SUPPLEMENTARY MATERIALS:
Heritage of Hunger: Ethnicity, Resource Allocation and Child Poverty

Eduard van der Merwe*

March 6, 2026

*Department of Economics, Rice University, Houston, Texas, United States. e.vdmerwe@rice.edu.

Appendix A: Brief Overview of LPW

Here, I provide a brief overview of the linear reframing proposed by Lechene et al. (2022) (LPW). I denote all household observables with the subscript h and W_h^t as the observed household budget share of the assignable good for member type t. Let z_h be a vector of covariates including both the sharing rule determinants z^{sr} and the preference shifters z^p. Lastly, let ϵ_h^t be an additive error term. Then, using the Piglog preference structure (3) from the main text, with the SAP assumption ($\beta = \beta_t$), the system of equations can be written as

$$W_h^t = a_h^t + b_h^t \ln y_h + \epsilon_h^t \;\; \forall \;\; t, \tag{A1}$$

where

$$a_h^t = \eta_t(z_h)\alpha^t(z_h) + \eta_t(z_h)\beta(z_h)\ln \eta_t(z_h) - \eta_t(z_h)\beta(z_h)\ln N_h^t \tag{A2}$$

and

$$b_h^t = \eta_t(z_h)\beta(z_h). \tag{A3}$$

By rearranging (A3), an equation for the resource shares is

$$\eta_t(z_h) = \frac{b_h^t}{\beta(z_h)}. \tag{A4}$$

Since estimating (A2) and (A3) directly is not feasible with a high-dimensional conditioning vector z_h, LPW approximate[1] these terms with the linear indices

$$a_h^t = a_0^t + a_{\ln N^t}^t \ln N_h^t + \boldsymbol{a}_z^{t'} \boldsymbol{z}_h \tag{A5}$$

and

$$b_h^t = b_0^t + \boldsymbol{b}_z^{t'} \boldsymbol{z}_h. \tag{A6}$$

Substituting (A5) and (A6) into (A1) yields

$$W_h^t = a_0^t + a_{\ln N^t}^t \ln N_h^t + \boldsymbol{a}_z^{t'} \boldsymbol{z}_h + b_0^t \ln y_h + \boldsymbol{b}_z^{t'} \boldsymbol{z}_h \ln y_h + \epsilon_h^t,$$

which is the linear individual assignable goods Engel curve for each household member type. Using seemingly unrelated regression (SUR), one can get the estimates $\hat{b}_h^t = \hat{b}_0^t + \hat{\boldsymbol{b}}_z^{t'} \boldsymbol{z}_h$. Using these estimates along with the assumption that individual resource shares summed across member types is 1, $\sum_{t=1}^{T} \hat{b}_h^t$ can be used as an estimate for $\beta(z_h)$.[2] Lastly,

[1] This approximation for b_h^t is exact if η_t is linear in z_h and β is independent of z_h - i.e., if β is a constant.

[2] To see this, notice that $\sum_t \eta_t(z_h) = \sum_t b_h^t / \beta(z_h) = 1$ which implies that $\beta(z_h) = \sum_t b_h^t$.

using these estimates and (A4), resource shares can be estimated by

$$\hat{\eta}_t(z_h) = \frac{\hat{b}^t_h}{\sum_{t=1}^{T}\hat{b}^t_h} = \frac{\hat{b}^t_0 + \hat{\boldsymbol{b}}^{t'}_z \boldsymbol{z}_h}{\sum_{t=1}^{T}\left(\hat{b}^t_0 + \hat{\boldsymbol{b}}^{t'}_z \boldsymbol{z}_h\right)}. \tag{A7}$$

LPW provides a note of caution on using (A7) since resource shares are vulnerable to a denominator with a lot of variation or one close to zero. For this reason, they propose the linear restriction

$$\sum_t b^t_{\tilde{z}} = 0, \tag{A8}$$

which implies that $\sum_{t=1}^{T}\hat{b}^t_h = \sum_t\left(\hat{b}^t_0 + \hat{b}^t_{N_m}N^m_h + \hat{b}^t_{N_f}N^f_h + \hat{b}^t_{N_c}N^c_h\right)$. Using this restriction, resource shares can be estimated using

$$\hat{\eta}_t(z_h) = \frac{\hat{b}^t_0 + \hat{\boldsymbol{b}}^{t'}_z \boldsymbol{z}_h}{\sum_t\left(\hat{b}^t_0 + \hat{b}^t_{N_m}N^m_h + \hat{b}^t_{N_f}N^f_h + \hat{b}^t_{N_c}N^c_h\right)}$$

separately for each household composition in each country.

To perform the linear identification tests, first notice from (A4) that resource shares cannot be identified if $\beta(z_h) = 0$. Now, the overall household assignable-goods Engel curve can be written as

$$W_h = a_h + b_h \ln y_h + \epsilon_h, \tag{A9}$$

which is the sum of individual member Engel curves.[3] Therefore, estimating (A9) with ordinary least squares (OLS) provides an estimate $\hat{b}_h$ of $\beta(z_h)$. When performing this identification test, I assume the linear restriction, i.e., (A8), holds. For every household in the sample, I test whether $\hat{b}_h$ is zero and report the total fraction of households for which it is not.

The second test is whether the overall assignable-goods Engel curve, evaluated at the mean of $\boldsymbol{z}_h$, has a non-zero slope. Letting $\bar{z}_h$ denote the vector of mean covariate values,

$$E[\hat{b}_h] = \hat{b}_0 + \hat{\boldsymbol{b}}'_z \bar{z}_h$$

can be used to perform a Z-test where the rejection of the null is a necessary condition for identification.

[3]By sum of individual member Engel curves, I mean: $W_h = \sum_t W^t_h$; $a_h = \sum_t a^t_h$; $b_h = \sum_t b^t_h$; and $\epsilon_h = \sum_t \epsilon^t_h$.

Appendix B: Data and Descriptive Statistics

This section outlines information about the data in the analysis. It contains information on sample construction for each country, the definition of assignable goods, descriptive statistics and the procedure for matching households to ancestral ethnic groups.

B.1: Data

Table B1 provides a detailed breakdown of the sample sizes and household compositions. While LPW utilized earlier rounds of the LSMS for several of these countries, they restricted their final analysis to Malawi due to identification constraints. By using more recent data that leverage improvements in survey methodology—such as

Table B1: Sample Descriptors

Country	Survey Year	Total Sample	Single-Adult Sample	Nuclear Sample	Estimation Sample	Household Compositions
Ethiopia	2019	6770	926	3891	5542	mfc, mf, fc, mc, m, f, c
Ghana	2016-2017	14009	3314	6451	9766	mfc, mf, fc, mc, m, f
Malawi	2019-2020	11434	1024	7315	10066	mfc, mf, fc, mc, m, f
Niger	2014	3617	153	2611	3191	mfc, mf, fc, mc, m, f
Nigeria	2018	5025	551	2909	3773	mfc, mf, fc, mc, m, f
Tanzania	2020-2022	4709	487	2678	3619	mfc, mf, fc, mc, m, f, c
Uganda	2019-2020	3098	318	1883	2537	mfc, mf, fc, mc, m, f
All	2014 - 2022	48662	6773	27738	38494	mfc, mf, fc, mc, m, f, c

Note: The estimation sample excludes households with missing/implausible interview dates, demographics (age, sex, or education), or consumption and clothing expenditure data. It also omits single-adult households and those where the number of men, women, or children exceeds six, six, and eight, respectively. Household compositions are defined by member presence: m, f, and c denote the presence of a man, woman, and child, respectively.

Computer-Assisted Personal Interviewing (CAPI) and the LSMS research agenda to enhance individual-level data (Carletto & Gourlay, 2019)—I extend the analysis to Ethiopia, Ghana and Nigeria. The estimation sample is constructed through a rigorous cleaning process to ensure the reliability of the structural estimates. I exclude households with missing or implausible data on age, sex or adult education, as well as those with missing or implausible values for aggregate consumption or clothing expenditure. To mitigate the influence of extreme outliers in large households, I restrict the sample to households with no more than six men, six women and eight children.

Table B2: Information on Assignable Good

	Member Type						Clothing Type				Children's Ages
	Men	Women	Children	Boys	Girls	Infants	Clothes	Footwear	Materials	Repairs	
Ethiopia	Y	Y	N	Y	Y	N	Y^a	Y^a	Y^a	N	< 18
Ghana	Y	Y	Y	S	S	S	Y	Y	Y	Y	< 14
Malawi	Y	Y	N	Y	Y	S^c	Y	Y	N	N	$< 15^b$
Niger	Y	Y	Y	N	N	N	Y	Y	N	Y	< 15
Nigeria	Y	Y	N	Y	Y	S^c	Y	Y	N	N	< 15
Tanzania	Y	Y	Y	N	N	N	Y	Y	N	N	< 16
Uganda	Y	Y	Y	N	N	N	Y	Y	N	N	$< 18^b$

Note: Y - Yes, N - No, and S - Some. The table depicts the information on assignable goods that are available. S indicates that some assignable goods can be distinguished between household member types. [a] Ethiopia combines these categories into one clothing expenditure that cannot be separated. [b] age of children is not clearly defined. [c] Footwear or cloth materials and repairs for infants are not given.

As shown in Table B1, a significant portion of the sample in each country is non-nuclear. The inclusion of multigenerational and complex household structures (indicated by compositions such as mfc, mf and fc) is essential for capturing the reality of social structures in Sub-Saharan Africa (SSA). The final estimation sample ranges from $2,537$ households in Uganda to $10,066$ in Malawi, totaling over $38,000$ households across the region. The primary assignable good across all countries is clothing and footwear. Table B2 summarizes data availability for these items. To maintain uniformity, I aggregate clothing and footwear expenditures into three categories: adult men, adult women and children. Furthermore, the definition of a "child" follows the specific country's survey design, ranging from under 14 in Ghana to under 18 in Ethiopia.

B.2: Descriptive Statistics

Table B3 presents summary statistics for the covariates in the final estimation sample. Demographic patterns show adult men are consistently older than adult women, with mean ages of approximately 37 and 34 years, respectively. There is a persistent gender education gap in the region, with adult men averaging 6.2 years of education compared to 4.6 years for adult women. The average age of children in the sample, ranging from 6.5 years in Ghana to 8.0 years in Ethiopia, reflects the varying thresholds used by the respective national surveys to define a child. Geographic distribution reveals that the

Table B3: Descriptive Statistics of Covariates

	Ethiopia	Ghana	Malawi	Nigeria	Total
Mean Age of Men	38.06 (12.11)	34.87 (12.02)	33.74 (12.19)	37.32 (11.47)	37.11 (11.88)
Mean Age of Women	34.06 (10.23)	34.38 (10.29)	33.17 (11.73)	33.91 (10.36)	33.95 (10.42)
Mean Age of Children	7.959 (4.086)	6.475 (3.287)	6.763 (3.423)	6.813 (3.256)	7.208 (3.651)
Minimum Age of Children	4.769 (4.401)	4.279 (3.810)	4.217 (3.713)	3.845 (3.879)	4.257 (4.086)
Mean Education of Men	4.182 (4.531)	7.336 (4.420)	6.894 (3.531)	7.594 (5.151)	6.241 (5.017)
Mean Education of Women	2.462 (3.781)	5.895 (4.275)	5.712 (3.481)	5.884 (5.087)	4.591 (4.747)
Area Type (0 - Rural, 1 - Urban)	0.247 (0.431)	0.498 (0.500)	0.158 (0.365)	0.275 (0.447)	0.275 (0.447)
Number of Men in a Household	1.264 (0.565)	1.557 (0.907)	1.408 (0.730)	1.715 (1.052)	1.509 (0.886)
Number of Women in a Household	1.218 (0.541)	1.627 (0.908)	1.380 (0.669)	1.753 (0.984)	1.513 (0.851)
Number of Children in a Household	2.923 (1.623)	2.392 (1.363)	2.411 (1.268)	3.298 (1.842)	3.011 (1.715)
Proportion of Boys in a Household	0.500 (0.346)	0.522 (0.379)	0.496 (0.371)	0.510 (0.336)	0.506 (0.346)

Note: This table displays the means and standard deviations (in parentheses) for all households. Values for mean age and mean education are the average age and education across member types within each household. Values for minimum child age are for the youngest child per household. Values for the number of men, women, and children are averages per household. Note that children's ages (as defined by the surveys) vary across samples.

samples are predominantly rural, with Malawi being the most rural (84%) and Ghana being the most urbanized (50% split). Finally, household compositions reveal near parity in the number of adult men and women per household, with an average of about 1.5 per household. These households also contain an average of 3 children, with the highest concentrations observed in Nigeria (3.3) and Ethiopia (2.9).

Table B4: Descriptive Statistics of Budget Information

	Ethiopia	Ghana	Malawi	Nigeria	Total
Share of Men's Clothes of Total Budget	0.0306 (0.0418)	0.0231 (0.0310)	0.00644 (0.0146)	0.0142 (0.0164)	0.0209 (0.0309)
Share of Women's Clothes of Total Budget	0.0264 (0.0351)	0.0166 (0.0186)	0.0125 (0.0177)	0.0147 (0.0132)	0.0190 (0.0248)
Share of Children's Clothes of Total Budget	0.0364 (0.0477)	0.0126 (0.0163)	0.0127 (0.0147)	0.0258 (0.0237)	0.0271 (0.0344)
Share of Men's Clothes of Clothes Budget	0.299 (0.285)	0.383 (0.353)	0.158 (0.284)	0.247 (0.256)	0.279 (0.287)
Share of Women's Clothes of Clothes Budget	0.286 (0.273)	0.349 (0.301)	0.369 (0.369)	0.272 (0.231)	0.292 (0.267)
Share of Children's Clothes of Clothes Budget	0.415 (0.341)	0.268 (0.265)	0.473 (0.390)	0.481 (0.298)	0.428 (0.323)

Note: This table displays the means and standard deviations (in parentheses). Clothes and budget shares are the mean values summed across household member types.

Table B4 displays summary statistics for assignable clothing shares. Across the sample, clothing accounts for less than 7% of the total household budget. This small magnitude does not hinder estimation, as identifying resource shares depends on the marginal responsiveness (the slope) of these goods to changes in the total budget, rather than their absolute level.[4] Within the clothing budget, children command the largest share on average (42.8%), driven heavily by households in Nigeria (48.1%), Malawi (47.3%) and Ethiopia (41.5%). Among adults, the overall share of the clothing budget allocated to women (29.2%) slightly exceeds that allocated to men (27.9%). Ghana, however, presents a notable exception to these regional trends. Men's clothing accounts for the largest share of the clothing budget (38.3%), surpassing both women (34.9%) and children (26.8%).

Table B5 presents child anthropometrics for those under five. Demographically, the sam-

[4] Resource shares identify household members' access to the total household budget, not just their consumption of the assignable good.

Table B5: Child Anthropometrics

	Ethiopia	Ghana	Malawi	Nigeria	Total
Sex (Male=0, Female=1)	0.491 (0.500)	0.484 (0.500)	0.500 (0.500)	0.501 (0.500)	0.494 (0.500)
Age (in completed months)	32.03 (14.54)	31.77 (16.56)	30.07 (16.73)	23.72 (16.16)	29.94 (16.47)
Weight (kg)	11.89 (3.599)	12.35 (4.612)	11.73 (3.299)	11.63 (4.053)	11.94 (3.936)
Height (cm)	84.64 (12.56)	86.98 (14.90)	83.99 (13.11)	83.82 (13.68)	85.02 (13.77)
Height-for-Age Z-score	-1.945 (2.265)	-0.927 (2.670)	-1.421 (1.569)	-1.573 (1.799)	-1.377 (2.155)
Weight-for-Age Z-score	-1.033 (1.681)	-0.600 (2.155)	-0.646 (1.179)	-0.910 (1.530)	-0.739 (1.687)
Weight-for-Height Z-score	0.167 (2.852)	0.0323 (3.513)	0.240 (1.421)	0.0367 (2.041)	0.130 (2.585)

Note: This table shows the means and standard deviations (in parentheses) for children under five years old. The sample includes children with plausible Z-scores ($[-5, 5]$).

ple is evenly split between boys and girls. The average age across the pooled sample is approximately 30 months, though the Nigerian cohort is noticeably younger on average (about 24 months). To evaluate child undernutrition, the table reports average Z-scores, which measure deviations from the World Health Organization (WHO) reference population. Specifically, height-for-age Z-scores (HAZ) capture stunting, weight-for-age Z-scores (WAZ) measure underweight status and weight-for-height Z-scores (WHZ) indicate wasting. The WHO designates a child as stunted, underweight or wasted if their respective Z-score falls below the -2 standard deviation threshold.

The average child in Ethiopia has a Height-for-Age Z-score (HAZ) of -1.945, perilously close to the formal WHO threshold for stunting, indicating severe and widespread developmental delays in that cohort. In fact, nearly half of all children in Ethiopia (47%) are stunted. Table B6 confirms that chronic undernutrition (stunting) affects 34% of the children in the sample, whereas acute undernutrition (wasting) affects only 9%. This corroborates findings that chronic developmental delays (stunting) are the primary nutritional challenge in the region (Brown et al., 2019).

Table B6: Child Malnutrition Categories

	Ethiopia	Ghana	Malawi	Nigeria	Total
		Stunting			
Healthy	947	2730	2319	1139	7135
$(HAZ > -1)$	(0.31)	(0.45)	(0.36)	(0.38)	(0.38)
Marginal	534	1212	1929	685	4360
$(-1 > HAZ > -2)$	(0.17)	(0.20)	(0.30)	(0.23)	(0.23)
Moderate	516	761	1316	527	3120
$(-2 > HAZ > -3)$	(0.17)	(0.12)	(0.21)	(0.18)	(0.17)
Severe	915	707	828	651	3101
$(HAZ < -3)$	(0.30)	(0.12)	(0.13)	(0.22)	(0.17)
		Underweight			
Healthy	1359	3205	3779	1478	9821
$(WAZ > -1)$	(0.44)	(0.51)	(0.60)	(0.49)	(0.53)
Marginal	756	1633	1676	774	4839
$(-1 > WAZ > -2)$	(0.25)	(0.26)	(0.27)	(0.26)	(0.26)
Moderate	464	657	558	403	2082
$(-2 > WAZ > -3)$	(0.15)	(0.10)	(0.09)	(0.13)	(0.11)
Severe	351	286	131	236	1004
$(WAZ < -3)$	(0.11)	(0.05)	(0.02)	(0.08)	(0.05)
Overweight	125	329	124	113	691
$(WAZ > 2)$	(0.04)	(0.05)	(0.02)	(0.04)	(0.04)
		Wasting			
Healthy	1593	3102	4853	2025	11573
$(WHZ > -1)$	(0.54)	(0.56)	(0.78)	(0.68)	(0.65)
Marginal	409	988	666	445	2508
$(-1 > WHZ > -2)$	(0.14)	(0.18)	(0.11)	(0.15)	(0.14)
Moderate	222	415	146	159	942
$(-2 > WHZ > -3)$	(0.07)	(0.07)	(0.02)	(0.05)	(0.05)
Severe	187	279	76	83	625
$(WHZ < -3)$	(0.06)	(0.05)	(0.01)	(0.03)	(0.04)
Overweight	378	414	426	192	1410
$(WHZ > 2)$	(0.13)	(0.07)	(0.07)	(0.06)	(0.08)

Note: This table displays the frequencies and fraction of the total sample (in parentheses) for children's malnutrition rates. The sample only includes children under five years whose Z-scores are plausible ($[-5, 5]$). There are $16,621$ children in the sample, with Ethiopia, Ghana, Malawi, and Nigeria contributing $2,703$, $4,952$, $6,084$, and $2,882$ children, respectively.

B.3: Matching to Ethnic Groups

Matching Strategy. Matching households to ethnic groups, specifically those in the Ethnographic Atlas (EA), is a well-established practice in the literature (Alesina et al., 2013; Aminjonov et al., 2024; Ashraf et al., 2020; Nunn & Wantchekon, 2011). Generally, it takes on one of the following strategies: (1) direct ethnic self-reports, (2) primary language spoken or (3) geographic location. Because the availability of these identifiers varies across the LSMS data, I adopt a strategy that maximizes matching accuracy while remaining as consistent as possible across the sample. For Ethiopia, Malawi and Nigeria, I intersect the cluster-level GPS coordinates with Ethnologue shapefiles (Giuliano & Nunn, 2018). I acknowledge that spatial matching can introduce measurement error, particularly given the LSMS privacy protocols.[5] However, because ethnic groups heavily cluster, this noise is unlikely to systematically bias the matches.

While some recent studies utilize language matching (Aminjonov et al., 2024), I match households spatially to the ancestral homelands defined in the EA. Thus, mitigating potential endogeneity concerns that recent migrants or minority households might adopt the dominant local language for economic or social assimilation (while still retaining their ancestral norms). For Ghana (no GPS coordinates), I utilize the self-reported ethnicity and a sequential matching strategy (Aminjonov et al., 2024; Nunn & Wantchekon, 2011). First, I match the member directly if a one-to-one match exists. Second, I use synonyms and known variations of ethnic group names, cross-referenced from the Afrobarometer, Demographic and Health Surveys (DHS) and the Joshua Project. Third, for members reporting broader ethnolinguistic groups, I use the classifications compiled by Giuliano and Nunn (2018). I successfully match approximately 95% of households in each country.

Validating Malawi Matching Strategy. To validate spatial matching over language matching, Table B7 cross-tabulates these methods for Malawi. While agreement is high at 74.6% (Kappa = 0.53, Cramér's V = 0.63), the divergence is systematic and illustrative of endogenous language assimilation. Specifically, many households in ancestral Makua territory now report Nyanja (Chichewa), the national common language in Malawi, as their primary language. Relying on language matching would erroneously misclassify these distinct ethnic minorities. GPS matching serves as a link to ancestral homelands

[5] The GPS coordinates of household locations are randomly offset by $0-2km$ for urban areas, $0-5km$ for rural areas and $0-10km$ in extreme cases (roughly 1% of enumeration areas). Consequently, the locations represent a set of coordinates, representative at the cluster level, that fall within known limits of accuracy.

and the deep-seated norms that may persist there despite language shifts.

Table B7: Malawi: Language vs. GPS Matching

Language Matched Ethnic Group	GPS Matched Ethnic Group							
	Laketonga	Makua	Ngonde	Nyanja	Safwa	Sena	Tumbuka	Yao
Laketonga	230	0	0	36	0	0	1	0
Makua	3	156	0	108	0	1	1	16
Ngonde	6	3	125	4	3	1	6	0
Nyanja	136	1,057	6	6,269	6	47	42	197
Safwa	7	0	36	3	242	0	7	0
Sena	1	36	0	134	0	252	0	0
Tumbuka	290	0	83	199	57	1	630	0
Yao	3	9	0	283	0	0	6	408

Note: This table displays the frequencies of households assigned to each Ethnic group in Malawi based on language matching (rows) and GPS matching (columns).

Stability and Persistence of Ancestral Traits. Although migration, technology advancement and societal transformation have certainly occurred, a large body of literature demonstrates that ancestral cultural traits exhibit significant persistence over centuries through inherited social norms and traditions (Alesina et al., 2013; Bahrami-Rad et al., 2021; Giuliano & Nunn, 2018, 2021). Following Bahrami-Rad et al. (2021), I verify the contemporary relevance of the EA data in this specific context by using ancestral traits (e.g., matrilocality, patrilocality, polygamy, animal husbandry and agriculture) to predict modern counterparts in the LSMS data (e.g., marriage structures, livestock ownership and the use of modern agricultural tools). While the statistical precision of these estimates is understandably lower than in broader cross-continental studies (given the small sample size and spatial noise), the consistent positive associations confirm that inherited social norms continue to govern intrahousehold dynamics today, even as households transition to modern agricultural tools or urban settings.

Spatial Heterogeneity of Ancestral Traits. To ensure the supervised machine learning procedure identifies specific ethnic traits rather than broad regional proxies, I examine variation in these traits among geographically adjacent ethnic groups. The data reveals significant cultural heterogeneity even among groups sharing the same regional characteristics, providing the necessary variation to disentangle culture from geography. For example, in North Ethiopia, the Tigrinya and Falasha exhibit relatively high ancestral agricultural dependence ($66 - 75\%$), whereas the Amhara exhibit medium dependence ($56 - 65\%$) and the Afar exhibit minimal agricultural reliance ($0 - 5\%$). Marriage customs also vary between the Tigrinya (dowry system), Afar and Macha (bride price) and the Koma (sister/relative exchange). Furthermore, settlement patterns range from the seminomadic patterns of the Afar to the compact permanent settlements of the Tigrinya and the separated hamlets of the Amhara.

Ethnic Matches in Urban Areas. The inclusion of urban households—representing a highly economically integrated segment of the population—is vital to avoid selection bias. Although urban residents may no longer engage in ancestral agricultural occupations, the social norms forged by those historical practices are frequently transmitted across generations. Additionally, African urban centers often exhibit "ethnic clustering", where migrants move into neighborhoods populated by co-ethnics, helping preserve cultural links. While I acknowledge that urban matching via GPS coordinates is inherently noisier than in rural settings, the majority of children in Ethiopia (60%), Ghana (68%), Malawi (85%) and Nigeria (74%) reside in rural areas. Nevertheless, to ensure that this urban noise does not bias the estimates, I control for area type in the sharing rule determinants.

Appendix C: Baseline Results

This section presents extra results from the Phase 1 estimation. Specifically, it details predicted poverty rates, the relationship between estimated resource shares and child undernutrition, concentration curves and the observed differences across ethnic groups.

C.1: Poverty Rates

Table C1 presents the predicted poverty rates for Ethiopia, Ghana, Malawi and Nigeria, comparing standard per-capita measures against both unadjusted and calorie-adjusted individual poverty rates. The unadjusted individual estimates reveal severe intrahousehold inequality, demonstrating that children and women experience systematically higher rates of poverty than men across all four countries. For instance, in Nigeria, the unadjusted child poverty rate reaches 76%, far exceeding the household per-capita rate of 42%. However, the intrahousehold gap narrows considerably when individual poverty lines are rescaled to account for age- and gender-specific caloric requirements (Brown et al., 2021). Because young children require fewer baseline calories than prime-age adults, their calorie-adjusted poverty rates fall (e.g., dropping from 56% to 36% in Ethiopia and from 72% to 46% in Malawi). Despite this adjustment, these rates remain significantly higher than global averages, highlighting the urgent need for poverty alleviation policies that effectively target the unique vulnerabilities of children.

Table C1: Predicted Poverty Rates

Country	Per-Capita Poverty Rate	Unadjusted Individual Poverty Rate			Calorie-Adjusted Individual Poverty Rate		
		Men	Women	Children	Men	Women	Children
Ethiopia	0.40	0.29	0.33	0.56	0.32	0.26	0.36
Ghana	0.37	0.24	0.33	0.56	0.27	0.27	0.34
Malawi	0.70	0.61	0.74	0.72	0.65	0.65	0.46
Nigeria	0.42	0.17	0.32	0.76	0.21	0.24	0.49

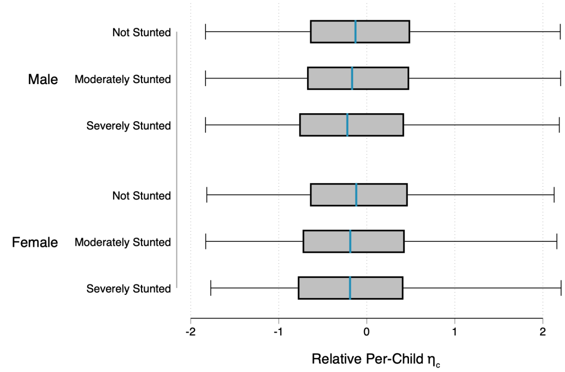
Note: This table displays estimated poverty headcount rates based on an extreme poverty line of $2.15/day (2017 US$ PPP). Unadjusted poverty rates are calculated as the individual's estimated resource share multiplied by total household expenditure. For calorie-adjusted rates, I utilize a calorie-based poverty line following Brown et al. (2021) assuming $2.15/day is the baseline threshold for adults aged 15-45. Individual poverty lines are rescaled according to the 2015–2020 Dietary Guidelines for Americans based on age and gender. For example, a male aged 16–25 has a recommended intake of 2,800 calories, resulting in a scaled poverty line of $2.51/day.

C.2: Resource Shares and Undernutrition

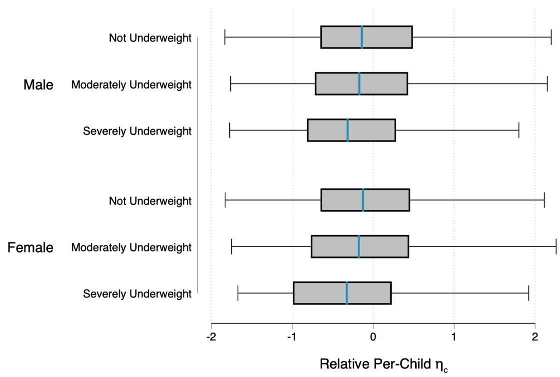
As noted by Brown et al. (2019), stunting, underweight and wasting represent distinct nutritional challenges with differing causes and consequences. Stunting (HAZ) and underweight (WAZ) serve as indicators of prolonged, persistent undernutrition where recovery is complex and often carries adverse long-term developmental consequences. In contrast, wasting (WHZ) reflects acute, short-term undernutrition, typically driven by seasonal food deprivation or recent illness. To explore the correlations between intrahousehold resource shares and undernutrition, Figure C1 displays boxplots categorizing children by nutritional severity. Across all three anthropometric indicators and for both boys and girls, children classified as "severely undernourished" exhibit lower median relative resource shares compared to their "moderately undernourished" or "not undernourished" peers. This provides descriptive evidence that structurally estimated resource allocations map onto observable physical outcomes.

To examine this relationship across the broader expenditure distribution, the logit es-

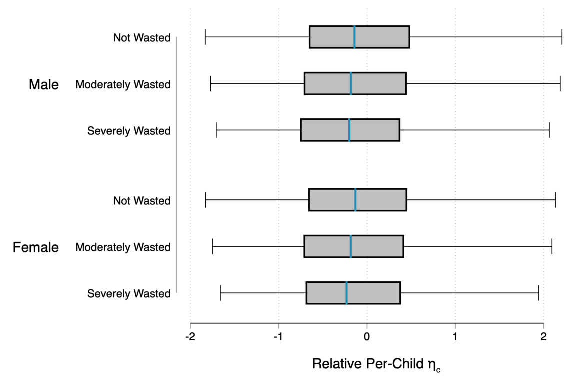
Figure C1: Children Undernutrition and Resource Shares



(a) Stunting



(b) Underweight

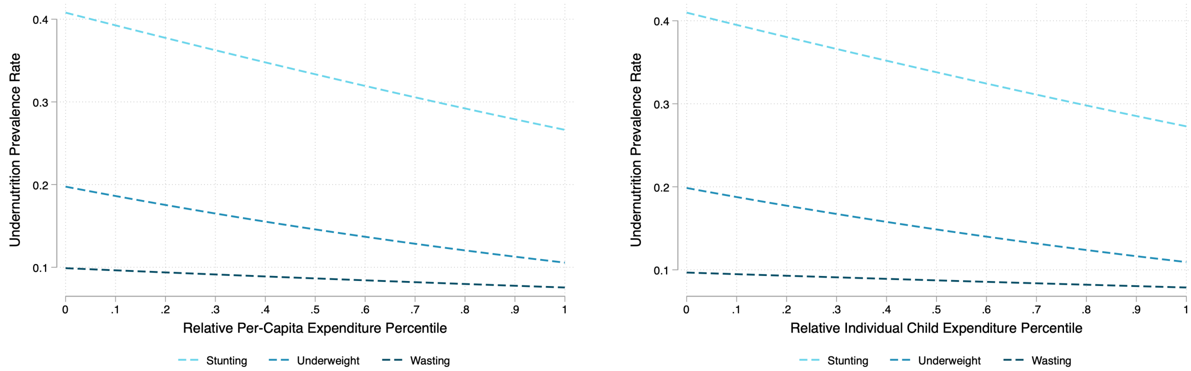


(c) Wasting

Note: These figures display the distribution of country-standardized per-child resource shares (Z-scores of η_c) across three anthropometric indicators: (a) stunting, (b) underweight and (c) wasting. Undernutrition categories are defined by child-specific Z-scores: Severe ($Z < -3$), Moderate ($-3 \leq Z < -2$) and Not Undernourished (≥ -2). The sample includes children aged 0–5 years, excluding observations with biologically implausible Z-scores (outside the $[-5, 5]$ range).

timation curves (Figure C2) plot the marginal predicted probability of undernutrition against relative within-country expenditure percentiles. Both the per-capita and individual expenditure graphs exhibit a clear downward trend, consistent with the expected income effect. For stunting, the predicted prevalence falls from approximately 41% at the lowest expenditure percentile to 27% at the highest. However, these rates remain strikingly high even at the top of the distribution (27% for stunting and 11% for underweight). This dynamic is further visualized by the concentration curves in Figure C3. While the curves lie above the 45-degree line of equality—confirming that undernutrition is concentrated among poorer households—the deviation from the line is relatively marginal.

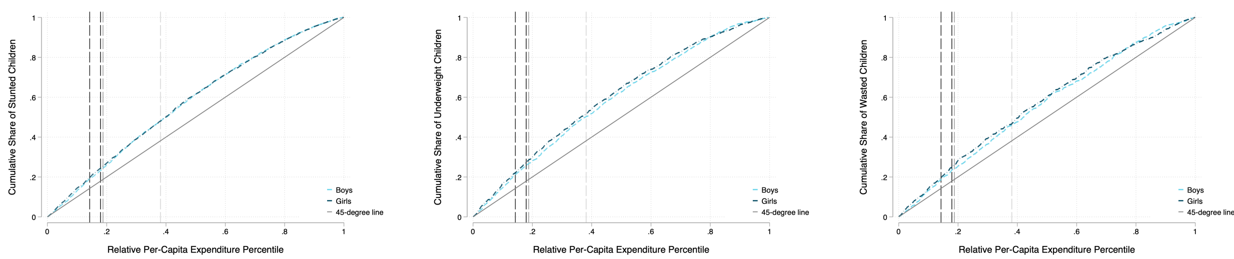
Figure C2: Undernutrition Prevalence by Expenditure Percentile



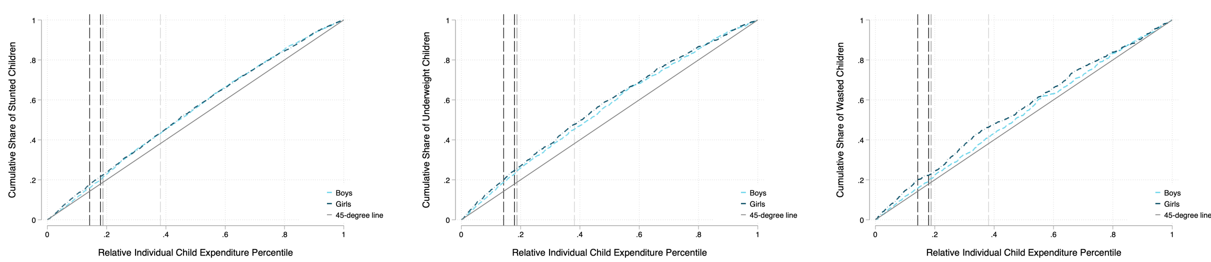
(a) Per-Capita Expenditure (b) Individual Child Expenditure

Note: These figures illustrate the marginal predicted probability of child undernutrition across country-specific expenditure percentiles for (a) household per-capita expenditure and (b) individual child expenditure. Probabilities are estimated via logit regressions controlling for child demographics (age, sex), household composition (size, average member age), gender education gaps, rural residency and country fixed effects. Undernutrition is defined as a Z-score below -2 for the respective indicator. The sample includes children aged 0–5 years, excluding observations with biologically implausible Z-scores (outside the $[-5, 5]$ range).

Figure C3: Undernutrition Concentration Curves



(a) Stunting (Per-Capita) (b) Underweight (Per-Capita) (c) Wasting (Per-Capita)



(d) Stunting (Individual) (e) Underweight (Individual) (f) Wasting (Individual)

Note: These figures depict undernutrition concentration curves disaggregated by gender (light blue dashed for boys; dark blue dashed for girls). Panels (a)-(c) plot rates against relative within-country household per-capita expenditure percentiles, while panels (d)-(f) utilize relative within-country individual child expenditure. Individual child expenditure is calculated as total annual household expenditure (2017 US$ PPP) multiplied by the estimated per-child resource share (η_c). Vertical dashed lines represent percentiles corresponding to the World Bank extreme poverty line ($1.29/day), mapped for Nigeria (dark grey), Ghana/Ethiopia (medium grey) and Malawi (light grey). Children are classified as stunted (HAZ), underweight (WAZ) or wasted (WHZ) if their Z-scores fall below -2. The sample includes children aged 0–5 years, excluding observations with biologically implausible Z-scores (outside the $[-5, 5]$ range).

These results suggest that the burden of undernutrition is not exclusively confined to the poorest households. There are two primary explanations for this persistence. First, the sample consists of low-income SSA countries. Therefore, the top expenditure percentile locally does not equate to a high standard of living globally. Second, children in the sample reside predominantly in rural communities and face poor sanitary infrastructure and disease environments that can impede nutrition regardless of food expenditure. Ultimately, this evidence highlights that relying solely on aggregate household wealth masks the reality that a child's individual share of intrahousehold resources correlates with their nutritional well-being.

C.3: Differences Across Ethnic Groups

A premise of this study is that child nutrition and intrahousehold resource shares exhibit significant variation across ethnic groups. If these outcomes were uniform within a country, the role of ethnic-specific cultural traits would be negligible. To test this, I perform one-way ANOVA tests to examine mean differences in estimated resource shares and nutrition outcomes across ethnic groups. Table C2 reports the ANOVA results for per-individual men, women and children resource shares and the aggregate per-household share for children. Across all four countries, the null hypothesis of equal means across

Table C2: ANOVA Tests for Intrahousehold Resource Shares

Country	Number of Ethnic Groups	Per-Individual η_m	Per-Individual η_w	Per-Individual η_c	Per-Household η_c
Ethiopia	28	5.85 (0.000)	8.54 (0.000)	22.10 (0.000)	5.36 (0.000)
Ghana	24	38.47 (0.000)	38.13 (0.000)	71.23 (0.000)	7.61 (0.000)
Malawi	8	3.98 (0.000)	16.11 (0.000)	11.00 (0.000)	2.40 (0.019)
Nigeria	55	7.47 (0.000)	12.49 (0.000)	26.11 (0.000)	2.23 (0.000)

Note: This table reports F-statistics and p-values (in parentheses) from one-way ANOVA tests examining mean differences in estimated resource shares across ethnic groups within each country. Columns (1)–(3) test the equality of mean individual shares (η_t) for men, women and children, respectively. Column (4) tests whether the total household resource share allocated to all children differs across ethnic groups.

ethnic groups is rejected at the 1% level for every member type. Similarly, Table C3 shows that mean Z-scores (HAZ, WAZ, WHZ) and the prevalence rates of stunting, underweight and wasting differ significantly across ethnic groups. These results provide a strong empirical foundation for the study, confirming that resource allocation patterns and child nutrition outcomes are fundamentally heterogeneous by ethnicity.

Table C3: ANOVA Tests for Child Undernutrition

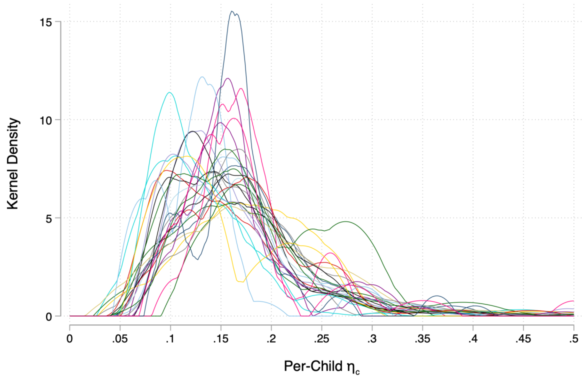
Country	Number of Ethnic Groups	HAZ	WAZ	WHZ	Stunted	Underweight	Wasted
Ethiopia	27	3.77 (0.000)	6.52 (0.000)	4.96 (0.000)	3.90 (0.000)	4.11 (0.000)	2.98 (0.000)
Ghana	24	7.40 (0.000)	5.67 (0.000)	4.13 (0.000)	8.64 (0.000)	4.06 (0.000)	4.43 (0.000)
Malawi	8	4.20 (0.000)	7.87 (0.000)	8.51 (0.000)	2.68 (0.009)	3.83 (0.000)	3.51 (0.001)
Nigeria	54	6.97 (0.000)	4.13 (0.000)	2.61 (0.000)	7.26 (0.000)	4.61 (0.000)	1.49 (0.012)

Note: This table reports F-statistics and p-values (in parentheses) from one-way ANOVA tests for mean differences in child anthropometrics across ethnic groups within each country. Columns (1)-(3) analyze continuous Z-scores (HAZ, WAZ, WHZ), while columns (4)-(6) test differences in undernutrition prevalence (binary indicators for stunting, underweight and wasting, defined as $Z < -2$). The sample includes children aged 0–5 years, excluding observations with biologically implausible Z-scores (outside the $[-5, 5]$ range).

Beyond differences in means, I inspect Kernel density estimates to assess variation in the full distributions of resource shares and child HAZ (see Figures C4 and C5). While the density shapes for resource shares are somewhat consistent across some groups, there are clear horizontal shifts in the distributions. These shifts indicate that certain ethnic groups systematically allocate a higher proportion of resources to children regardless of household income. Furthermore, the kernel densities exhibit first-order stochastic dominance between specific ethnic groups within the same country. This implies that at any given level of resource allocation, children in certain ethnic groups are statistically more likely to receive a higher share than their peers in other groups.

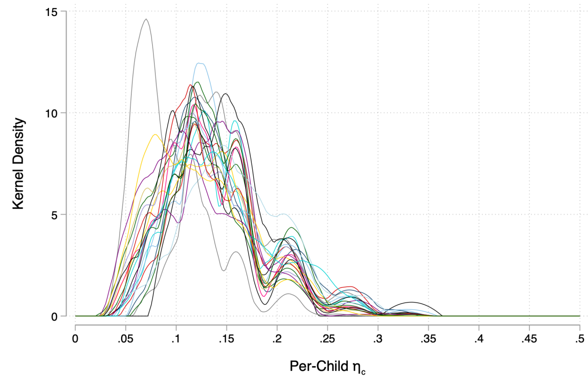
The variation is even more pronounced for child HAZ, with distributions showing substantial differences in both central tendency and dispersion. A critical observation in the child HAZ densities is the variation in left-tail density. While the modal child in two different ethnic groups might have similar Z-scores, the thickness of the left tail—

Figure C4: Kernel Densities of Predicted Per-Child Resource Shares by Ethnic Group



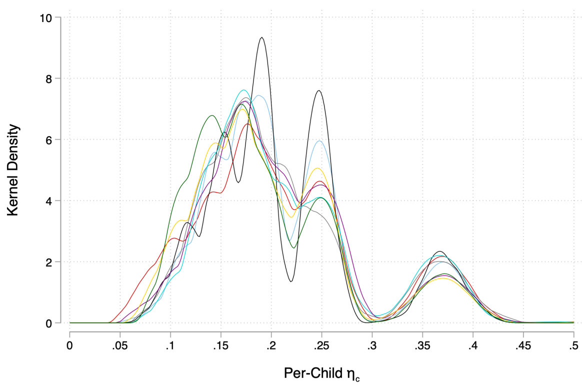
(a) Ethiopia

Note: This figure displays kernel density estimates of estimated per-child resource shares (η_c) across 27 ethnic groups in Ethiopia. To ensure estimation stability, groups with fewer than 20 child observations are excluded.



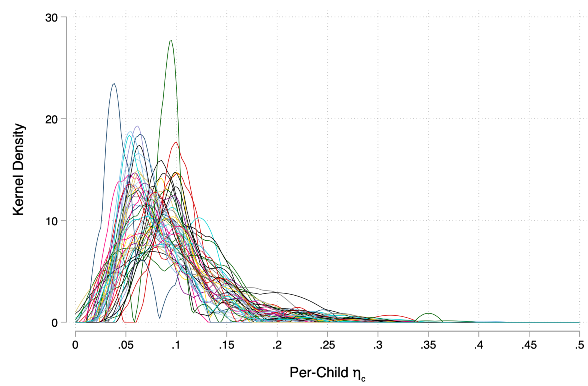
(b) Ghana

Note: This figure displays kernel density estimates of estimated per-child resource shares (η_c) across 23 ethnic groups in Ghana. To ensure estimation stability, groups with fewer than 20 child observations are excluded.



(c) Malawi

Note: This figure displays kernel density estimates of estimated per-child resource shares (η_c) across 8 ethnic groups in Malawi. To ensure estimation stability, groups with fewer than 20 child observations are excluded.

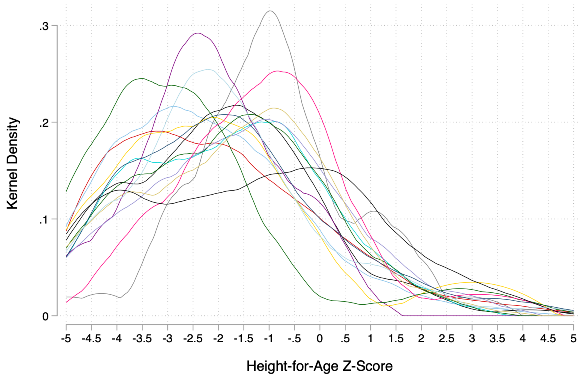


(d) Nigeria

Note: This figure displays kernel density estimates of estimated per-child resource shares (η_c) across 50 ethnic groups in Nigeria. To ensure estimation stability, groups with fewer than 20 child observations are excluded.

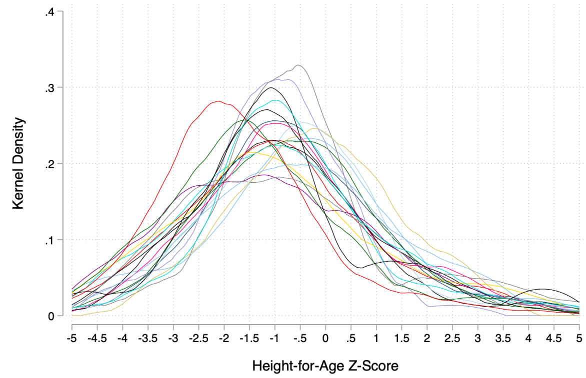
representing the proportion of children falling into severe undernutrition—varies significantly. In sum, the kernel densities confirm that ethnicity does not merely shift the average outcome but it reshapes the entire distribution of child welfare. Across the four-country sample, the observed differences in skewness, kurtosis and horizontal displacement provide descriptive evidence that underlying ethnic traits may be a primary driver of the structural variation in child welfare.

Figure C5: Kernel Densities of Height-for-Age Z-Scores (HAZ) by Ethnic Group



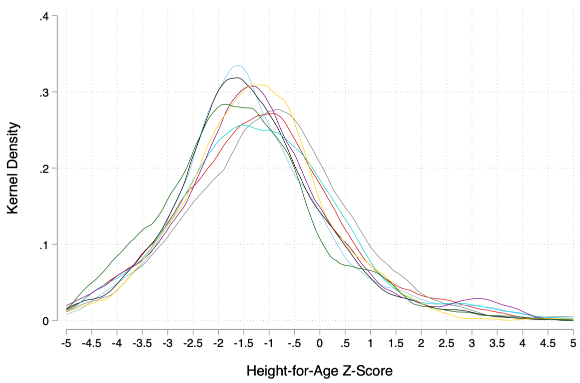
(a) Ethiopia

Note: This figure displays kernel density estimates of HAZ for 15 ethnic groups in Ethiopia. The sample is restricted to biologically plausible Z-scores ($[-5, 5]$). Groups with fewer than 20 observations are excluded to ensure reliable density estimates.



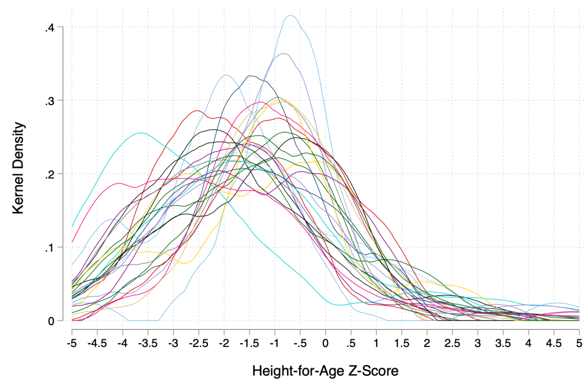
(b) Ghana

Note: This figure displays kernel density estimates of HAZ for 20 ethnic groups in Ghana. The sample is restricted to biologically plausible Z-scores ($[-5, 5]$). Groups with fewer than 20 observations are excluded to ensure reliable density estimates.



(c) Malawi

Note: This figure displays kernel density estimates of HAZ for 8 ethnic groups in Malawi. The sample is restricted to biologically plausible Z-scores ($[-5, 5]$). Groups with fewer than 20 observations are excluded to ensure reliable density estimates.



(d) Nigeria

Note: This figure displays kernel density estimates of HAZ for 27 ethnic groups in Nigeria. The sample is restricted to biologically plausible Z-scores ($[-5, 5]$). Groups with fewer than 20 observations are excluded to ensure reliable density estimates.

Appendix D: Feature Selection

This section details the feature selection results used to identify the relevant ethnic traits. First, I describe the ancestral ethnic characteristics selected from the Ethnographic Atlas (EA). I then present the results across 20 distinct feature selection specifications covering various outcome variables and country samples. Finally, I provide robustness checks using Recursive Feature Elimination with Cross-Validation (RFECV) and a Group Lasso framework with K-Nearest Neighbors (KNN) imputation.

D.1: Variables from the Ethnographic Atlas

Table D1 displays the 56 ethnic practices from the EA included in the feature selection process. These categories were selected based on three criteria. First, I excluded categories with no variation across the ethnic groups in the sample. Second, I removed redundant or secondary variables, such as alternative forms of marriage. Third, I excluded variables with high rates of missing data or those lacking a clear theoretical link to child welfare (such as linguistic affiliation). While most EA practices are nominal, some are ordinal ($v1$ - $v5$, $v31$, $v32$ and $v33$). For nominal traits, I generate indicator (dummy) variables for each specific characteristic. I rescale all ordinal variables to the $[0, 1]$ interval to ensure comparability during the feature selection process. For a comprehensive description of these categories and the underlying coding scheme, see Murdock et al. (1999).

Table D1: Ethnic Characteristics

Identifier	Ethnic Characteristic Category	Identifier	Ethnic Characteristic Category
v1	gathering	v40	predominant type of animal husbandry
v2	hunting	v42	subsistence economy
v3	fishing	v43	descent: major type
v4	animal husbandry	v44	sex differences: metal working
v5	agriculture	v45	sex difference: weaving
v6	mode of marriage	v46	sex differences: leather working
v8	domestic organization	v47	sex differences: pottery making
v9	marial composition: monogamy and polygamy	v48	sex differences: boat building
v11	transfer of residence at marriage: after first years	v49	sex difference: house building
v12	marital residence with kin: after first years	v50	sex differences: gathering
v15	community marriage organization	v51	sex differences: hunting
v17	largest patrilineal kin group	v52	sex differences: fishing
v19	largest matrilineal exogamous group	v53	sex differences: animal husbandry
v21	largest martrilineal kin group	v54	sex differences: agriculture
v23	cousin marriages (allowed)	v66	class stratification
v25	preferred rather than just permitted cousin marriages	v70	type of slavery
v27	kin terms for cousins	v71	former presence of slavery
v28	intensity of agriculture	v72	succession to the office of local headman
v29	major crop type	v73	hereditary succession type to the local headman office
v30	settlement patterns	v74	inheritance rule for real property (land)
v31	mean size of local communities	v75	inheritance distribution for real property (land)
v32	jurisdictional hierarchy of local community	v76	inheritance rule for movable property
v33	jurisdictional hierarchy beyond local community	v77	inheritance distribution for movable property
v34	high gods	v78	norms of premarital sexual behavior of girls
v35	games	v90	political integration
v36	post	v94	political succession for the local community
v37	male genital mutilations	v112	trance states
v38	segregation of adolescent boys	v113	societal rigidity

Note: This table depicts the 56 variable categories selected from the Ethnographic Atlas for the feature selection process. For the exhaustive list of categories and the specific characteristics within each, please refer to Murdock et al. (1999).

D.2: Random Forest Regressor

The results of the feature selection process using the Random Forest regressor are detailed in Table D2. I denote each ethnic characteristic as $v\kappa\_\rho$, where κ is the variable identifier according and ρ represents the specific subcategory. For example, $v6\_1$ indicates that bride price or bride wealth is the prevailing mode of marriage transaction. The Random Forest algorithm was estimated using $5,000$ trees. To identify the most relevant traits across 20 distinct specifications, Table D2 highlights the top-performing features using two criteria. Panel A displays characteristics that appear most frequently (at least four times) in the top 10 features of each specification. Panel B lists characteristics prioritized by their cumulative predictive power, listing traits with a cumulative feature importance score greater than 0.28.

To provide context on ethnic trait prevalence, the table reports the number of households with non-missing data, along with the subset that actually possesses the trait. For instance, out of $24,671$ households, $22,498$ have valid records for male genital mutilation. Within that sample, $2,520$ households belong to ethnic groups where the procedure occurs between ages 6 and 10 ($v37\_5$). This specific characteristic appears in the top 10 most important features three times, achieving a total importance score of 0.40. Similarly, one of the highest-performing traits, bride price ($v6\_1$), is practiced by $14,518$ households and appears in the top 10 a maximum of seven times.

D.3: Feature Selection Robustness

To ensure stability and robustness, I perform Recursive Feature Elimination with Cross-Validation (RFECV) (Table D3). This procedure identifies optimal characteristics that are selected at least nine times across the 20 specifications. The identified characteristics broadly align with jurisdictional hierarchy and community size ($v33$, $v31$), marriage and domestic organization ($v8\_4$, $v15\_2$, $v6\_1$, $v23\_3$), subsistence agriculture ($v5$), sex differences in labor ($v45\_1$, $v47\_5$, $v54\_2$) and social stratification or slavery ($v70\_4$, $v66\_3$).

As an additional check, Table D4 displays results using a Group Lasso regressor. Because the Group Lasso model cannot process missing values, I impute the missing EA data using a KNN algorithm. The results are broadly consistent with the Random Forest regressor, emphasizing the importance of major crop types ($v29\_3$ and $v29\_2$) and marriage transactions (such as bride price, $v6\_1$). Characteristics with high rates of missing

Table D2: Random Forest Regressor Feature Selection

Label	Ethnic Characteristic	Sample Size	Frequency	Sum of Feature Importance Scores
Panel A: By Frequency				
v6_1	mode of marriage: bride price or wealth, to bride's family	24671 (14518)	7	0.67
v4	subsistence economy: animal husbandry	24671 (.)	6	0.37
v32	jurisdictional hierarchy of local community	23530 (.)	5	0.28
v33	jurisdictional hierarchy beyond local community	23499 (.)	5	0.22
v54_2	sex differences: agriculture: males appreciably more	19677 (5690)	4	0.59
v54_5	sex differences: agriculture: female appreciably more	19677 (7445)	4	0.47
v30_4	settlement patterns: neighborhoods of dispersed family homesteads	23576 (8967)	4	0.18
v25_3	preferred rather than just permitted cousin marriages: duolateral, patrilineal preference	22785 (561)	4	0.17
v23_3	cousin marriages (allowed): Nonlateral evidence only for first cousins	22785 (5405)	4	0.16
Panel B: By Sum of Scores				
v66_3	class stratification: dual	21859 (5594)	3	1.00
v6_1	mode of marriage: bride price or wealth, to bride's family	24671 (14518)	7	0.67
v34_4	high gods: supportive of human morality	12770 (3078)	3	0.59
v54_2	sex differences: agriculture: males appreciably more	19677 (5690)	4	0.59
v54_5	sex differences: agriculture: female appreciably more	19677 (7445)	4	0.47
v47_5	sex differences: pottery making: activity present	18911 (2470)	2	0.44
v29_3	major crop type: cereal grains	23601 (16630)	3	0.43
v34_1	high gods: absent or not reported	12770 (972)	2	0.41
v37_5	male genital mutilations: 6 to 10 years	22498 (2520)	3	0.40
v4	subsistence economy: animal husbandry	24671 (.)	6	0.37
v30_2	settlement patterns: seminomadic	23576 (425)	3	0.33
v53_1	sex differences: animal husbandry: males only	7697 (2183)	3	0.29
v32	jurisdictional hierarchy of local community	23530 (.)	5	0.28

Note: This table presents the ethnic characteristics identified via a Random Forest feature selection algorithm. Each characteristic is denoted as $v\kappa\_\rho$, where κ is the primary ethnic variable identifier and ρ indicates the specific subcategory. Feature selection was executed across 20 distinct specifications, encompassing four child welfare outcomes (poverty, stunting, underweight and wasting) and five samples (the pooled multi-country sample, plus Ethiopia, Ghana, Malawi and Nigeria). Panel A highlights traits ranking in the top 10 features in at least four specifications. Panel B displays traits with a cumulative feature importance score exceeding 0.28 across all specifications. The "Sample Size" column reports households with non-missing data for the parent variable, with the count of households possessing the specific trait in parentheses. The total sample $(24,671)$ comprises estimation-sample households matched to ethnic groups containing at least one child.

Table D3: Recursive Feature Elimination with Cross-Validation

Identifier	Ethnic Characteristic	Count
v33	jurisdictional hierarchy beyond local community:	11
v31	mean size of local communities:	10
v8_4	domestic organization: polygynous: usual co-wives	10
v15_2	community marriage organization: segmented communities without local exogamy	10
v45_1	sex difference: weaving: males only	10
v70_4	type of slavery: hereditary and socially significant	9
v5	agriculture:	9
v47_5	sex differences: pottery making: activity present	9
v54_2	sex differences: agriculture: males appreciably more	9
v6_1	mode of marriage: bride price or wealth, to bride's family	9
v23_3	cousin marriages (allowed): Nonlateral evidence only for first cousins	9
v66_3	class stratification: dual	9

Note: This table displays ethnic characteristics identified as optimal via Recursive Feature Elimination with Cross-Validation (RFECV) in at least nine of the 20 feature selection specifications. Feature selection was executed across 20 distinct specifications, encompassing four child welfare outcomes (poverty, stunting, underweight and wasting) and five samples (the pooled multi-country sample, plus Ethiopia, Ghana, Malawi and Nigeria). "Count" represents the number of times a characteristic was selected as optimal across the 20 specifications.

values primarily drive differences between the models. For instance, the Group Lasso model places higher importance on the historical presence of slavery (*v*71_1, *v*71_2), inheritance distributions (*v*77_1, *v*77_2) and the segregation of adolescent boys (*v*38_1, *v*38_2). These divergences likely stem from the KNN imputation process or the structural differences in how Group Lasso penalizes categorical blocks compared to tree-based estimators. Nevertheless, the broad convergence of these two distinct methods on key variables reinforces the robustness of the primary feature selection.

Table D4: Group Lasso Regressor Feature Selection

Label	Ethnic Characteristic	Sample Size	Frequency	Sum of Absolute Coefficients
Panel A: By Frequency				
v29_3	major crop type: cereal grains	23601 (16630)	6	0.18
v29_2	major crop type: roots or tubers	23601 (6610)	5	0.17
v71_1	former presence of slavery: absent or exists currently and in past	23571 (11318)	4	0.07
v71_2	former presence of slavery: formerly present but not currently existing	23571 (12253)	4	0.07
v77_2	inheritance distribution for movable property: primogeniture	22656 (17558)	4	0.05
v38_1	segregation of adolescent boys: absent	9084 (3192)	4	0.04
v38_2	segregation of adolescent boys: partial	9084 (2024)	3	0.05
v34_4	high gods: supportive of human morality	12770 (3078)	3	0.05
v77_1	inheritance distribution for movable property: equal or relatively equal	22656 (4610)	3	0.04
v52_2	sex differences: fishing: males appreciably more	17306 (1409)	2	0.12
v28_4	intensity of agriculture: intensive irrigated agriculture	23601 (598)	2	0.08
v28_2	intensity of agriculture: extensive or shifting agriculture	23601 (17890)	2	0.07
v53_1	sex differences: animal husbandry: males only	7697 (2183)	2	0.07
Panel B: By Sum of Absolute Coefficients				
v29_3	major crop type: cereal grains	23601 (16630)	6	0.18
v29_2	major crop type: roots or tubers	23601 (6610)	5	0.17
v52_2	sex differences: fishing: males appreciably more	17306 (1409)	2	0.12
v28_4	intensity of agriculture: intensive irrigated agriculture	23601 (598)	2	0.08
v71_2	former presence of slavery: formerly present but not currently existing	23571 (12253)	4	0.07
v71_1	former presence of slavery: absent or exists currently and in past	23571 (11318)	4	0.07
v6_1	mode of marriage: bride price or wealth, to bride's family	24671 (14518)	1	0.07
v28_2	intensity of agriculture: extensive or shifting agriculture	23601 (17890)	2	0.07
v53_1	sex differences: animal husbandry: males only	7697 (2183)	2	0.07
v47_5	sex differences: pottery making: activity present	18911 (2470)	1	0.06
v66_3	class stratification: dual	21859 (5594)	2	0.06
v38_2	segregation of adolescent boys: partial	9084 (2024)	3	0.05
v77_2	inheritance distribution for movable property: primogeniture	22656 (17558)	4	0.05
v52_1	sex differences: fishing: males only	17306 (12051)	1	0.05

Note: This table presents the ethnic characteristics identified via a Group Lasso selection algorithm. Each characteristic is denoted as $v\kappa\_\rho$, where κ is the primary ethnic variable identifier and ρ indicates the specific subcategory. Feature selection was executed across 20 distinct specifications, encompassing four child welfare outcomes (poverty, stunting, underweight and wasting) and five samples (the pooled multi-country sample, plus Ethiopia, Ghana, Malawi and Nigeria). Panel A highlights traits ranking in the top 10 features in at least four specifications. Panel B displays traits with a cumulative absolute coefficient exceeding 0.05 across all specifications. The "Sample Size" column reports households with non-missing data for the parent variable, with the count of households possessing the specific trait in parentheses. The total sample (24,671) comprises estimation-sample households matched to ethnic groups containing at least one child. Missing data was imputed using K-Nearest Neighbors (KNN) imputation.

Appendix E: The Role of Ancestral Ethnic Traits

E.1: Resource Shares and Poverty

Table E1 presents the per-individual resource shares for the Phase 2 estimation. Although the sharing rules now account for diverse ancestral traits, the structural estimates remain stable and qualitatively similar to the baseline. In Malawi, the per-child share remains identical at 21% across both phases. Conversely, in Ethiopia, Ghana and Nigeria, accounting for ethnic traits leads to a slight upward revision of men's resource shares by $2-4$ percentage points. This indicates that traditional ethnic norms may be a significant driver of male bargaining power—a factor not typically included in standard household-level analysis.

While the shifts in resource shares are small (2–3 percentage points), Table E2 reveals that these small reallocations have a disproportionate impact on poverty headcounts, particularly for children in Ghana and Nigeria. In Nigeria, the calorie-adjusted child poverty rate rises from 49% to 60% when ethnic traits are included. This suggests that a significant number of children live very close to the poverty threshold. Hence, even a marginal 2-percentage-point decrease in their resource share (from 0.09 to 0.07) is sufficient to push a large segment of the population below the individual poverty line. This finding underscores the importance of precisely identifying intrahousehold allocation rules, as

Table E1: Predicted Resource Shares Accounting for Ethnic Characteristics

Country	Men, Women, Children Households				All Family-Type Households			
	Number of Households	Per-Man Share	Per-Woman Share	Per-Child Share	Number of Households	Per-Man Share	Per-Woman Share	Per-Child Share
Ethiopia	3864	0.25 (0.12)	0.18 (0.09)	0.16 (0.06)	5508	0.30 (0.16)	0.22 (0.13)	0.17 (0.07)
Ghana	5906	0.25 (0.10)	0.20 (0.11)	0.10 (0.04)	8895	0.28 (0.12)	0.24 (0.16)	0.11 (0.04)
Malawi	6514	0.19 (0.08)	0.16 (0.09)	0.21 (0.08)	8914	0.22 (0.12)	0.20 (0.13)	0.21 (0.08)
Nigeria	2716	0.25 (0.08)	0.19 (0.12)	0.07 (0.03)	3543	0.26 (0.09)	0.22 (0.15)	0.08 (0.04)

Note: This table displays the mean and standard deviation of estimated per-person resource shares, η_t/N_t for men, women, and children. Family-Type households encompass nuclear, single-parent, and childless-couple compositions. Contrary to the baseline results, these estimates control for the identified ethnic characteristics.

small biases in share estimates can lead to significant errors in assessing the depth of child poverty.

Table E2: Predicted Poverty Rates Accounting for Ethnic Characteristics

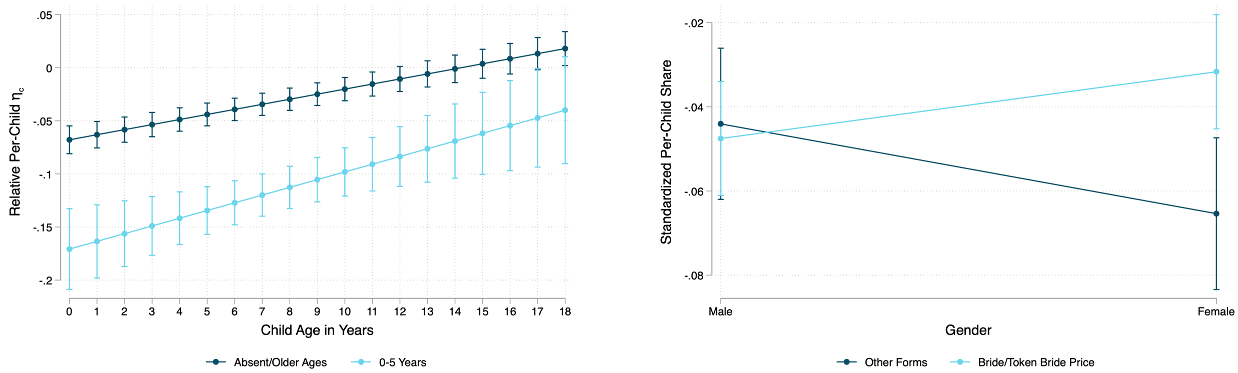
Country	Per-Capita Poverty Rate	Unadjusted Individual Poverty Rate			Calorie-Adjusted Individual Poverty Rate		
		Men	Women	Children	Men	Women	Children
Ethiopia	0.40	0.26	0.39	0.56	0.28	0.31	0.36
Ghana	0.37	0.21	0.33	0.68	0.23	0.27	0.46
Malawi	0.70	0.62	0.73	0.70	0.66	0.64	0.44
Nigeria	0.42	0.11	0.34	0.84	0.14	0.26	0.60

Note: This table displays estimated poverty headcount rates based on an extreme poverty line of \$2.15/day (2017 US\$ PPP). Unadjusted poverty rates are calculated as the individual's estimated resource share multiplied by total household expenditure. For calorie-adjusted rates, I utilize a calorie-based poverty line following Brown et al. (2021) assuming \$2.15/day is the baseline threshold for adults aged 15-45. Individual poverty lines are rescaled according to the 2015–2020 Dietary Guidelines for Americans based on age and gender. For example, a male aged 16–25 has a recommended intake of 2,800 calories, resulting in a scaled poverty line of \$2.51/day.

When analyzing the intersection of ethnic traits with child demographics, two compelling patterns emerge: (1) gender interacts significantly with marriage transactions (θ_6) and (2) age interacts with male circumcision practices (θ_5). For male circumcision, Figure E1a illustrates an initial "entry cost." In societies where circumcision historically occurs between ages 0 and 5, infants and young toddlers receive a significantly lower resource share compared to their peers in societies where the practice is absent or occurs later in life. However, as these boys mature, their resource shares converge toward the baseline. This "catch-up" highlights the strength of age-set societies, where the lifecycle ritual serves as a prerequisite for resource pooling and social status, which eventually stabilizes the child's welfare.

Regarding marriage transactions, the results in Figure E1b reveal that the observed premium in societies practicing bride price is entirely driven by female children. While resource shares for boys remain statistically identical across marriage systems, girls in bride-price societies receive a significantly higher allocation than those in societies with other marriage forms. This suggests that in cultures where daughters represent a future resource transfer to the household, they are treated as high-value human capital investments, narrowing the traditional intrahousehold gender gap.

Figure E1: Impact of Age and Gender



(a) Male Circumcision

(b) Marriage Transactions

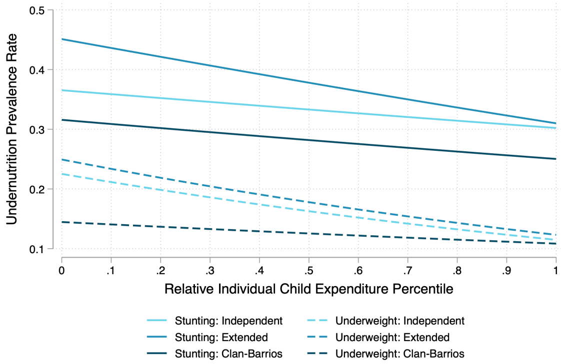
Note: These figures depict the predictive margins of the standardized per-child resource shares derived from OLS estimates. Panel (a) illustrates the interaction between child age and the timing of ancestral male circumcision. Panel (b) displays the interaction between child gender and ancestral bride price marriage transactions. All specifications control for child age, sex, household demographic composition, wealth indicators and country fixed effects. Vertical bars represent 95% confidence intervals and standard errors are clustered at the primary sampling unit level.

E.2: Undernutrition

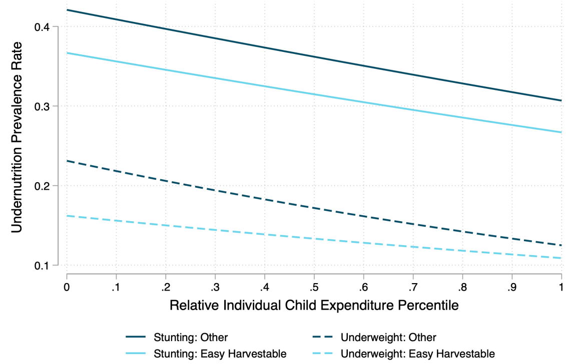
Figure E2 examines how ethnic characteristics relate to child undernutrition by tracking stunting (solid lines) and underweight (dashed lines) prevalence rates across different individual child expenditure percentiles. The underlying logistic regression controls for household composition, child demographics and country fixed effects. While undernutrition rates generally decline as individual child expenditure increases, significant prevalence remains across all expenditure levels, indicating that undernutrition is not solely a problem of the extreme poor. The social structure of an ethnic group also appears to serve as a protective buffer (Figure E2a). Children in clan-barrio societies are less likely to be undernourished across all expenditure percentiles, implying more equitable resource sharing within those communities. In contrast, children from groups with extended family structures face higher rates of stunting and underweight, regardless of how much is spent on them.

Ethnic groups that historically relied on high-intensity agriculture (Figure E2c) or "easy-harvestable" crops (Figure E2b) see lower undernutrition today. Similarly, while bride price practices are associated with lower stunting (Figure E2e), settlement pattern history presents a trade-off (Figure E2f) where children from historically nomadic groups are more likely to be underweight, whereas those from settled, non-nomadic groups are more prone to stunting. Finally, the data reveal a paradox regarding male circumcision

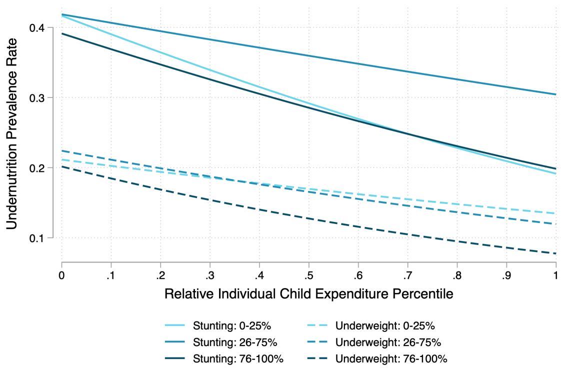
Figure E2: Undernutrition Prevalence by Individual Child Expenditure Percentile



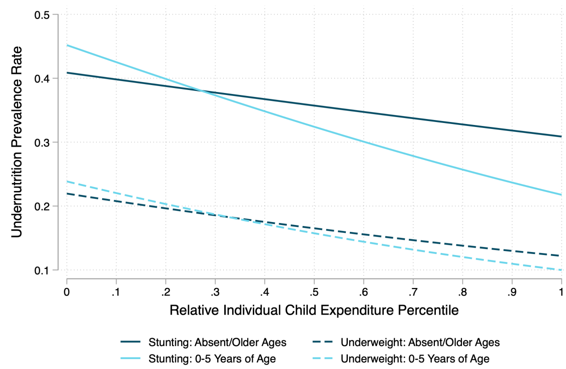
(a) Jurisdictional Hierarchy



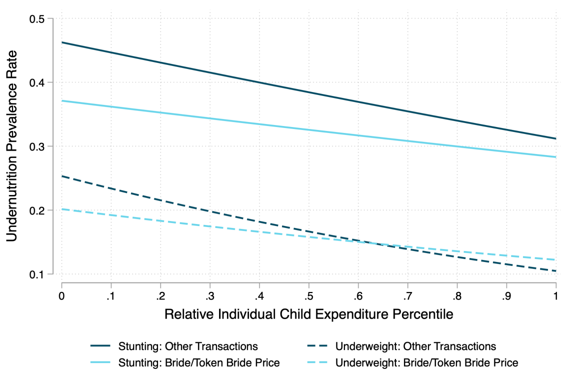
(b) Principal Type of Crop



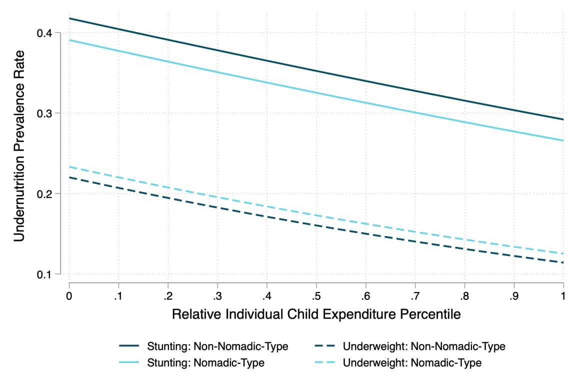
(c) Dependence on Agriculture



(d) Male Circumcision



(e) Transactions at Marriage



(f) Settlement Patterns

Note: These figures depict the predicted undernutrition rates (stunting and underweight) for children across individual child expenditure percentiles, stratified by ethnic characteristics. Undernutrition outcomes are defined according to WHO growth standards (Z-scores < -2). Estimates are derived from a logistic regression model controlling for household composition (number and average age of men, women and children), the proportion of boys, the gender education gap, urban/rural residence and the child's age and sex. Vertical dashed lines represent percentiles corresponding to the World Bank extreme poverty line ($1.29/day), mapped for Nigeria (dark grey), Ghana/Ethiopia (medium grey) and Malawi (light grey). Children are classified as stunted (HAZ), underweight (WAZ) or wasted (WHZ) if their Z-scores fall below -2. The sample includes children aged 0–5 years, excluding observations with biologically implausible Z-scores (outside the $[-5, 5]$ range).

rituals (Figure E2d). In societies where these rituals occur at a very young age, poor children are actually more likely to be stunted and underweight. This disadvantage disappears once expenditure crosses the 30^{th} percentile. Beyond this point, these children actually fare better than their peers.

Appendix F: Robustness

This section outlines the robustness results concerning per-child resource shares and children's height-for-age Z-scores (HAZ). To ensure that the observed variations are driven by cultural traits rather than spatial confounding or correlated socioeconomic factors, I use Kolmogorov-Smirnov (KS) tests to assess the equality of distributions across eth-

Table F1: Resource Share Robustness - Border Analysis

	Kolmogorov-Smirnov Tests						
	1*km*	5*km*	10*km*	15*km*	20*km*	25*km*	50*km*
	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)
Jurisdictional Hierarchy							
2 - Independent Families	0.256*** (0.000)	0.099*** (0.000)	0.072*** (0.000)	0.073*** (0.000)	0.070*** (0.000)	0.070*** (0.000)	0.047*** (0.000)
Principal Type of Crop							
Easy Harvestable	0.131*** (0.000)	0.087*** (0.000)	0.090*** (0.000)	0.101*** (0.000)	0.099*** (0.000)	0.101*** (0.000)	0.112*** (0.000)
Dependence on Agriculture							
$76-100\%$	0.118** (0.012)	0.118*** (0.000)	0.110*** (0.000)	0.106*** (0.000)	0.125*** (0.000)	0.166*** (0.000)	0.217*** (0.000)
Male Curcumcision							
$0-5$ Years of age	0.101*** (0.000)	0.070*** (0.000)	0.061*** (0.000)	0.090*** (0.000)	0.098*** (0.000)	0.103*** (0.000)	0.144*** (0.000)
Transactions at Marriage							
Bride/Token Bride Price	0.097*** (0.001)	0.040** (0.016)	0.055*** (0.000)	0.080*** (0.000)	0.088*** (0.000)	0.080*** (0.000)	0.080*** (0.000)
Settlement Patterns							
Nomadic-Type Societies	0.078* (0.051)	0.108*** (0.000)	0.135*** (0.000)	0.143*** (0.000)	0.152*** (0.000)	0.158*** (0.000)	0.177*** (0.000)

Note: This table displays the *D*-Statistic and corresponding *p*-values (in parentheses) from the Kolmogorov-Smirnov test for equality of distribution for relative per-child resource shares across ethnic characteristics. Only children from Ethiopia, Malawi, and Nigeria are included. ***, **, * - indicates significance at a 1%, 5%, and 10% level.

nic characteristics. Specifically, I test for distributional differences among ethnic groups living in close geographic proximity (border analysis) and among those sharing similar socioeconomic status (SES).

Table F2: Undernutrition Robustness - Border Analysis

	Kolmogorov-Smirnov Tests						
	$1km$	$5km$	$10km$	$15km$	$20km$	$25km$	$50km$
	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)
Jurisdictional Hierarchy							
2 - Independent Families	0.077 (0.643)	0.058 (0.281)	0.070** (0.024)	0.079*** (0.002)	0.074*** (0.002)	0.073*** (0.001)	0.035 (0.170)
Principal Type of Crop							
Easy Harvestable	0.078 (0.599)	0.100*** (0.002)	0.110*** (0.000)	0.118*** (0.000)	0.119*** (0.000)	0.134*** (0.000)	0.149*** (0.000)
Dependence on Agriculture							
$76-100\%$	0.072 (0.941)	0.051 (0.832)	0.050 (0.530)	0.076** (0.041)	0.084** (0.011)	0.104*** (0.000)	0.122*** (0.000)
Male Circumcision							
$0-5$ Years of age	0.085 (0.450)	0.058 (0.216)	0.059* (0.061)	0.072*** (0.002)	0.068*** (0.002)	0.074*** (0.000)	0.074*** (0.000)
Transactions at Marriage							
Bride/Token Bride Price	0.128** (0.049)	0.051 (0.195)	0.053** (0.029)	0.057*** (0.003)	0.053*** (0.003)	0.051*** (0.002)	0.043*** (0.001)
Settlement Patterns							
Nomadic-Type Societies	0.106 (0.404)	0.110*** (0.003)	0.111*** (0.000)	0.136*** (0.000)	0.137*** (0.000)	0.151*** (0.000)	0.149*** (0.000)

Note: This table displays the D-Statistic and corresponding p-values (in parentheses) from the Kolmogorov-Smirnov test for equality of children's height-for-age Z-scores. Only children from Ethiopia, Malawi and Nigeria are included. ***, **, * - indicates significance at a 1%, 5%, and 10% level.

Tables F1 and F2 present the border analysis, which restricts the sample to households living within tight spatial bandwidths (ranging from $1km$ to $50km$) of ethnic boundaries. This strategy effectively controls for unobserved local variations in climate, geography and institutional access. For per-child resource shares, the null hypothesis of equal distributions is overwhelmingly rejected ($p < 0.01$) across almost all ancestral ethnic traits and all distance bandwidths. This confirms that intrahousehold allocation rules are deeply cultural and persist even among close geographic neighbors. For children's HAZ, the results are more nuanced. At very tight bandwidths ($1km$ and $5km$), the distributions of HAZ for several traits (such as jurisdictional hierarchy and male circumcision) are not statistically different. This is economically intuitive since neighboring households share

Table F3: Resource Shares Robustness - Socioeconomic Status

	Kolmogorov-Smirnov Tests			
	Electricity, Non-Agriculture Business, and Phone	Employed, Own a Home, and Phone	No Plough, Low Skilled, and No Financial Asset	PCA Groups
	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)
Jurisdictional Hierarchy				
2 - Independent Families	0.335*** (0.000)	0.486*** (0.000)	0.501*** (0.000)	0.364*** (0.000)
Principal Type of Crop				
Easy Harvestable	0.109*** (0.000)	0.306*** (0.000)	0.321*** (0.000)	0.149*** (0.000)
Dependence on Agriculture				
$76-100\%$	0.326*** (0.000)	0.368*** (0.000)	0.421*** (0.000)	0.346*** (0.000)
Male Circumcision				
$0-5$ Years of age	0.328*** (0.000)	0.391*** (0.000)	0.413*** (0.000)	0.349*** (0.000)
Transactions at Marriage				
Bride/Token Bride Price	0.336*** (0.000)	0.527*** (0.000)	0.576*** (0.000)	0.341*** (0.000)
Settlement Patterns				
Nomadic-Type Societies	0.569*** (0.000)	0.459*** (0.000)	0.526*** (0.000)	0.765*** (0.000)

Note: This table displays the D-Statistic and corresponding p-values (in parentheses) from the Kolmogorov-Smirnov test for equality of distribution for per-child resource shares across ethnic characteristics. ***, **, * - indicates significance at a 1%, 5%, and 10% level.

the same local disease environment and health infrastructure, which heavily dictate anthropometric outcomes. However, as the spatial bandwidth expands beyond $10km$, ethnic differences in child health outcomes become highly significant.

Tables F3 and F4 test whether ethnic characteristics are merely proxying for socioeconomic status. Even when grouping households by similar socioeconomic status—such as households with electricity, owning a non-agricultural business and a phone—the distribution of per-child resource shares remains significantly different ($p < 0.01$) across all ancestral cultural traits. When examining HAZ, controlling for SES absorbs more of the variation. For instance, the distributional differences in undernutrition across male circumcision practices disappear when comparing households of similar wealth. Conversely, traits like marriage transactions (bride price) remain a highly significant differentiator of child nutrition ($p < 0.01$) across almost all socioeconomic groupings.

Table F4: Undernutrition Robustness - Socioeconomic Status

	Kolmogorov-Smirnov Tests			
	Electricity, Non-Agriculture Business, and Phone	Employed, Own a Home, and Phone	No Plough, Low Skilled, and No Financial Asset	PCA Groups
	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)	D-Stat (p-value)
Jurisdictional Hierarchy				
2 - Independent Families	0.086 (0.530)	0.134** (0.015)	0.120 (0.485)	0.108 (0.463)
Principal Type of Crop				
Easy Harvestable	0.078*** (0.001)	0.105*** (0.000)	0.058 (0.432)	0.095*** (0.006)
Dependence on Agriculture				
$76 - 100\%$	0.077* (0.076)	0.064* (0.059)	0.039 (0.985)	0.167*** (0.001)
Male Circumcision				
$0 - 5$ Years of age	0.042 (0.500)	0.045 (0.177)	0.056 (0.679)	0.071 (0.237)
Transactions at Marriage				
Bride/Token Bride Price	0.177*** (0.000)	0.114*** (0.000)	0.160*** (0.000)	0.167*** (0.000)
Settlement Patterns				
Nomadic-Type Societies	0.213 (0.192)	0.207*** (0.000)	0.145 (0.511)	0.263 (0.985)

Note: This table displays the *D*-Statistic and corresponding *p*-values (in parentheses) from the Kolmogorov-Smirnov test for equality of distribution for children's height-for-age z-scores. ***, **, * - indicates significance at a 1%, 5%, and 10% level.

References

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, *128*(2), 469–530.

Aminjonov, U., Colacce, M., Bargain, O., & Tiberti, L. (2024). Culture, Intrahousehold Distribution, and Individual Poverty. *Economic Development and Cultural Change*, *73*(1), 127–165.

Ashraf, N., Bau, N., Nunn, N., & Voena, A. (2020). Bride Price and Female Education. *Journal of Political Economy*, *128*(2), 591–641.

Bahrami-Rad, D., Becker, A., & Henrich, J. (2021). Tabulated Nonsense? Testing the Validity of the Ethnographic Atlas. *Economics Letters*, *204*, 109880.

Brown, C., Calvi, R., & Penglase, J. (2021). Sharing the Pie: An Analysis of Undernutrition and Individual Consumption in Bangladesh. *Journal of Public Economics*, *200*, 104460.

Brown, C., Ravallion, M., & Van De Walle, D. (2019). Most of Africa's Nutritionally Deprived Women and Children Are Not Found in Poor Households. *The Review of Economics and Statistics*, *101*(4), 631–644.

Carletto, C., & Gourlay, S. (2019). A Thing of the Past? Household Surveys in a Rapidly Evolving (Agricultural) Data Landscape: Insights from the LSMS-ISA. *Agricultural Economics*, *50*, 51–62.

Giuliano, P., & Nunn, N. (2018). Ancestral Characteristics of Modern Populations. *Economic History of Developing Regions*, *33*(1), 1–17.

Giuliano, P., & Nunn, N. (2021). Understanding Cultural Persistence and Change. *The Review of Economic Studies*, *88*(4), 1541–1581.

Lechene, V., Pendakur, K., & Wolf, A. (2022). Ordinary Least Squares Estimation of the Intrahousehold Distribution of Expenditure. *Journal of Political Economy*, *130*(3), 681–731.

Murdock, G. P., Textor, H. B., White, D. R., Gray, J., & Divale, W. T. (1999). A Corrected Ethnographic Atlas. *World Cultures*, *10*(1), 24–85.

Nunn, N., & Wantchekon, L. (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252.