same range as in Latin America and South Asia, and notably higher than in Europe and other high-income economies.

Figure 4 also confirms the well-known conclusion that IOp is positively correlated with inequality overall.[18] For this comparison, we choose to rely on our "comparable IOp" estimates, which are the IOp estimates based on only parental education and region of birth, because these circumstances are available in all countries in our sample, and the estimates we have selected from other regions use a similar set of circumstances.[19]

Figure 4: Correlation between Comparable IOp and Inequality



The IOp estimates for SSA are the "comparable" estimates that only use birth region and parental education. Gini is taken from the World Development Indicators in the survey year, and if survey year is not provided, the most recent year is used.

In Figure A2, we visualize the global relationship between IOp and output per capita. This relationship was previously found to follow an inverted-U relationship, in which IOp first increases and then decreases in output per capita (Bussolo et al. 2013). However, given that IOp in SSA is higher than previously thought, the relationship observed here is more linearly negative – countries that are wealthier per-capita also tend to have lower

[18] Because our analysis is limited to 18 countries, we chose not to devote too much attention to the relationship between IOp and other country characteristics specifically within SSA.

[19] IOp estimates for Latin America and high-income economies come from *The World Database on Equality of Opportunity and Social Mobility*, which was accessed from www.equalchances.org in December 2023, and are based on a non-parametric regression method, use data from 2008-2015, and use birthplace, parental education, and parental occupation as circumstances. IOp estimates from South Asia are from Bussolo et al. (2023), are measured between 2011-2019, and are based on parental education and current region.

IOp.[20] IOp and output per capita also has a stronger negative relationship than total inequality and output-per-capita, consistent with existing evidence that IOp is relatively more important for growth than total inequality (Marrero et al. 2013, Bradbury et al. 2016, Carranza 2020).

Finally, in Figure A3, we show the relationship between comparable IOp and intergenerational mobility (IGM) calculated for our countries of analysis by van der Weide et al. (2024). We measure intergenerational mobility as 1 minus the correlation coefficient between respondents' and parents' years of schooling. Globally there is a negative relationship between IGM and IOp, typically seen as driven by the fact that IOp decreases as parental socioeconomic characteristics become a less relevant predictor of income inequality (Brunori et al., 2013). This pattern persists as our estimates from SSA are included with global data.[21] However, the negative global correlation between IOp and IGM is relatively moderate (R^2=.45), highlighting the value of IOp as a standalone measure of inherited inequality that carries additional information.

Section 6: Conclusion

In this paper, we estimated IOp across 18 countries in SSA (the largest sample among existing studies on IOp in Africa) using the most recently available surveys, a large set of circumstances, and a machine-learning method that trades off upward and downward bias, reduces the need for arbitrary specification choices, and reveals the unique ways in which birth characteristics explain inequality across countries. The results show that more than half of inequality in SSA is explained by circumstances at birth. On average, IOp explains 54 percent of overall inequality in consumption, which is 15 percent higher than previous estimates in the region.

The outsized role of conditions at birth in determining lifelong income is concerning from a moral perspective, but also because existing evidence indicates that IOp is a more important determinant of growth than total inequality (Marrero et al. 2013, Bradbury et al. 2016, Carranza 2020). There is extensive variation across countries in the level of IOp and the role played by different circumstances, highlighting the unique challenges that countries face in addressing IOp. In general, birthplace, parental education, and ethnicity

[20] It goes without saying that these are simply correlations and do not capture the causal relationship between IOp and income.

[21] The relationship is similarly negative if we measure IGM as the share of people with a strictly higher educational category than their parents excluding those whose parents have tertiary education.

tend to be the most important factors, while South Africa stands out as a country where race plays an outsized role in explaining inequality.

These results underscore the pressing need for policy makers to intensify efforts to create more equal access to opportunities. IOp reflects disparities that begin in early childhood and continue to accumulate later in life. Therefore, there are multiple entry points for policy makers to address IOp, including for children as they build their productive capacities, for adults participating in income earning activities, as well as for retirees who provide important services at home for their families and communities.

The current evidence indicates that public funds spent on early childhood interventions, basic education, and infant and child health have high returns and are crucial for children to reach their productive potential later in life (Narayan et al. 2018). In order to accelerate this positive impact and reduce inequalities, direct health and education spending should be prioritized towards the poor. In education, for example, this may imply better targeting educational investment towards lagging regions and disadvantaged populations. Policy makers can improve access to pre-primary enrollment in rural and poor communities through low-cost, community-based approaches with high quality standards (Bashir et al. 2018).

Expansion of basic services such as water, sanitation, and electricity should also prioritize disadvantaged populations. This will not only help to address the issues of coverage and inequality, but also to address budget constraints widespread in African countries. According to Foster and Briceño-Garmendia (2010), achieving universal access to services such as piped water and electricity in Africa has been impeded by an overreliance on subsidized services. The authors suggest focusing on connecting people to existing infrastructure networks, increasing coverage by focusing on underserved populations living near existing networks, and replacing one-time upfront connection charges with long-term financing options like general tariffs.

Policies for income-earning adults may include, among others, investing in skills building, technical, and vocational education programs for women and disadvantaged youth, adopting and enforcing labor standards, eliminating discriminatory labor laws, and provision of childcare. Fiscal policies can also play an important role, including more progressive personal income and property taxation, targeted adaptative social assistance to provide protection against shocks, and replacing fertilizer subsidies with public goods intended to limit environmental degradation and improve yields in areas with depressed agriculture.

Reducing unequal access to opportunities in SSA will require a coordinated and holistic approach with policies and interventions to help disadvantaged individuals in all stages of life. This is a daunting task, but necessary to foster societies that are more equitable and to unlock the full potential for growth in the region.

Bibliography

Aaberge, R., Mogstad, M., & Peragine, V. (2011). Measuring Long-Term Inequality of Opportunity. *Journal of Public Economics*, *95*(3-4), 193-204.

Bashir, S., Lockheed, M., Ninan, E., & Tan, J. P. (2018). *Facing forward: Schooling for learning in Africa*. World Bank Publications.

Bradbury, K., & Triest, R. K. (2016). Inequality of Opportunity and Aggregate Economic Performance. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *2*(2), 178-201.

Brunori, P., Ferreira, F. H., & Peragine, V. (2013). Inequality of Opportunity, Income Inequality, and Economic Mobility: Some International Comparisons. In *Getting Development Right: Structural Transformation, Inclusion, and Sustainability in the Post-Crisis Era* (pp. 85-115). New York: Palgrave Macmillan US.

Brunori, P., Peragine, V., & Serlenga, L. (2018). Upward and Downward Bias when Measuring Inequality of Opportunity. *Social Choice and Welfare*, *52*, 635-661.

Brunori, P., Palmisano, F., & Peragine, V. (2019). Inequality of Opportunity in Sub-Saharan Africa. *Applied Economics*, *51*(60), 6428-6458.

Brunori, P., Hufe, P., & Mahler, D. (2023). The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests. *The Scandinavian Journal of Economics*, 125(4), 900-932.

Bussolo, M., Peragine, V., & Reutzel, F. (2023). Inequality of Opportunity in South Asia. *Mimeo.*

Carletto, G., Tiberti, M., & Zezza, A. (2022). Measure for Measure: Comparing Survey-Based Estimates of Income and Consumption for Rural Households. *The World Bank Research Observer*, 37(1), 1-38.

Carranza, R. (2020). *Inequality of outcomes, inequality of opportunity, and economic growth* (No. 534). ECINEQ, Society for the Study of Economic Inequality.

Ferreira, F. H. (Ed.). (2005). *World Development Report 2006: Equity and Development* (Vol. 28). World Bank Publications.

Ferreira, F. H., & 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., Lakner, C., Lugo, M. A., & Özler, B. (2018). Inequality of Opportunity and Economic Growth: How Much Can Cross-Country Regressions Really Tell Us? *Review of Income and Wealth*, *64*(4), 800-827.

Foster, V., & Cecilia. Briceño-Garmendia. (2010). *Africa's infrastructure: a time for transformation*. Agence Française de Développement.

Kanbur, R., & Wagstaff, A. (2016). How Useful is Inequality of Opportunity as a Policy Construct? In *Inequality and Growth: Patterns and Policy: Volume I: Concepts and Analysis* (pp. 131-150). London: Palgrave Macmillan UK.

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

Narayan, A., Van der Weide, R., Cojocaru, A., Lakner, C., Redaelli, S., Mahler, D. G., ... & Thewissen, S. (2018). *Fair progress? Economic mobility across generations around the world*. World Bank Publications.

Palmisano, F., & Peragine, V. (2022). Inequality of Opportunity: Theoretical Considerations and Recent Empirical Evidence. In *Advances in Economic Measurement: A Volume in Honour of DS Prasada Rao* (pp. 349-386). Singapore: Springer Nature Singapore.

Roemer, J. E. (1993). A Pragmatic Theory of Responsibility for the Egalitarian Planner. *Philosophy & Public Affairs*, 146-166.

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

Tetteh-Baah, S., K. Harttgen, R. Lahoti, I. Gunther. (2024). Horizontal Inequalities in Africa. *Review of Income and Wealth* (Early View).

Thorbecke, E. (2023). The Interrelationships Among Growth, Inequality and Poverty in Africa. *Journal of African Economies* 32, ii81–ii86.

Van De Gaer, D. (1993). Equality of Opportunity and Investment in Human Capital.

Van der Weide, R., Lakner, C., Mahler, D. G., Narayan, A., & Gupta, R. (2024). Intergenerational Mobility Around the World: A New Database. *Journal of Development Economics*, *166*, 103167.

World Bank. (2022a). *Poverty and Shared Prosperity 2022: Correcting Course*. The World Bank.

World Bank. (2022b). *Inequality in Southern Africa: An Assessment of the Southern African Customs Union*. The World Bank.

World Bank. (2023). *Expanding Opportunities: Toward Inclusive Growth.* The World Bank.

Wu, H., Atamanov, A, Bundervoet, T. and Paci, P. (2024). *The Growth Elasticity of Poverty: Is Africa Any Different?* Policy Research Working Paper Series 10690. The World Bank.

Data Appendix

Region of Birth: It was not always known where every individual was born. In some cases, if the respondent had ever moved, then only their previous location is known, and this was assumed to be their location of birth (up to 25% in 8 countries). In other cases, the precise region was not known and instead we were able to find out whether the location was rural or urban outside the capital (up to 30% in 7 countries).

Born urban: This was constructed analogously to region of birth and indicates whether the birthplace was rural or urban.

Parental Education: This was compiled from information on household members when the parent is alive and in the household, and from questions on parental characteristics when the parent is dead or not in the household. The categories used are No Education, Primary (complete or incomplete), Secondary (complete or incomplete), Tertiary (complete or incomplete). We do not use finer education groups because for many countries they could not be constructed for parents outside the household. To maintain consistency across countries and between parents in and out of the household, the following categories are also classified as "No Education" in cases where they are observed: Adult Education, Literacy, Quranic Education.

Parental Industry: This is constructed analogously to parental education. It is missing when the parent did not work, hence the larger number of missing observations for these variables.

Ethnicity: In South Africa, race is used instead of ethnicity.

Religion: If there were small categories of religions with around .1% of the population or less, then these were combined with the "Other" category. This only affected 5 countries.

Table A1: Circumstance Availability by Country

Country	Region of Birth	Born Urban	Father Educ	Mother Educ	Father Industry	Mother Industry	Religion	Ethnicity
Benin	X	X	X	X	X	X	X	X
Burkina Faso	X	X	X	X	X	X	X	X
Côte d'Ivoire	X	X	X	X	X	X	X	X
Ethiopia	X		X	X	X	X	X	
Gabon	X		X	X	X	X	X	X
Gambia, The	X	X	X	X	X	X	X	X
Ghana	X	X	X	X			X	X
Guinea-Bissau	X	X	X	X	X	X	X	X
Liberia	X	X	X	X				X
Malawi	X	X	X	X			X	
Mali	X	X	X	X	X	X	X	
Niger	X	X	X	X	X	X	X	
Nigeria	X	X	X	X			X	
Senegal	X	X	X	X	X	X	X	X
South Africa	X	X	X	X			X	X (race)
Tanzania	X	X	X	X	X	X		
Togo	X	X	X	X	X	X	X	X
Uganda	X		X	X				X

Table A2: Share of Missing Observations by Country

Country	Region of Birth	Born Urban	Father Educ	Mother Educ	Father Industry	Mother Industry	Religion	Ethnicity
Benin	0.20	0	0.02	0.01	0.04	0.08	0	0.05
Burkina Faso	0.19	0	0.01	0.01	0.07	0.26	0	0
Côte d'Ivoire	0.13	0	0.03	0.01	0.05	0.3	0	0.18
Ethiopia	0		0.08	0.07	0.07	0.51	0.03	
Gabon	0		0.18	0.1	0.34	0.59	0	0.13
Gambia, The	0	0	0.01	0	0	0	0	0.04
Ghana	0.02	0.34	0.05	0.03			0	0.02
Guinea-Bissau	0.05	0	0.07	0.05	0.11	0.19	0	0.02
Liberia	0	0.53	0.03	0.01				0
Malawi	0	0	0.1	0.06			0	
Mali	0.17	0	0.05	0.07	0.23	0.72	0	
Niger	0.05	0	0.01	0	0.14	0.48	0	
Nigeria	0.31	0.37	0.03	0.02			0	
Senegal	0.21	0	0.05	0.03	0.2	0.53	0	0.01
South Africa	0.05	0.42	0.24	0.12			0.12	0
Tanzania	0	0.32	0.08	0.05	0.22	0.29		
Togo	0.31	0	0.02	0.01	0.03	0.13	0	0.04
Uganda	0		0.55	0.4				0

The share of the sample missing information ranges from 0% to 72% (mother industry in Mali).

Appendix Figures

Figure A1: Random Forest IOp Estimates using Gini and MLD

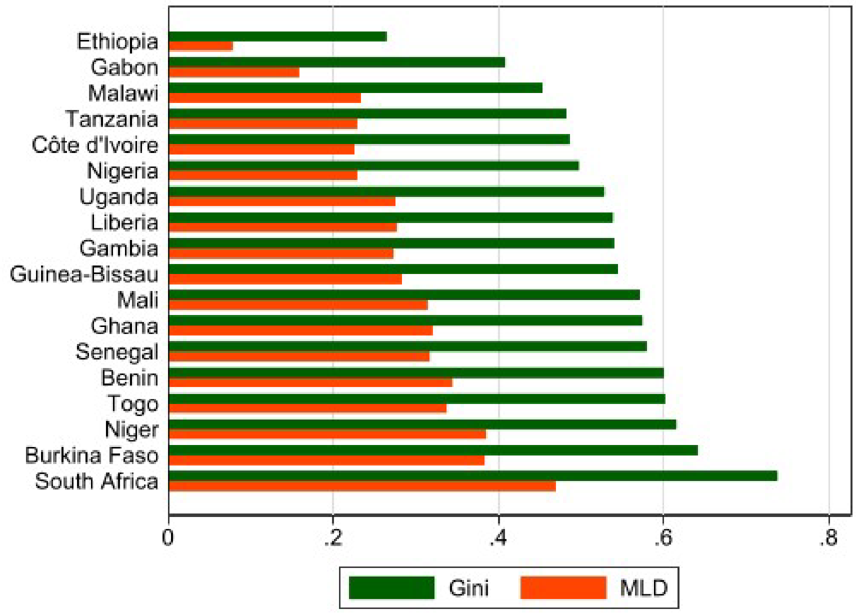
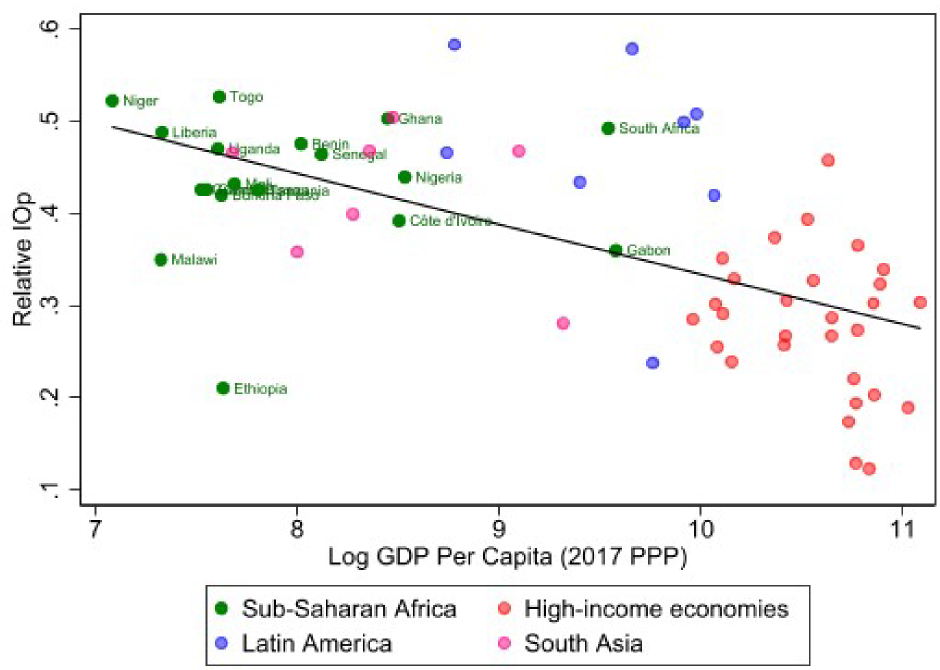
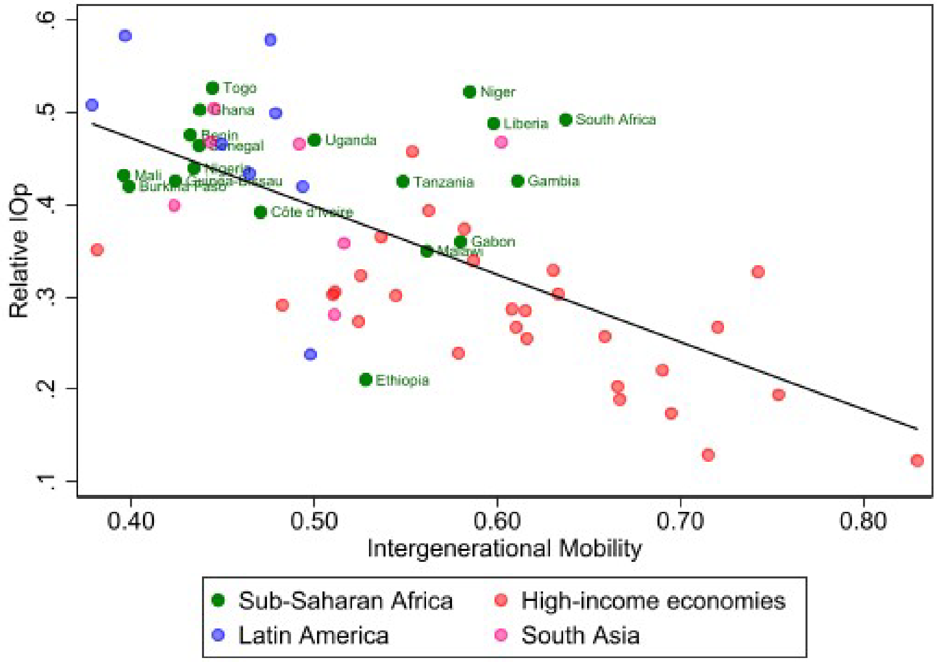


Figure A2: Correlation between Comparable IOp and GDP Per Capita



The IOp estimates for SSA are the "comparable" estimates that only use birth region and parental education. Refer to the paper for the sources of IOp in other regions. GDP per capita is taken from the World Development Indicators in the survey year, and if survey year is not provided, the most recent year is used.

Figure A3: Correlation between Comparable IOp and Intergenerational Mobility



The IOp estimates for SSA are the "comparable" estimates that only use birth region and parental education. Refer to the paper for the sources of IOp in other regions. Intergenerational mobility is 1 minus the correlation coefficient between respondents' and parents' years of schooling. Estimates are taken from van der Weide et al. (2024) and averaged for the 1980s cohort (our sample includes the 1950s-1980s cohorts, but information on IGM for pre-1980s cohorts is missing in many countries).

Model Parameters

	Tree	Forest
Mincriterion (1 – p-value used to make splits)	.99	0
Minbucket (minimum number of observations allowed in a terminal node)	1% of sample	0.1% of sample
Maxdepth (maximum depth of the tree)	10	-
Number of trees	-	200
Mtry (number of circumstances considered at each split point)	-	.9*Number of Circumstances
Fraction (fraction of observations used in each tree)	-	.5

Appendix Tree Diagrams and Labels

Labels common across countries:

Born Urban Rural	Parent Education	Parent Industry
0 Rural 1 Urban 96 Unknown 98 Other country	1 Tertiary 2 Secondary 3 Primary 4 None	1 Agriculture 2 Industry 3 Services 4 Other

For the estimation used to construct these tree diagrams, ethnicities that make up <3% of the population have been combined into an "Other" category. This may lessen the role of ethnicity in IOp but it allows the reader to clearly see the role of major ethnic groups. This is only for expositional purposes and the results presented in the paper are based on all ethnic groups as defined by each survey.

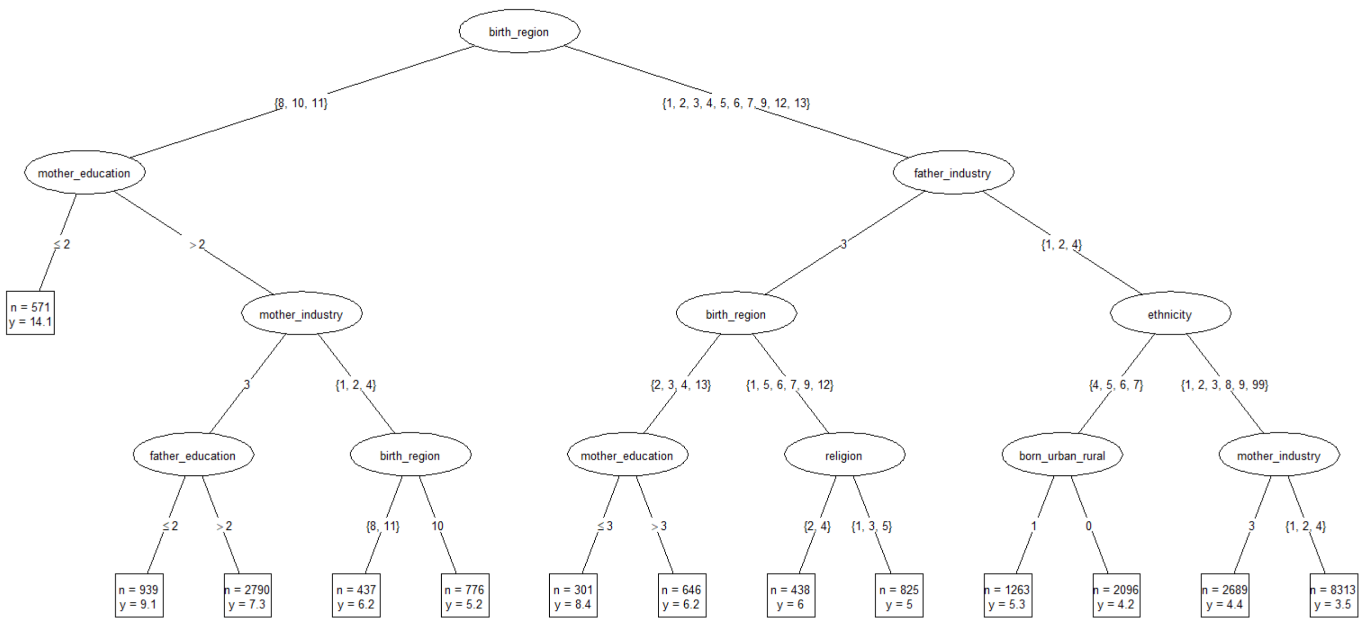
These diagrams use a maximum depth of 4 whereas the estimation in the paper uses a maximum depth of 10.

n = number of observations

y = average daily per-capita consumption (2017 PPP-adjusted USD)

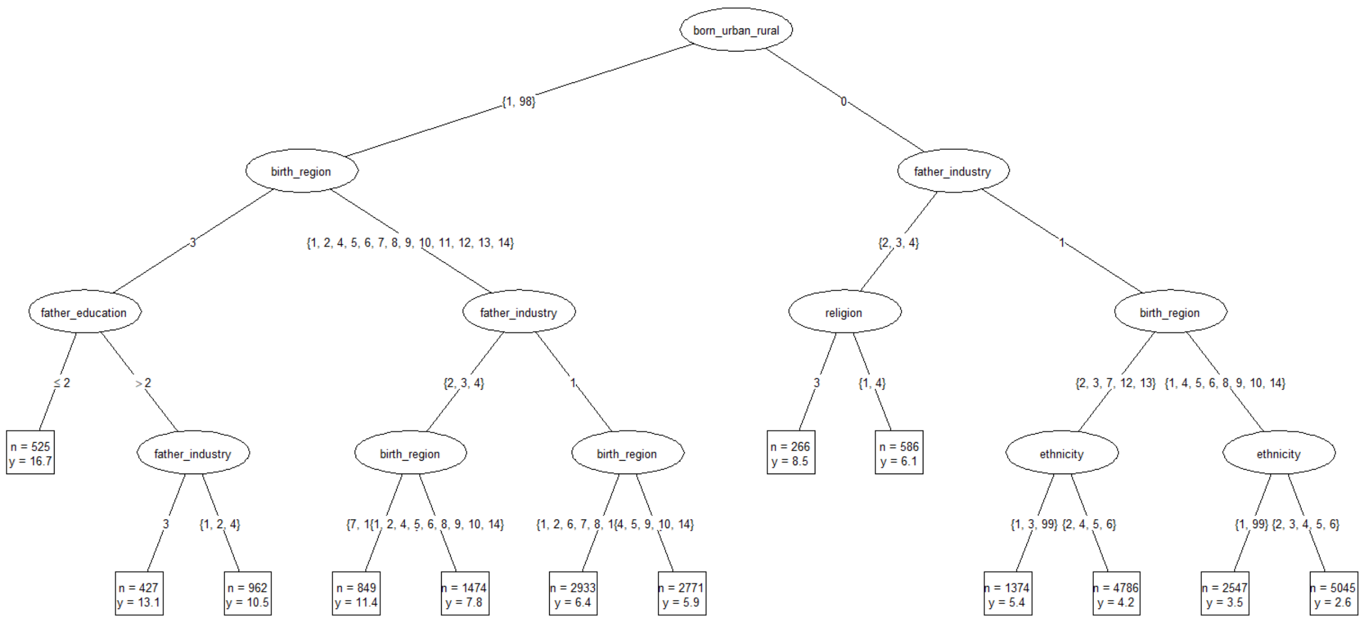
(Note – Higher consumption groups are split to the left, lower consumption groups to the right)

Benin



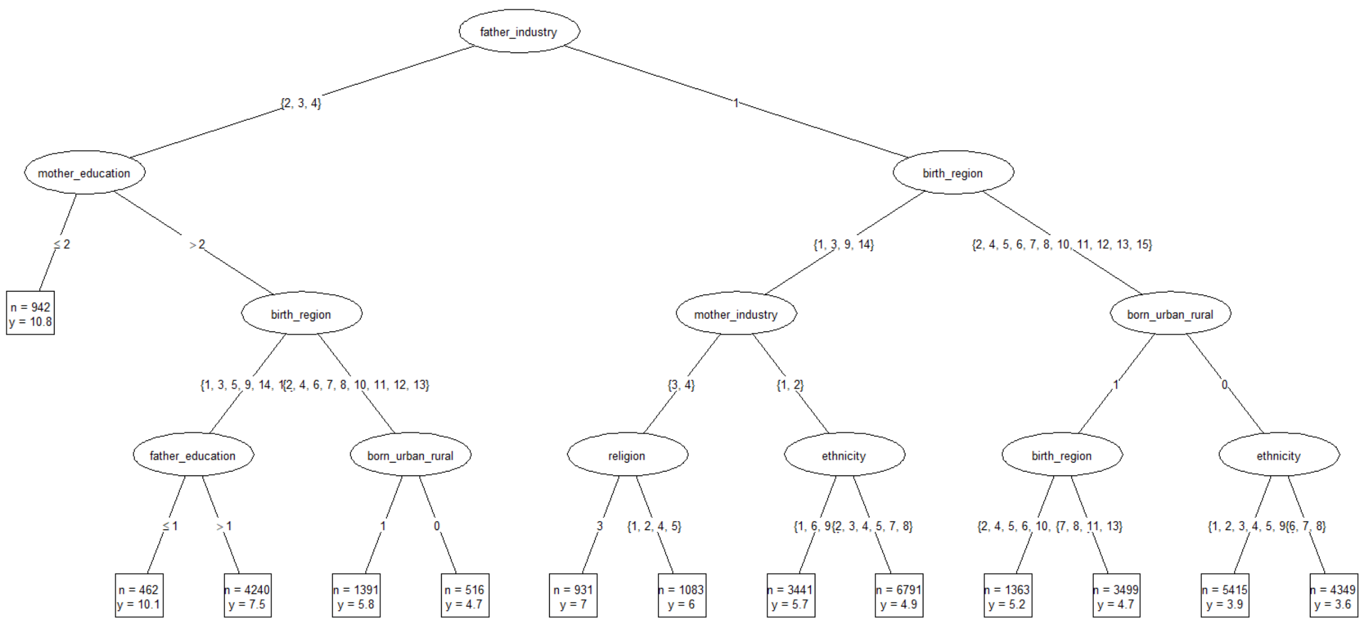
Ethnicity	Religion	Birth Region
1 Adja 2 Aïzo 3 Bariba 4 Fon 5 Goun 6 Mahi 7 Nago 8 Peulh 9 Sahouè 99 Other	1 Animiste 2 Autre Réligion 3 Chrétien 4 Musulman 5 Sans Réligion	1 Alibori 2 Atacora 3 Atlantique 4 Borgou 5 Collines 6 Couffo 7 Donga 8 Littoral 9 Mono 10 Oueme 11 Other Country 12 Plateau 13 Zou

Burkina Faso



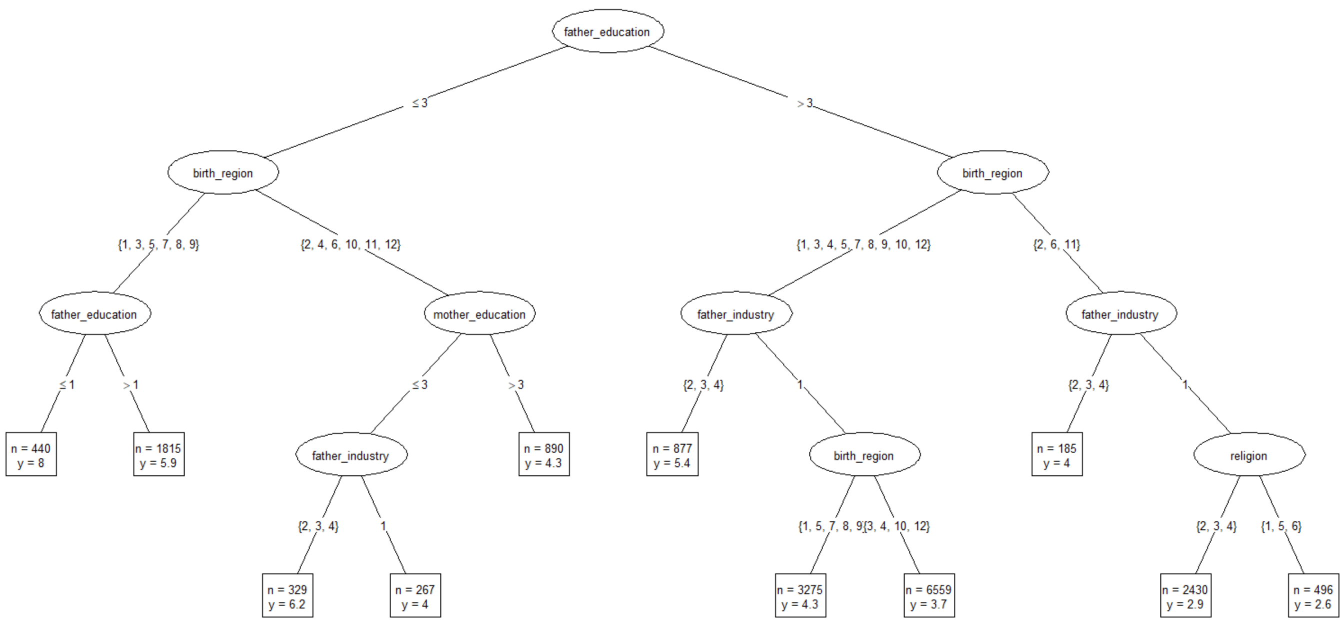
Ethnicity	Religion	Birth Region
1 Bobo	1 Animiste	1 Boucle du Mouhoun
2 Dagara	2 Autre Réligion	2 Cascades
3 Gourmatché	3 Chrétien	3 Centre
4 Mossi	4 Musulman	4 Centre-Est
5 Peulh	5 Sans Réligion	5 Centre-Nord
6 Senoufo		6 Centre-Ouest
99 Other		7 Centre-Sud
		8 Est
		9 Hauts Bassins
		10 Nord
		11 Other Country
		12 Plateau-Central
		13 Sahel
		14 Sud-Ouest

Côte d'Ivoire



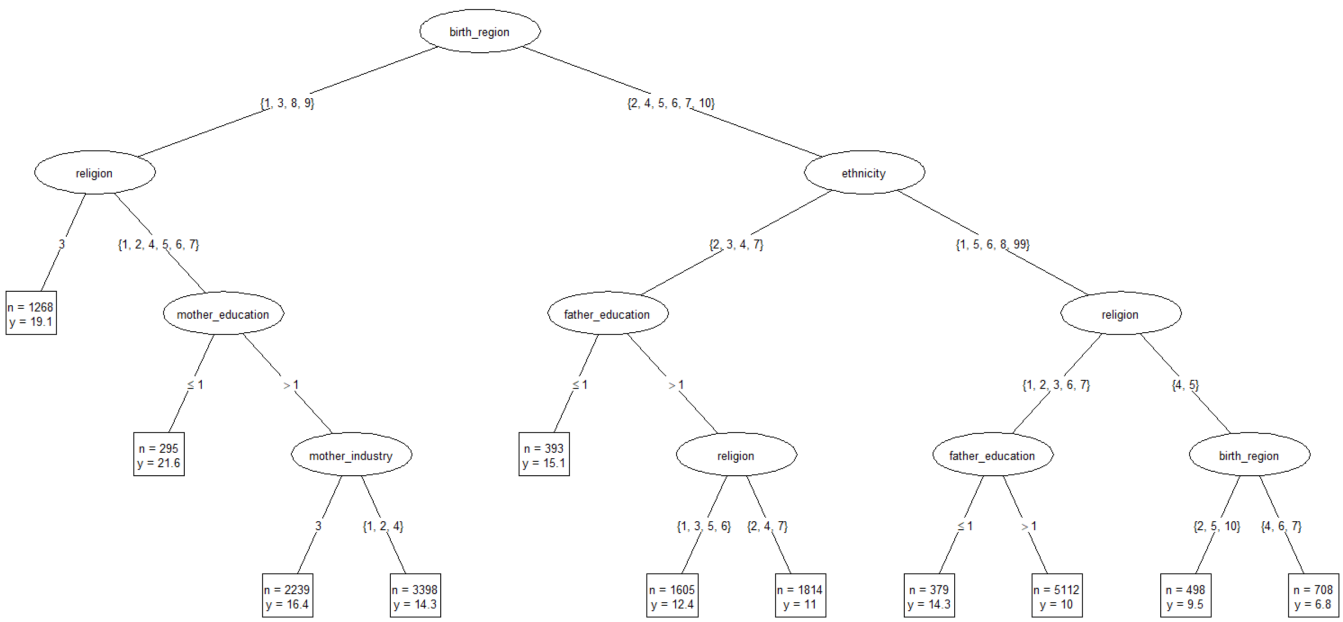
Ethnicity	Religion	Birth Region
1 AGNI	1 Animiste	1 Autonome D'Abidjan
2 BAOULE	2 Autre Réligion	2 Bas-Sassandra
3 DIOULA	3 Chrétien	3 Comoé
4 KOYAKA OU KOYARA	4 Musulman	4 Denguélé
5 LOBI	5 Sans Réligion	5 Gôh-Djiboua
6 MALINKE OU MANINKA		6 Lacs
7 SENOUFO		7 Lagunes
8 YACOUBA OU DAN		8 Montagnes
99 Other		9 Other Country
		10 Sassandra-Marahoué
		11 Savanes
		12 Vallée du Bandama
		13 Woroba
		14 Yamoussoukro
		15 Zanzan

Ethiopia



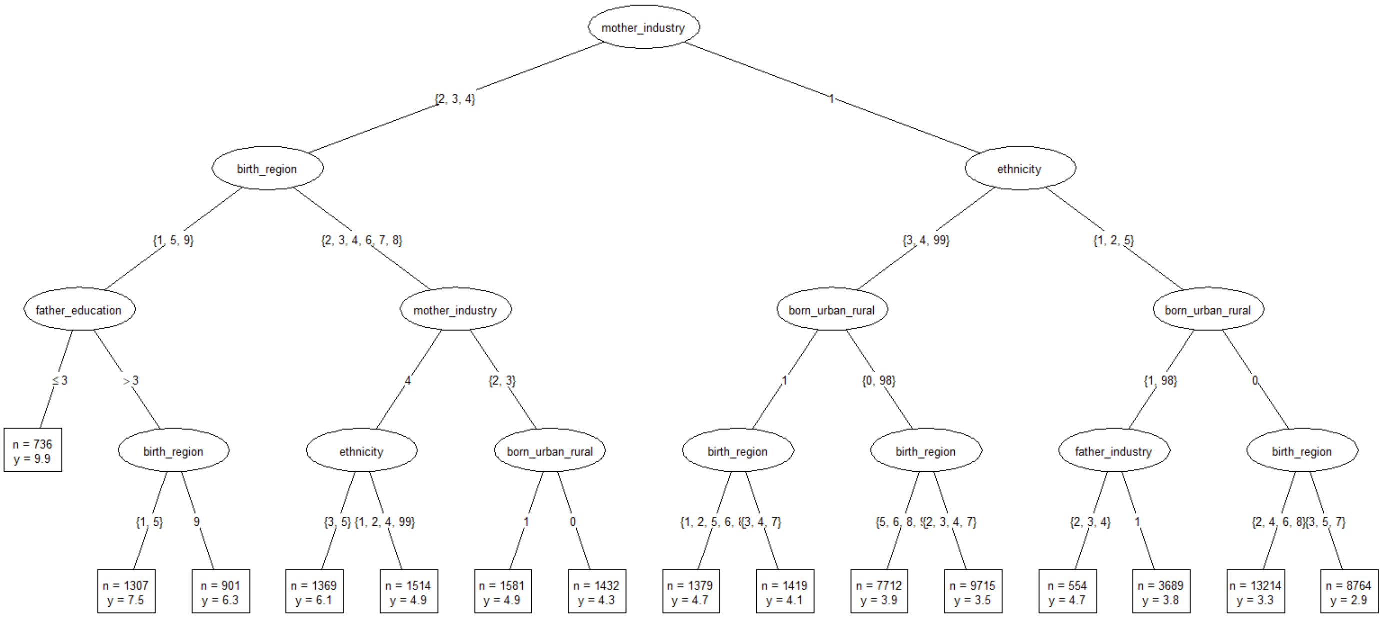
Religion	Birth Region
1 Catholic	1 Addis Ababa
2 Muslim	2 Afar
3 Orthodox	3 Amhara
4 Other	4 Benishangul-Gumuz
5 Protestant	5 Dire Dawa
6 Traditional	6 Gambela
	7 Harari
	8 Oromia
	9 Other Country
	10 SNNP
	11 Somali
	12 Tigray

Gabon



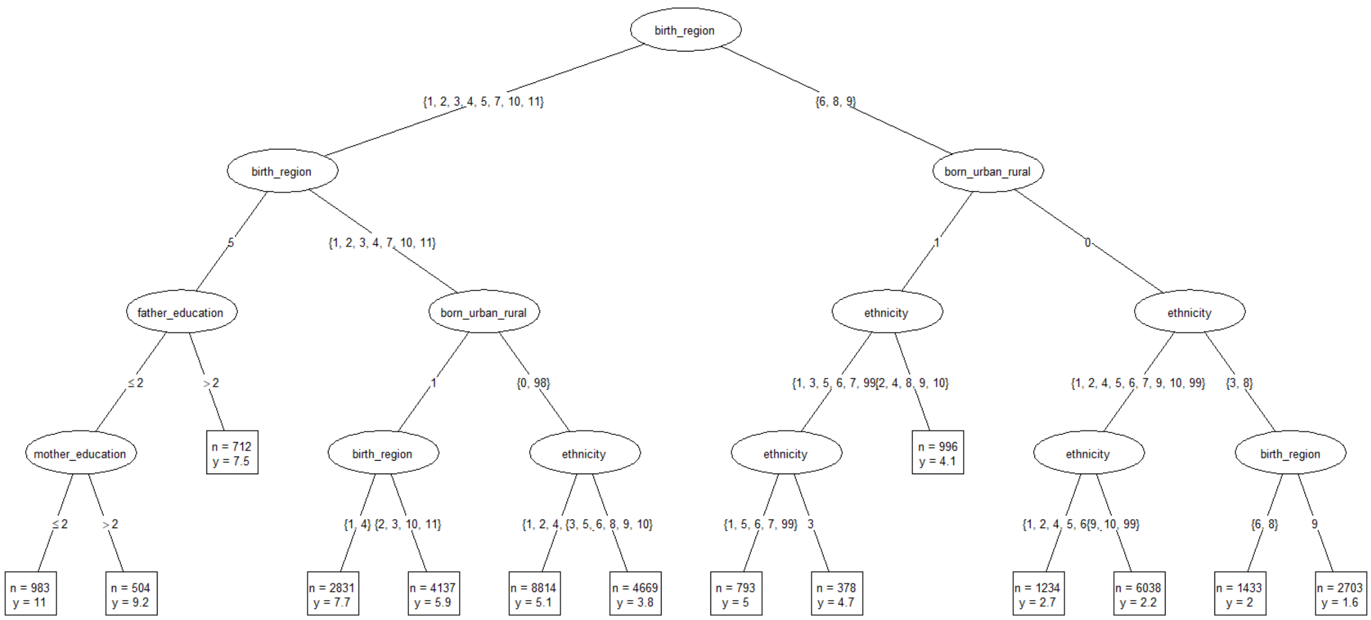
Ethnicity	Religion	Birth Region
1 Ambamba-Obamba 2 Ashira-Shira-Eshira-Guisir 3 Fang 4 Mbede-Teke 5 Nzabi-Nzebi 6 Oteghe-Bateke 7 Punu 8 Tsogho-Mitsogho 99 Other	1 Catholic 2 Church of Revival 3 Muslim 4 No Religion 5 Other 6 Other Christian religion 7 Protestant	1 Estuaire 2 Haut-Ogooué 3 Moyen-Ogooué 4 Ngounié 5 Nyanga 6 Ogooué-Ivindo 7 Ogooué-Lolo 8 Ogooué-Maritime 9 Other Country 10 Woleu-Ntem

Gambia, The



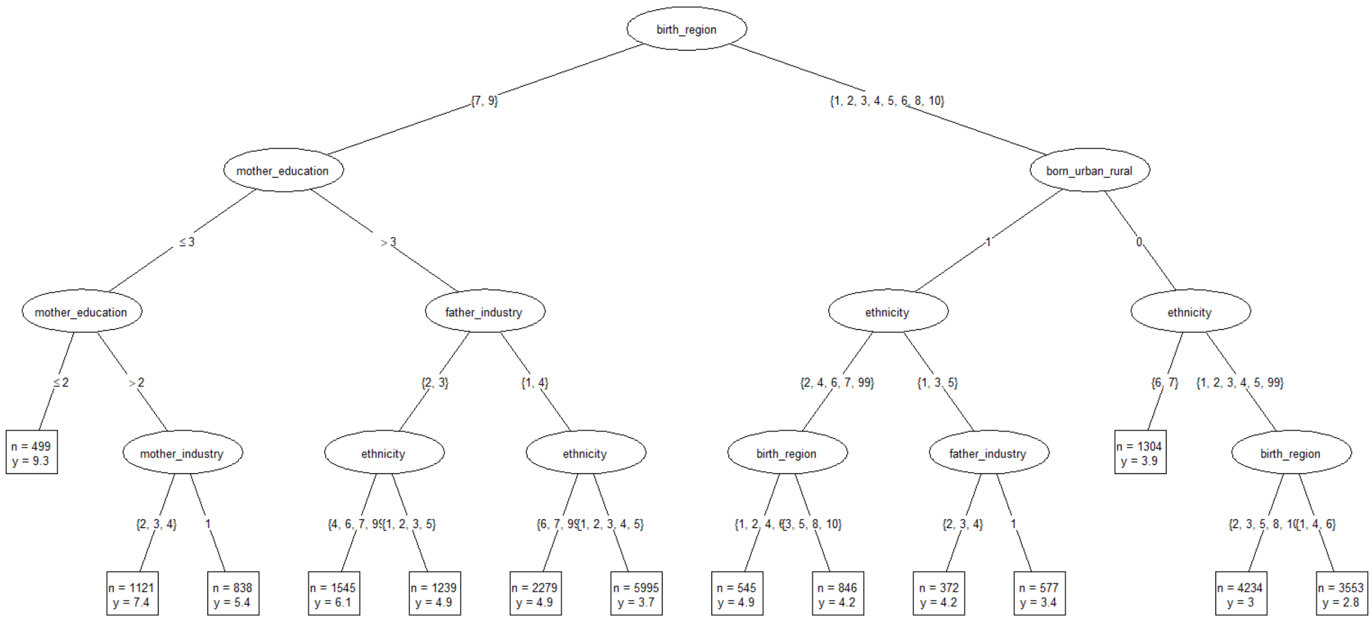
Ethnicity	**Religion**	**Birth Region**
1 Fula/tukulur/lorobo 2 Jola/karoninka 3 Mandinka/jahanka 4 Serahulleh 5 Wollof 99 Other	1 Christianity 2 Islam 3 Other 4 Traditional	1 Banjul 2 Basse 3 Brikama 4 Janjanbureh 5 Kanifing 6 Kerewan 7 Kuntaur 8 Mansa Konko 9 Other Country

Ghana



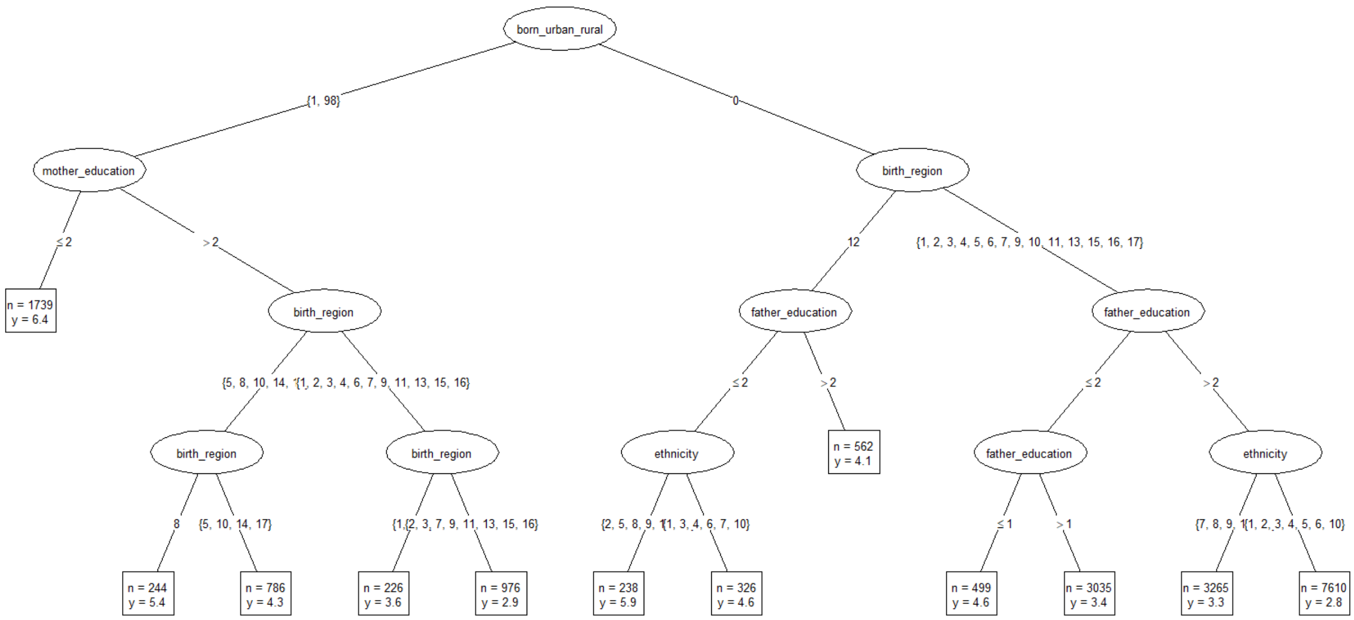
Ethnicity	Religion	Birth Region
1 Asante 2 Boron (Brong) (including Banda) 3 Dagarte (Dagaba), Lobi , Wali (Wala) 4 Dagomba 5 Dangme (Ada, Shai, Krobo, Osudoku,Ningo) 6 Ewe 7 Fante 8 Kokomba 9 Kusasi 10 Nankansi, Talensi & Gurense (Frafra) 99 Other	1 Catholic 2 Islam 3 No religion 4 Other 5 Other Christian 6 Pentecostal/Charismatic 7 Protestant 8 Traditionalist	1 Ashanti 2 Brong Ahafo 3 Central 4 Eastern 5 Greater Accra 6 Northern 7 Other Country 8 Upper East 9 Upper West 10 Volta 11 Western

Guinea-Bissau



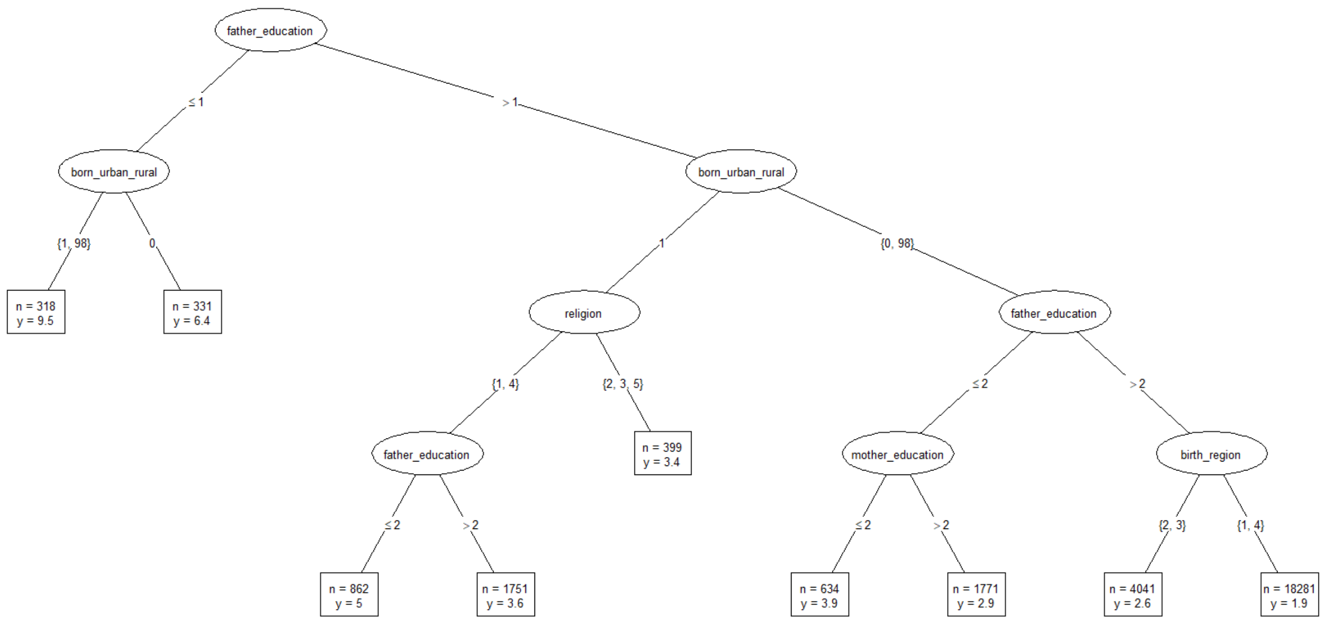
Ethnicity	Religion	Birth Region
1 Balanta	1 Animista	1 Bafata
2 Beafada	2 Cristão(ã)	2 Biombo
3 Bijagos	3 Muçulmano(a)	3 Bolama Bijagos
4 Fula	4 Outra religião	4 Cacheu
5 Mandinga	5 Sem religião	5 Gabu
6 Manjaco		6 Oio
7 Papel		7 Other Country
99 Other		8 Quinara
		9 Sab
		10 Tombali

Liberia



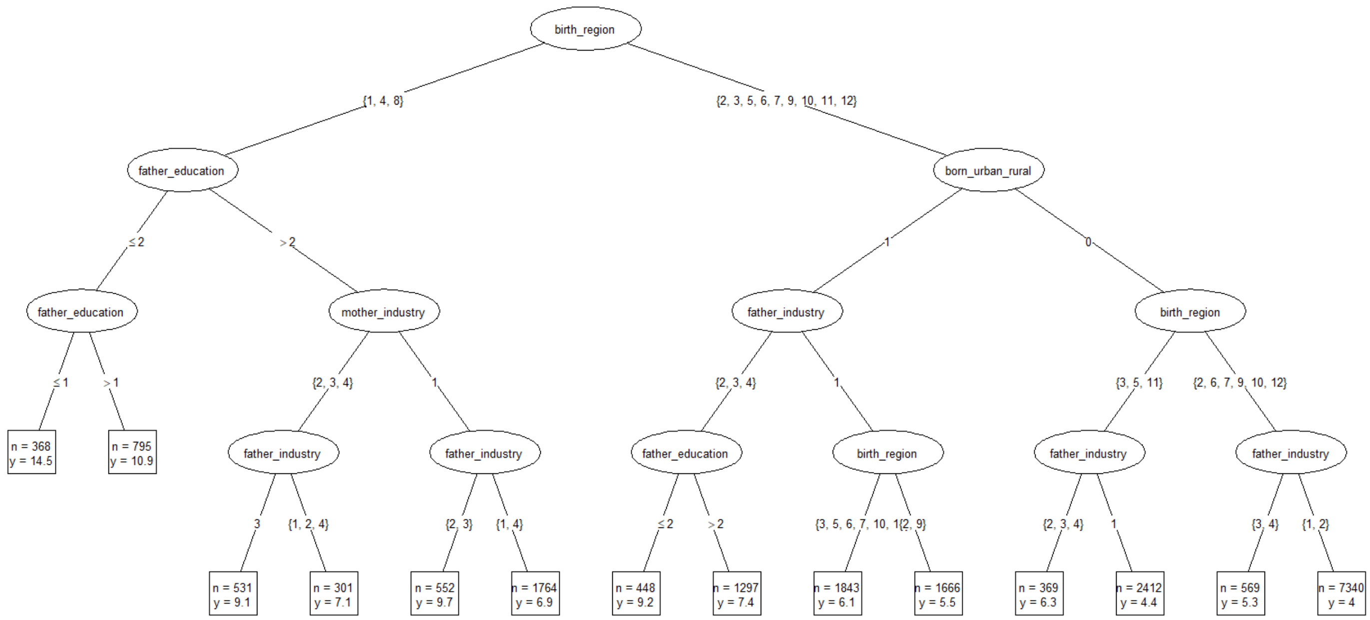
Ethnicity	**Birth Region**
1 Bassa 2 Gio 3 Gola 4 Grebo 5 Kissi 6 Kpelle 7 Krahn 8 Kru 9 Lorma 10 Mano 11 Vai 99 Other	1 Bomi 2 Bong 3 Gbarpolu 4 Grand Bassa 5 Grand Cape Mount 6 Grand Gedeh 7 Grand Kru 8 Greater Monrovia 9 Lofa 10 Margibi 11 Maryland 12 Montserrado 13 Nimba 14 Other Country 15 River Cess 16 River Gee 17 Sinoe

Malawi



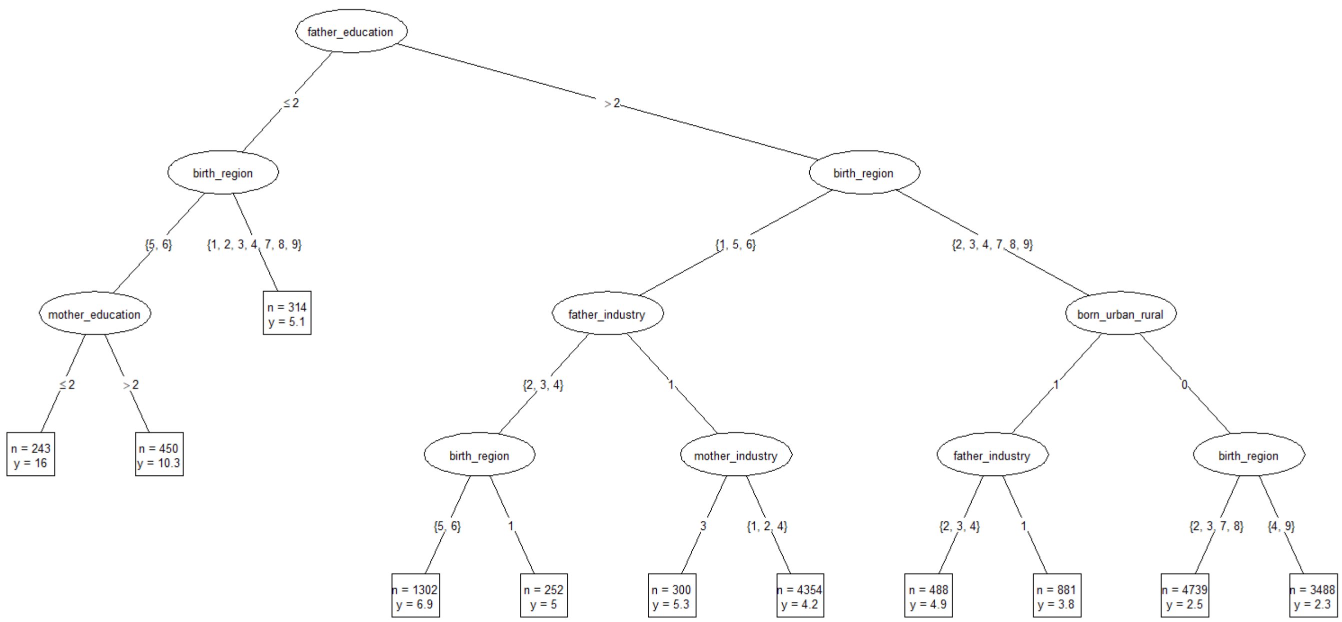
Religion	Birth Region
1 Christianity	1 Central
2 Islam	2 North
3 None	3 Other Country
4 Other	4 South
5 Traditional	

Mali



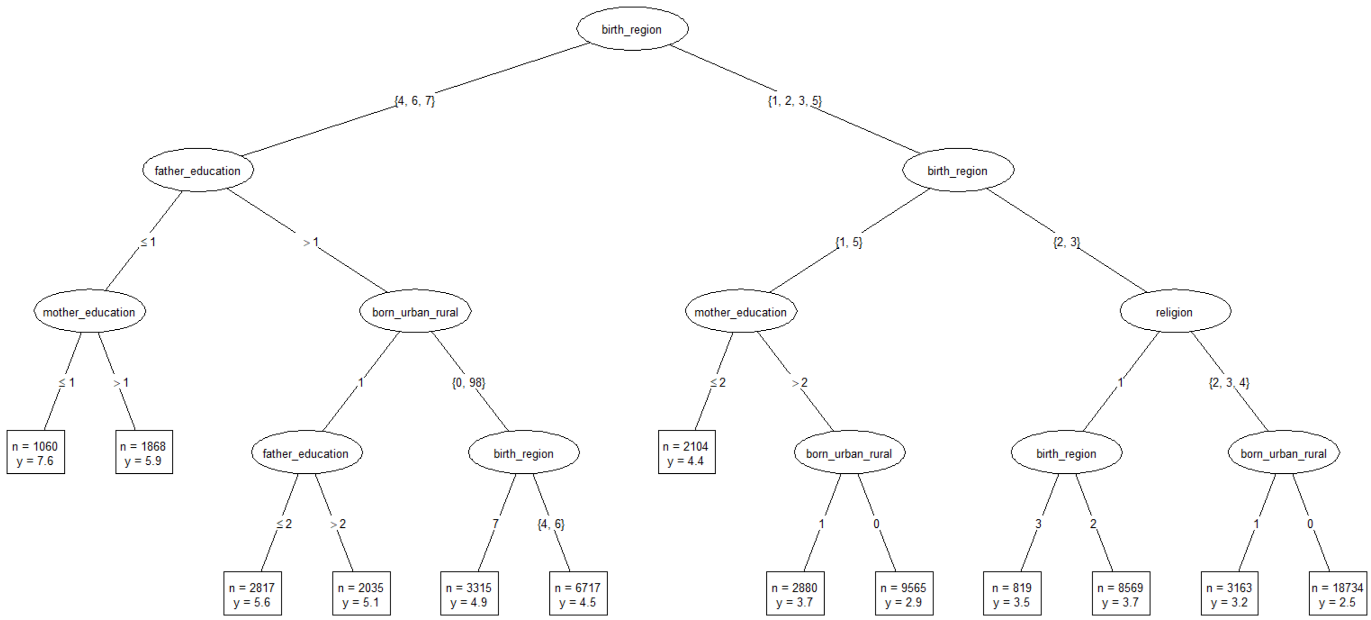
Religion	Birth Region
1 Animiste	1 Bamako
2 Autre Réligion	2 Gao
3 Chrétien	3 Kayes
4 Musulman	4 Kidal
5 Sans Réligion	5 Koulikoro
	6 Menaka
	7 Mopti
	8 Other Country
	9 Sikasso
	10 Ségou
	11 Taoudénit
	12 Tombouctou

Niger



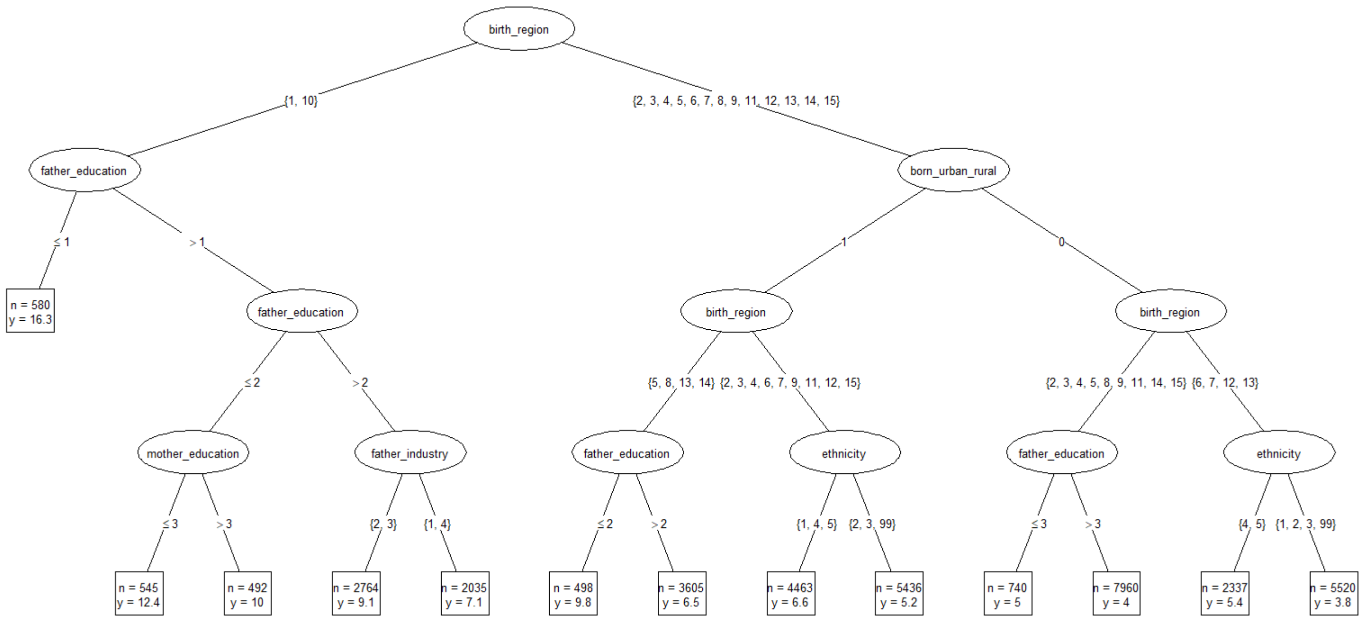
Religion	Birth Region
1 Animiste	1 Agadez
2 Autre Réligion	2 Diffa
3 Chrétien	3 Dosso
4 Musulman	4 Maradi
5 Sans Réligion	5 Niamey
	6 Other Country
	7 Tahoua
	8 Tillaberi
	9 Zinder

Nigeria



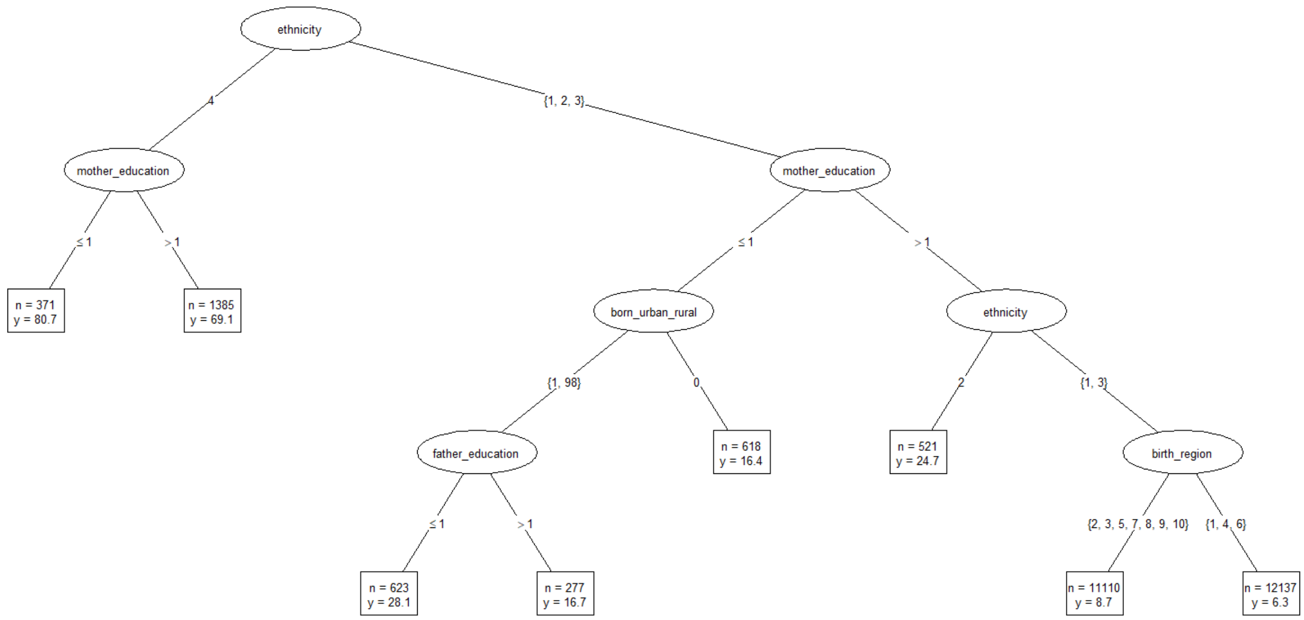
Religion	Birth Region
1 Christian	1 North Central
2 Islam	2 North East
3 Other	3 North West
4 Traditional	4 Other Country
	5 South East
	6 South South
	7 South West

Senegal



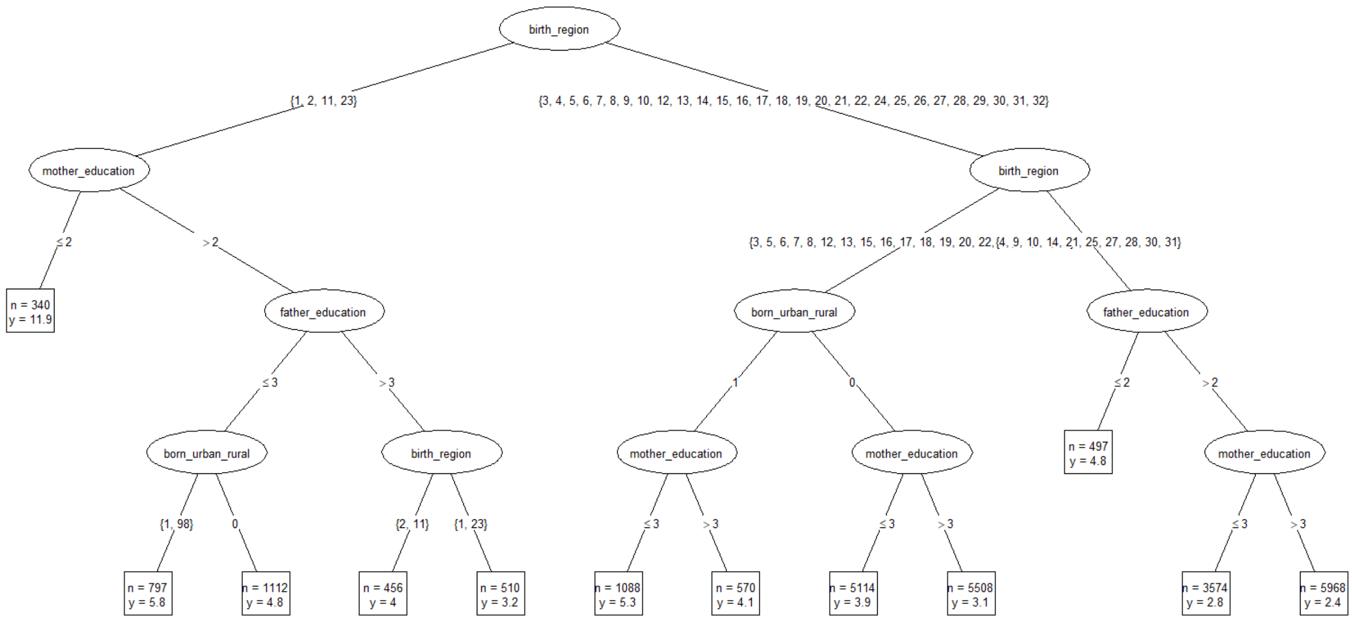
Ethnicity	Religion	Birth Region
1 Diola	1 Animiste	1 Dakar
2 Mandingue/Socé	2 Autre Réligion	2 Diourbel
3 Poular	3 Chrétien	3 Fatick
4 Sérère	4 Musulman	4 Kaffrine
5 Wolof/Lébou	5 Sans Réligion	5 Kaolack
99 Other		6 Kedougou
		7 Kolda
		8 Louga
		9 Matam
		10 Other Country
		11 Saint-Louis
		12 Sedhiou
		13 Tambacounda
		14 Thies
		15 Ziguinchor

South Africa



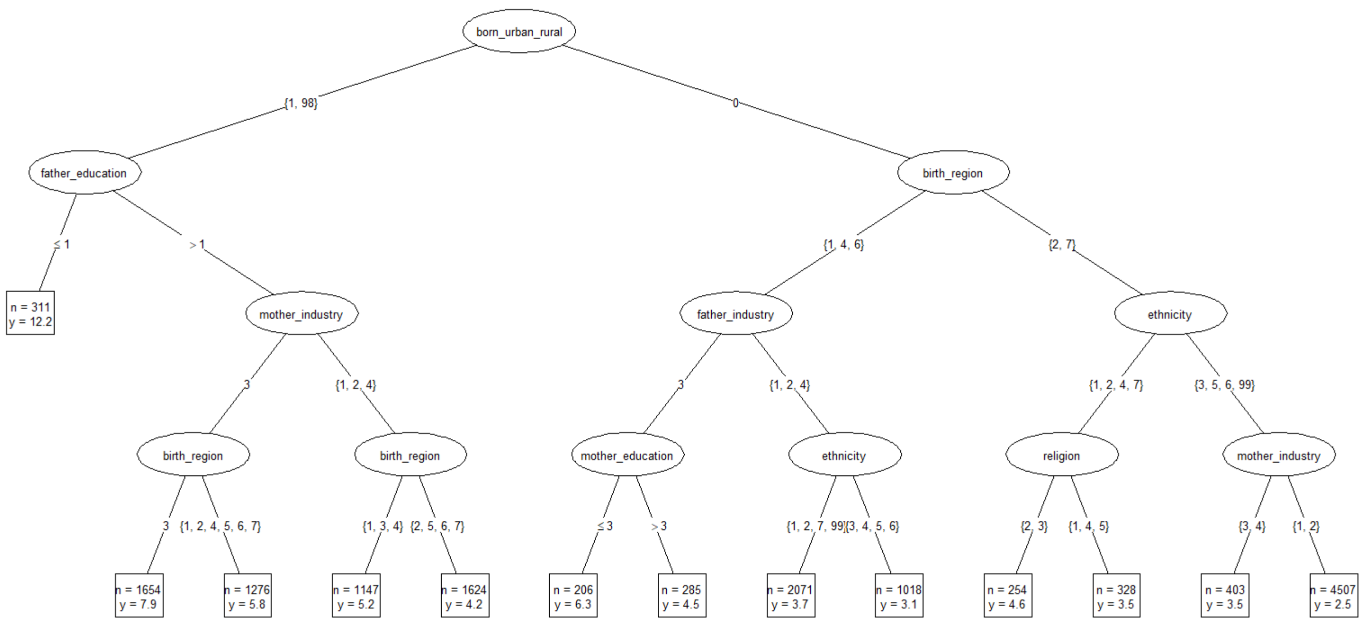
Race	Religion	Birth Region
1 African 2 Asian/Indian 3 Coloured 4 White	1 African traditional spiritual 2 Christian 3 Hindu 4 Jewish 5 Muslim 6 No religion 7 Other	1 Eastern Cape 2 Free State 3 Gauteng 4 KwaZulu-Natal 5 Limpopo 6 Mpumalanga 7 North West 8 Northern Cape 9 Other Country 10 Western Cape

Tanzania



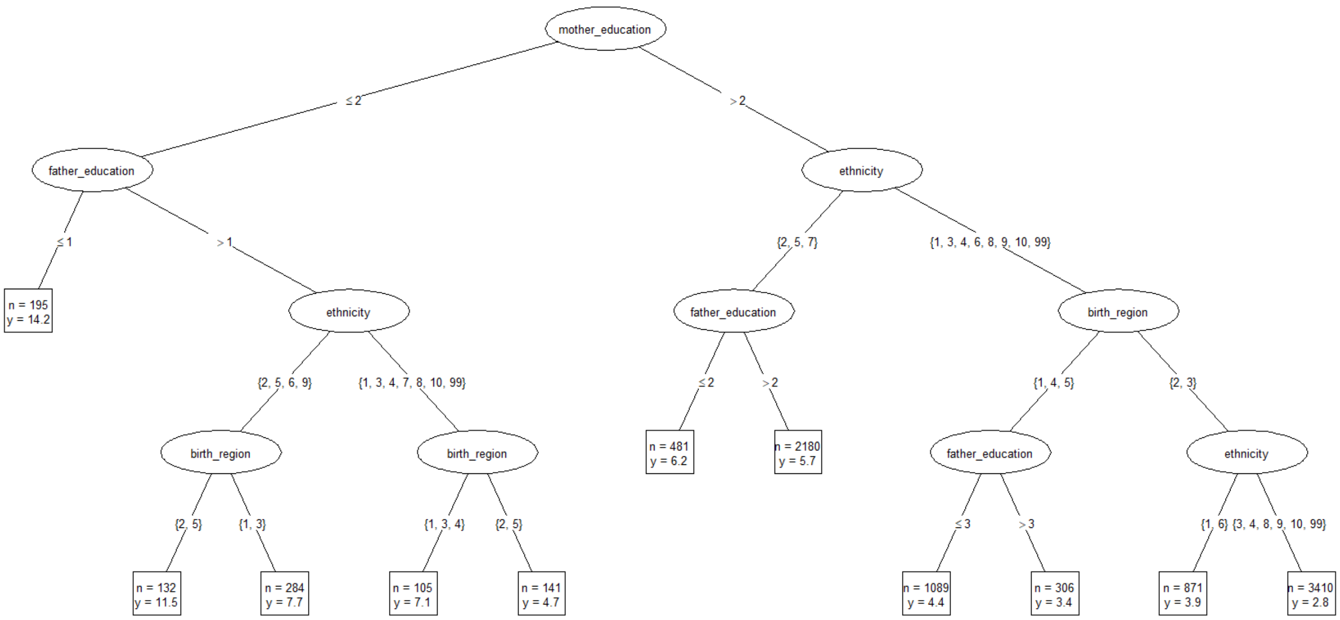
Birth Region	
1 Arusha	17 Mbeya
2 Dar Es Salaam	18 Mjini Magharibi
3 Dodoma	19 Morogoro
4 Geita	20 Mtwara
5 Iringa	21 Mwanza
6 Kagera	22 Njombe
7 Kaskazini Pemba	23 Other Country
8 Kaskazini Unguja	24 Pwani
9 Katavi	25 Rukwa
10 Kigoma	26 Ruvuma
11 Kilimanjaro	27 Shinyanga
12 Kusini Pemba	28 Simiyu
13 Kusini Unguja	29 Singida
14 Lindi	30 Songwe
15 Manyara	31 Tabora
16 Mara	32 Tanga

Togo



Ethnicity	Religion	Birth Region
1 Ana-Ifè 2 Ewé 3 Kabyè 4 Kotokoli 5 Lamba/ Nawdum 6 Moba 7 Ouatchi 99 Other	1 Animiste 2 Autre Réligion 3 Chrétien 4 Musulman 5 Sans Réligion	1 Centrale 2 Kara 3 Lomé commune 4 Maritime 5 Other Country 6 Plateaux 7 Savanes

Uganda



Ethnicity	Birth Region
1 Acholi	1 Central
2 Baganda	2 Eastern
3 Bagisu	3 Northern
4 Bakiga	4 Other Country
5 Banyakole	5 Western
6 Basoga	
7 Batoro	
8 Iteso	
9 Langi	
10 Lugbara	
99 Other	