Child Poverty Among Refugees

Theresa P. Beltramo[*]

Rossella Calvi[⋆]

Giacomo De Giorgi[†]

Ibrahima Sarr[‡]

December 2022

Abstract

There are now more violent conflicts globally than at any time in the past three decades, resulting in the largest forced displacement crisis ever recorded. Understanding at a granular level the well-being of refugees is essential to inform successful poverty alleviation strategies and unlock refugees' potential. As forced displacement can lead to a reorganization of a family's structure, we use a structural model in combination with data from refugee camps and surrounding communities in Uganda and Kenya to estimate the allocation of consumption within families. We compute poverty rates that account for intra-household inequality, finding that refugee children can be up to three times more likely to be poor than adults. So, refugee children not only suffer from the experience of forced migration, but also from potentially low nutrition and a disproportionately higher poverty risk. Using a supervised machine learning algorithm, we show that a small set of observable traits, such as a child's age, household composition, and access to sanitation and clean water, predict child poverty in refugee settlements and surrounding communities remarkably well, often better than per-capita household expenditure.

Keywords: refugees; collective model; intra-household allocation; child poverty; proxy-means test; poverty targeting.

[*]UNHCR.

[⋆]Rice University.

[†]University of Geneva.

[‡]UNHCR.

This paper has benefited from helpful comments from Caitlin Brown, Eeshani Kandpal, Stephen O'Connell, and Jacob Penglase. All errors are our own.

1 Introduction

Empirical evidence shows that consumption expenditures are not shared equally within families. Researchers have documented the inferior outcomes of vulnerable household members, and intra-household inequality in food intake, anthropometric measures, and non-food expenditures.[1] Intra-household inequality is potentially magnified in poor settings, where resources are limited and the competition among household members is high. The inaccurate measurement of individual consumption within a household can lead to the underestimation of poverty rates (Brown et al. (2019, 2021a)), which can falsely inform related poverty alleviation programs, and hence limit their ability to reach the world's poorest.

Refugees face an exceptionally high risk of living in poverty as they typically reside in developing countries, often in low-income border regions, and tend to work in the informal sector.[2] There is also reason to believe that refugees might allocate resources within households differently than non-refugees as forced displacement can lead to a reorganization of a family's structure.[3] In these contexts, understanding the intra-household allocation of expenses and its consequences for poverty measurement and targeting is of primary policy relevance. Additionally, understanding the relative deprivation of adults and children is essential as the lives of those children are shaped by the early events of forced migration as well as the potential poor nutrition and welfare due to the unequal resource sharing within the household.[4]

By the end of 2021, 89.3 million individuals were forcibly displaced worldwide as a result of persecution, conflict, violence or human rights violations. Africa has experienced a high number of new displacements, with the East and Horn of Africa, and the Great Lakes region hosting nearly 5 million refugees (67 percent of the refugees on the African continent and 20 percent of the global refugee population; UNHCR (2021)). An estimated 36.5 million (41 percent) of the forcibly displaced people are children below age 18 (UNHCR, 2021). To date, no study has analyzed con-

[1] See e.g. Chen and Drèze (1992), Drèze and Srinivasan (1997), Jensen (2005), van de Walle (2013), Djuikom and van de Walle (2018) for evidence on widows; Bicego et al. (2003), Case et al. (2004), Evans and Miguel (2007) for orphans; Subramanian and Deaton (1990), Lancaster et al. (2008), Oster (2009), Jayachandran and Kuziemko (2011) for girls; and Behrman and Tubman (1986), Behrman (1988), Black et al. (2005), Price (2008), Booth and Kee (2009), De Haan (2010), Black et al. (2011), Jayachandran and Pande (2017) for later-born children.

[2] A 2019 UNHCR survey of 111 countries representing 97 percent of the global refugee population found that 85 percent live in developing countries and 70 percent of refugees live in countries with restricted right to work.

[3] Refugee household composition can be unique as only a subset of the original members may be sustained post conflict (due to, e.g., conscription in the military for male members, and the death, kidnapping or separation of certain family members during displacement). Unaccompanied minors are also more prevalent in refugee camps.

[4] There is indeed ample evidence of how nature and nurture in the first few years of life are responsible for a large part of later in life development. See, among many others, Grantham-McGregor SM (1991), Shonkoff and Phillips (2000), Knudsen et al. (2006), Cunha and Heckman (2007) and Martorell (2017).

sumption inequality within refugee families and its consequences for poverty measurement.[5] We fill this gap by investigating the intra-household allocation of consumption in refugee settlements and the surrounding communities in rural East Africa. In doing so, we place special emphasis on the measurement and targeting of child poverty.

Typically, surveys do not collect consumption information at the individual level; they only record household-level consumption or expenditures. To overcome this limitation, a thriving literature has applied a structural approach based on the collective household model (Chiappori, 1988, 1992). This approach combines observable household-level expenditures on assignable goods (goods that are consumed exclusively by, e.g., women, men or children) with preference restrictions to recover individual-level consumption from household-level data (Dunbar et al., 2013).[6] Specifically, the structural approach allows one to identify and estimate *resource shares* (the fraction of *total* household consumption allocated to each family member), which are otherwise unobserved. Resource shares are then used to compute consumption and poverty rates at the individual (rather than the household) level. These poverty estimates are fundamentally different from standard poverty rates, which are based on observed household per-capita consumption and therefore assume an equal distribution of resources among family members. Our paper is the first to apply this approach to study individual poverty (child poverty in particular) in refugee settlements and the surrounding communities.

We focus on two East African countries: Kenya and Uganda. These two countries host a combined population of more than 2 million refugees and asylum seekers (UNHCR, 2021). Notably, Uganda is the largest refugee hosting country in Africa, with 1,529,272 refugees and asylum seekers as of June 2022 (UNHCR, 2022b), while Kenya is the third-largest refugee-hosting country in Africa, after Uganda and Ethiopia, with 555,183 refugees and asylum seekers as of June 2022 (UNHCR, 2022a). Most of refugees in Kenya are from Somalia (53 percent), South Sudan (25 percent), the Democratic Republic of Congo (9 percent), Ethiopia (6 per cent), and Burundi (4 percent). In Uganda, refugees are mainly from South Sudan (65 percent), the Democratic Republic of the Congo (31 percent), and Somalia (4 percent).

For Uganda, we use data from the 2018 Uganda Refugee and Host Communities Household

[5] Sozbir (2022) studies the intra-household effects of Syrian refugee inflows in Turkey. His focus, however, is on native families only.

[6] This approach has been used to study inequality between spouses or between parents and children (Dunbar et al., 2013; Bargain et al., 2021; Tommasi, 2019; Sokullu and Valente, 2021; Lechene et al., 2020; Casco, 2022; Hernandez-de Benito, 2022), the well-being of older women in India (Calvi, 2020), the treatment of foster children in Malawi (Penglase, 2020), and the allocation of resources among prime-aged adults, the elderly, and children by sex and birth-order in Bangladesh (Brown et al., 2021a). See Brown et al. (2021b) for an overview.

Survey, which was collected by the Uganda Bureau of Statistics and the World Bank, and covers refugee households in the largest settlements in the country and non-refugee (hereinafter referred to as host) families in the surrounding communities. This survey is among the first-ever detailed consumption surveys collected for refugees and hosts that are representative of both communities. For Kenyan refugees, data come from the joint World Bank and UNHCR 2018-2019 Kalobeyei Socio-Economic Assessment; for Kenyan nationals, we use data from the 2015-2016 Kenya Integrated Household Budget Survey led by the Kenyan National Bureau of Statistics (KNBS) and the World Bank. Importantly for our analysis, all these surveys contain consumption expenditures on assignable goods, which we use to estimate resource shares and individual-level poverty as we described above. They also include a battery of observable family characteristics, which we exploit in our analysis of poverty targeting.

According to our estimates, children are allocated a disproportionately low fraction of household consumption relative to adults (up to 80 percent lower). This finding holds true both in refugee and host communities and in Kenya and Uganda alike. In all refugee settlements but with varying intensity, adults' consumption is above the household per-capita consumption (which does not account for intra-household allocation), while children's consumption is substantially below. The relative importance of intra-household to between-household consumption inequality however is highly heterogeneous across our areas of study (which consist of a large refugee settlement in Kenya and several camps in the South West and West Nile regions in Uganda) and between host and refugee communities.

Differences in intra-household allocation and in incomes between areas and communities result in significant differences in women's, men's and children's relative likelihood to live in poverty. Even accounting for differences in needs by age and gender, we find that children face a substantially higher risk to achieve a level of consumption that is above the World Bank extreme poverty line of 1.90US$/day. For instance, the poverty rate among refugee children ranges from 39 in Kenya to 69 percent in the South West region of Uganda. Among hosts, child poverty ranges from 27 (in the West Nile region of Uganda) to 69 percent (in the South West Uganda). The intensity of child poverty among refugees is particularly notable, with the total poverty gap for children in refugee settlements estimated to be as much as five times larger than in the surrounding non-refugee communities. So, according to our estimates, the monetary disbursement required to eradicate child poverty in refugee camps can be up to five times higher than in the surrounding host communities.

Taken together, our findings call for tailored policies to reduce inequality and poverty among refugee and hosts, and to reach the most vulnerable individuals in such contexts. To this aim, we apply a supervised machine learning algorithm to identify the most critical predictors of child poverty. We show that a small set of observable traits, including a child's age, household composition, and access to sanitation and clean water, predict child poverty in refugee settlements and surrounding communities remarkably well. Based on these predictors, we develop low-cost and parsimonious approaches to target child poverty in our areas of study. We show that our proposed targeting approaches can outperform per-capita household expenditure and improve upon standard targeting strategies.

In terms of policy implications, our analysis adds to previous works showing that accounting for intra-household inequality is critical for poverty measurement (see, for instance, Dunbar et al. (2013); Brown et al. (2021a,b); Lechene et al. (2020)). Individual-level poverty measures are recommended by a recent World Bank report, which outlines the key considerations for monitoring global poverty (Atkinson, 2016). Measuring poverty at a granular level is also essential to meet the UN General Assembly Sustainable Development Goal 1 to "end poverty in all its forms" by 2030. Furthermore, as early events and nutrition can leave permanent marks on one's outcomes later in life (Shonkoff and Phillips, 2000; Cunha and Heckman, 2007), the finding that children are the poorest household members in both refugee and host communities has important policy implications for both humanitarian and development programs. Finally, our findings may help address the current drop in funding faced by UNHCR and other humanitarians organizations as a result of the Ukraine crisis and other macroeconomic factors. Over 7.7 million refugees have fled Ukraine since February 2022, which has called for substantial efforts towards the prioritization of targeting programs in other regions of the world. By identifying the most vulnerable individuals in refugee settlements and surrounding communities in Kenya and Uganda, our work contributes to this efforts.

The rest of the paper is organized as follows. Section 2 provides background information on the refugee populations in Kenya and Uganda. Section 3 discusses the methodology and data used. Section 4 details the main results of intra-household consumption and poverty. Section 5 outlines the results for poverty within households for children and considerations about poverty targeting. Section 6 concludes.

2 Background

Our analysis focuses on three areas in East Africa: the Kalobeyei refugee camp in Kenya and several refugee settlements in the West Nile and South West regions of Uganda. Together, these settlements host approximately 2 million refugees, accounting for 29 percent of refugees in the African continent and over 2 percent of all refugees worldwide. Figure A1 in the Appendix shows the location of refugee settlements in Kenya and Uganda and provides information about their size and composition.

The Kalobeyei Settlement. The Kalobeyei settlement was established in 2016 in Turkana County in North West Kenya, about twenty miles from Kakuma Town. Kalobeyei was established to promote the self-reliance of refugees and the surrounding host population and deliver integrated services to both. It is an example of an innovative approach designed to offer integrated market-based opportunities for refugees and hosts. Its development has been guided by the Kalobeyei Integrated Social and Economic Development Programme (KISEDP), led by the Government of Kenya (particularly the local Turkana County Government), UNHCR, and partners (UNHCR, 2018).[7] As of June 2022 (UNHCR, 2022a), Kalobeyei hosts 43,472 refugees, mainly from South Sudan (76 percent), Ethiopia (12 percent), and Burundi (6 percent). Most of the refugees arrived in the five years prior to the survey, either being displaced internally from the Dadaab Refugee Camp located in North Eastern Kenya, or arriving from South Sudan.

At the time of data collection in Kalobeyei, refugees were offered food assistance through a voucher program called Bamba Chakula (Swahili for "get your food"). The program, which accounted for 98 percent of food assistance in the settlement, allowed refugees to shop for food at designated facilities. Every month, refugee households in Kalobeyei received electronic transfers amounting to 1,400KES per person (approximately 13US$). All households were equally targeted, with the exact amount received by household based on family size.

[7] The Kalobeyei Integrated Socio-Economic Development Programme (KISEDP) offers a strategic roadmap for the evolution of Turkana West over 15 years. Phase I of the initial five year strategy has been completed with support from various donors and based on the commitments of, and deepening collaboration between, the Government of Kenya, the Turkana County Government (TCG), UNHCR, sister UN agencies and a range of humanitarian, development and private partners.

Refugee Settlements in Uganda. As of June 2022, Uganda has been hosting approximately 1.5 million refugees and has the highest refugee population in Africa. For the most part, refugees are located in the capital Kampala and in the North and South-Western regions. Refugees located outside urban areas live in settlements, where they co-exist with the host communities.[8] This approach, combined with progressive refugee laws and freedoms, provides refugees in Uganda with significant prospects for dignity and self-reliance.[9] The majority of the refugees are from South Sudan (65 percent), followed by 31 percent from the Democratic Republic of Congo, 4 percent from Somalia, 3 percent from Burundi.In 2015, Uganda experienced the most recent influx of refugees from South Sudan, which led to a five-fold increase in the total refugee population. Currently, the South Sudanese refugee population in the West Nile area amount to just under 900,000; and a little over half a million refugees are located in the South West region.

At the time of data collection in 2018, refugees were given food assistance based on the time since arrival. The World Food Program (WFP) provided a full ration to all refugees who arrived after July 2015, and a half ration to all refugees who arrived before July 2015 (NET, 2018). This targeting strategy, which was designed by UNHCR and WFP in conjunction with the Government of Uganda to address budget constraints, was controversial in its ability to protect the most vulnerable (McAloon, 2014). As a result, it was re-evaluated in October 2018 and replaced by equal rations to all refugees. In 2021, the targeting of food assistance was revisited again, with assistance being targeted based on a vulnerability assessment.

Challenges. In both Uganda and Kenya, UNHCR and partners have implemented integrated community-based management of acute malnutrition in the settlements, including in-patient and out-patient management of severe malnutrition, maternal and child health nutrition program and additional supplementary feeding program during the emergency phase (UNHCR, 2018; UNHCR et al., 2019; Asiimwe, 2021). Despite the numerous programs in place to fight poverty among refugees, the situation remains critical. In Kalobeyei, the incidence of malnutrition has increased in 2018 due to food ration cuts and breaks in the food pipeline. Evidence has shown these late disbursements created gaps in food planning and management, resulting in households taking out loans on credit from retailers and further impoverishment.

[8] Seven refugee settlements are located in the West Nile region and five are in the South West region.

[9] Uganda's approach to hosting refugees is one of the most generous and progressive of the world (World Bank, 2019). Ugandans have been themselves refugees in Sudan and the DRC during the 1980s and this experience is widely known to have influenced their welcoming environment to refugees.

According to the Standardized Expanded Nutrition Survey, the prevalence of stunting among refugee children aged 6 – 59 months living in Kalobeyei was 32 percent in 2019 (which UNHCR categorizes as *very high* or *critical*) (UNHCR, 2019). The anaemia prevalence in children was 57.5 percent (*critical*) (UNHCR, 2019). A similar assessment conducted between 2017 and 2018 by the Government of Uganda, UNHCR and WFP concluded that the food assistance provided to refugees in Uganda was insufficient to meet individuals' energy requirements and to provide essential micronutrients.[10] The prevalence of anaemia in children aged 6-59 months and in non-pregnant women of reproductive age (15-49 years) remained above 40 percent (the WHO threshold for raising public health concerns).

While nutrition and food consumption are clearly important components of individual well-being, other dimensions of consumption (such as healthcare and housing) may matter significantly. So, to correctly measure poverty and identify the most deprived individuals in these vulnerable contexts, one must determine the total (food and non-food) consumption each person can access.

Measuring individual-level consumption, however, is a challenge as surveys are typically conducted at the household level and goods can be shared. In what follows, we employ a structural approach to estimate how total consumption is divided among family members. As two out of three individuals in refugee camps in Kenya and Uganda are below the age of 18, we place special emphasis on the measurement of child consumption and poverty in and around refugee settlements. We wish to emphasize that our goal is not to evaluate the effect of any specific policy on poverty alleviation, but to identify who are the poorest people among refugees and hosts, improve upon current poverty measurement in these extremely vulnerable contexts, and guide the design of effective anti-poverty policies targeting refugee settlements and the surrounding communities in our areas of study.

3 Methods, Data, and Descriptive Statistics

In this section, we provide a brief non-technical description of how to identify and estimate the intra-household allocation of consumption using the collective household framework (a detailed

[10] The food assistance only provided 45 percent, 26 percent and 38 percent of vitamin B12, calcium and iron requirements, respectively. The ration only covered 32 percent of the cost of a nutritious diet for a household, and an estimated additional 4,800 – 11,330 Ugandan shillings per day would be required for a household to purchase all of their nutrient needs. A high percentage of households in the West Nile and South West camps reported not consuming vegetables, fruit, meat, eggs, fish and milk during the previous week. These foods are critical to meet essential micronutrient needs such as vitamin A, vitamin B12, iron and calcium.

formulation of the model is provided in the Appendix).[11] We also describe our data sources and descriptive statistics for refugee and host communities in Kalobeyei and in the South West and West Nile regions of Uganda.

3.1 Identification and Estimation

The collective household model assumes each family member has separate preferences and the intra-household allocation is Pareto efficient so that there is no waste of household resources (see Chiappori (1988, 1992) and Apps and Rees (1988) for seminal papers). The assumption of Pareto efficiency by itself, however, is not sufficient to identify how household expenses are allocated among family members from household-level consumption data.[12] Recent methodological advances have shown that one can rely on consumption data about personal (or *private-assignable*) goods and invoke semi-parametric restrictions on individual preferences over such goods to identify the intra-household allocation of resources (Dunbar et al., 2013).[13] Examples of such goods include toys and school material, which are private goods assignable to children, or alcohol and tobacco, which are assignable to adults. For our analysis, we follow previous works and focus on women's, men's, and children's clothing and footwear.

As shown in Browning et al. (2013) and Dunbar et al. (2013), under the convenient assumption of piglog (price independent generalized logarithmic) preferences, the Engel curves for women's, men's and children's assignable goods take the following simple forms:

$$X_j = J\eta_j(\alpha_j + \beta_j \ln \eta_j + \beta_j \ln y), \tag{1}$$

where $j = w, m, c$ denote women, men, and children and η_j are their *resource shares* (i.e., the fraction of household expenditure they consume); α_j and β_j are combinations of underlying pref-

[11] Technical discussions and formal identification proofs are provided in Browning et al. (2013) and Dunbar et al. (2013). Recent articles by Almås et al. (2021), Brown et al. (2021b), and Calvi et al. (2022) provide insightful overviews.

[12] While some papers provide evidence in favor of Pareto efficiency (see e.g. Attanasio and Lechene (2014) for Mexico and Brown et al. (2021a) for Bangladesh), some others cast doubt on this assumption (see e.g. Udry (1996) for Burkina Faso). Note that most rejections of Pareto efficiency are based on decisions about production, not consumption. As discussed in Rangel and Thomas (2019), confounding these two aspects may be misleading. Rangel and Thomas (2019) also show that in nuclear families in Burkina Faso, resource allocations are consistent with efficiency.

[13] A good is defined as private if it cannot be shared or consumed jointly by more than one person and assignable if consumed by a specific household member known to the researcher. Note that private assignable goods provide a number of advantages over goods that are jointly consumed. First, by construction, their demand is only driven by the preferences of those household members who consume them; so, if we observe a household consuming more of women's clothing, we can typically exclude men's preferences for women's clothing as a potential determinant of this difference. Moreover, as a consequence of them being privately consumed, their demand is not directly impacted by economies of scale.

erence parameters; X_m, X_w, X_c are the budget shares spent on men's, women's, and children's assignable clothing and y is the total household expenditure; and $J = W, M, C$ denote the number of men, women, and children residing in the family, respectively.

It is important to stress that budget shares on assignable clothing and resource shares are different objects. Specifically, the relative magnitude of the assignable goods budget shares (the fraction of expenditure devoted to women's, men's or children's clothing and footwear) does not necessarily determine the relative magnitude of resource shares (the fraction of total expenditure allocated to women, men or children). In other words, one cannot just use X_j as a measure of η_j because different household members may have very different tastes for their private assignable good.

Resource shares are identified by imposing similarities of preferences for private assignable goods across household members and the assumption that resource shares are independent of household expenditure, using the methodology developed by Dunbar et al. (2013).[14] These restrictions allow us to identify resource shares by comparing Engel curves for assignable clothing across people within households. In particular, when $\beta_j = \beta$, the slopes of the Engel curves can be identified by linear regression of X_j on a constant and $\ln y$. Resource shares are then identified by the relative slopes and by the constraint that resource shares within a family must sum to one. In practice, we estimate the assignable goods Engel curves. We then implicitly invert these Engel curves to recover the resource shares.

We estimate a system of Engel curves using the non-linear Seemingly Unrelated Regression method, which is iterated until the estimated parameters and the covariance matrix settle. In families with men, women, and children, the system includes three equations; in families without children, men, or women, there are two equations to be estimated.[15] We account for observed heterogeneity across households by allowing preference parameters and resource shares to vary with demographic and socio-economic traits, including the gender, age, years of education, and employment status of the household head, and the household composition by gender and age (number of women, men, girls and boys). For refugees, we also include the number of years since

[14] Empirical tests of the identifying assumptions of Dunbar et al. (2013) are provided, e.g., by Menon et al. (2012), Dunbar et al. (2013), and Bargain et al. (2021).

[15] Note that one can identify as many distinct resource shares as assignable goods. As we describe below, we do not observe clothing expenditures of, e.g., older vs. younger adults, so we cannot recover how resources are allocated among women or men within a household. While this is a limitation, since extended families are widespread, this approach certainly improves upon the assumption that resources are shared equally among all family members. In other words, instead of assuming equal sharing among all, we here assume equal sharing among family members of the same *type*.

they have arrived in the country to allow for heterogeneity along this dimension.

3.2 Data and Measurement

For our analysis, we rely on three data sources: the Kenya Integrated Household Budget Survey (2015/16 KIHBS), covering only Kenyan nationals (KNBS, 2018); the Kalobeyei Socio-Economic Assessment (2018/19 SEA-LV), covering 1,192 refugee households in Kalobeyei (Fix et al., 2019); and the 2018 Uganda Refugee and Host Communities Household Survey (2018 RHCS), covering 1,256 host households and 956 refugee households in Uganda in the Southwest and the West Nile regions (World Bank, 2019). For the Kenyan national sample, we restrict the analysis to communities from the hosting county – Turkana – and its neighboring counties – Marsabit, West Pokot, Samburu, Baringo. This subsample includes 1,930 households. Section B in the Appendix document contains a comprehensive description of our data.

In all three surveys, households are asked to recall their food consumption in the week prior to the survey and their non-food expenditure over various recall periods (one week, one month, three months, and one year). Importantly, the consumption module collects information on household expenditures on clothing and shoes for men, women, and children during the three months prior to the survey. Based on this information, we construct the assignable goods budget shares required in estimation (i.e., X_j, $j = w, m, c$). Consumption amounts also include the value of home-produced goods and services imputed at market value. We exploit information from the remaining survey modules to construct additional variables, including demographic traits (household composition, the head of household's gender and age, country of origin, region of residence), and variables related to employment, asset and livestock ownership. We later use these variables to identify the key predictors of child poverty.

Descriptive Statistics. Table A1 in the Appendix presents descriptive statistics for household and expenditure composition in the Kenyan and Ugandan samples together with two-sided t-test for differences in means across refugee and host communities. Refugee households tend to be larger in size relative to households in the host community and have higher youth dependency ratios (relative share of children to adults).[16] Female headed households are more prevalent in

[16] We wish to note that UNHCR estimates that for Kalobeyei less than one percent of the population is made up of unaccompanied minors. Wherever possible UNHCR works with the community and Government to place these children in extended families.

refugee communities: in both Uganda and Kenya refugee households are twice as likely than hosts to be female-headed. Notably, 72.3 percent of households in refugee communities in Kalobeyei are female-headed. Men are not at all present in about half of refugee families in settlements in South West Uganda. This figure is substantially lower (27 percent) in the surrounding host communities in the region.

Household expenditure (which we convert to US$ PPP) is highly heterogeneous across regions and communities. Notably, the average per-capita household expenditure is higher in Kalobeyei relative to its surrounding non-refugee communities; the opposite holds true in Uganda. Budget shares on clothes are small in general, but significantly smaller in refugee communities. Recall that the variation required for the identification of resource shares comes from the slope (rather that the intercept) of the budget share functions (see Section 3.1). So, that the expenditures on assignable clothing are limited is not a source of concern for identification.[17]

4 Intra-household Consumption and Poverty

In this section, we focus on consumption inequality between and within families. We then assess the incidence of individual-level poverty in refugee communities. Our inequality and poverty calculations are based on individual-level consumption, which we obtain using the model estimates: for each woman, man, and child in our samples of analysis, we compute individual-level consumption as the product between their estimated resource shares and total household expenditure; we then compare individual consumption to age and gender-adjusted poverty lines to evaluate the incidence of poverty among children and adults.

4.1 Intra-household Inequality among Refugees

Figure 1 summarizes our results. Panel (A) shows the average resource share of each woman, man, and child among refugees in the three regions.[18] The share of consumption is largely balanced between men and women (with women commanding slightly more in Uganda and less in Kalobeyei).

[17] Importantly, there is sufficient variation in the assignable clothing budget shares to be able to estimate the Engel curves in equation (1). The percent of the sample reporting zero expenditure on women's, children's or men's clothing is comparable to previous studies (Dunbar et al., 2013; Calvi, 2020; Brown et al., 2021a; Penglase, 2020; Hoehn-Velasco and Penglase, 2021).

[18] Recall that resource shares are estimated conditional on a set of observable household covariates (including household composition and characteristics of the household head) to allow for heterogeneity in intra-household allocation. The empirical distributions of the estimated resource shares for men, women and children are provided in Figures A2 in the online Appendix. The coefficients for each covariate and the corresponding heteroskedasticity-robust standard errors are available upon request.

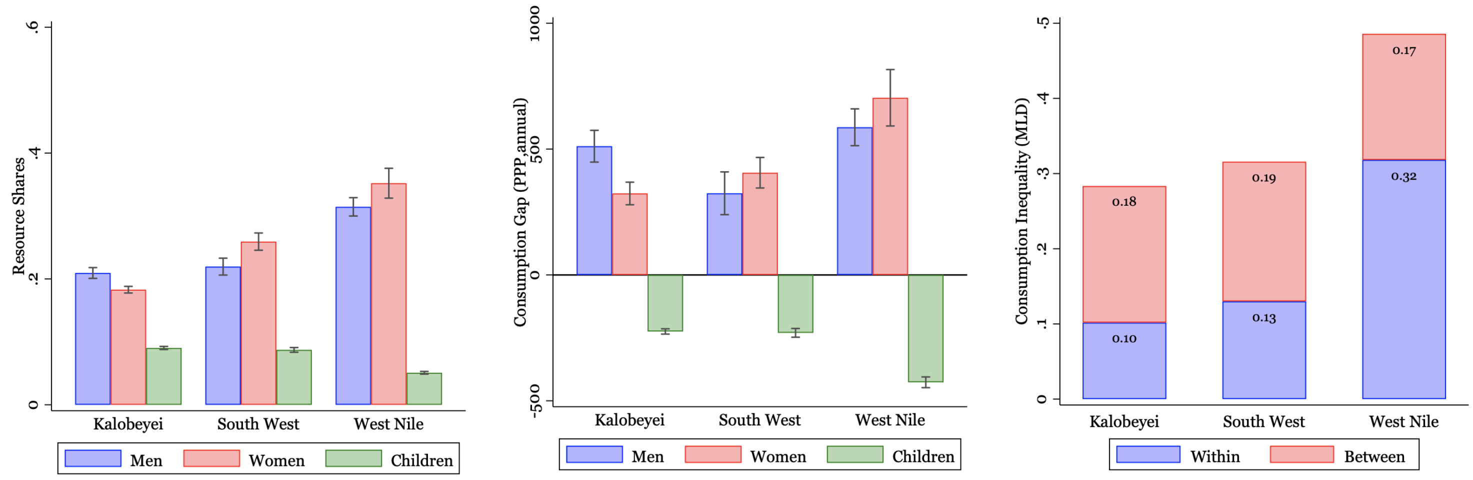
By contrast, children are allocated a strikingly smaller fraction of household expenses (9 percent each in Kalobeyei and South West Uganda, and 5 percent each in the West Nile settlements). The average age of children is lower in West Nile, which helps explain the differences across sites.

The difference between individual consumption (computed as the product of the estimated resource shares for men, women, and children, and their total household expenditure) and per-capita consumption provides further insight on the extent of intra-household inequality in our samples. We call this difference the *individual consumption gap*: if every family member was allocated the same share of consumption, the gap would equal zero for all; larger gaps indicate wider intra-household discrepancies in consumption, and reveal the potential mistargeting of anti-poverty programs that ignore intra-household inequality (an issue we investigate in more details in Section 5). In Panel (B) of Figure 1, we plot the average consumption gaps for women, men, and children in refugee families. In all refugee settlements but with varying intensity, adults' consumption is above the household per-capita consumption, while children's is substantially below. While part of the gap may be explained by differences in needs (with children requiring fewer resources than adults), in Section 5 we show that the shortfalls in children's consumption cannot be entirely explained by such differences.

To ease comparisons, Panels (A) and (B) of Figure 1 focus on refugee families with women, men, and children. These represent 64 percent of households in Kalobeyei, 49 percent in South West Uganda and and 72 percent in the West Nile region in Uganda. A sizeable share of households in these communities consists only of women and children - 32 percent in Kalobeyei, 46 percent in South West Uganda and 24 percent in West Nile Uganda. These figures reflect the extent to which a family structure's can be reorganized as a result of forced displacement. In these families, we estimate intra-household consumption inequality between women and children to be substantial in the Ugandan settlements, with women commanding twice as much consumption as children in the South West settlements and up to four times as much in the West Nile region. By contrast, in Kalobeyei, the average resource shares for women and children are almost identical, with each woman and child estimated being allocated approximately 20 percent of the total household budget (the full set of results is available upon request).

To better understand the scope and intensity of consumption inequality within and between households in refugee communities, we also calculate the mean log deviation for our estimates of

Figure 1: Intra-household Inequality among Refugees



(A) Resource Shares

(B) Individual Consumption Gap

(C) Within vs. Between Inequality

NOTES: Panel A shows the average estimated resource shares for women, men, and children with 95 percent confidence intervals. The consumption gap in Panel B is calculated as the difference between individual consumption (estimated) and per-capita household consumption. The mean log deviation (MLD) decomposition of consumption inequality between and within families is provided in Panel C. Only households with men, women and children are included in Panels A and B. Panel C includes both nuclear and extended families with and without children under 18 as well as single-parent families.

individual consumption. The mean log deviation (MLD) is a measure of income inequality, which takes on larger positive values as incomes become more unequal (Ravallion, 2016).[19] Unlike other measures of inequality such as the Gini index, the MLD can be decomposed into between and within-group components. As shown in Panel (C) of Figure 1, there is substantial heterogeneity in the overall level of inequality in the three regions. While the between-household inequality is similar across the three settlements, the total MLD is higher in the West Nile region relative to the camps in South West Uganda and the Kalobeyei settlement in Kenya. This gap is driven by the higher within household consumption inequality in the former region, which is two to three times as large as in South West Uganda and Kalobeyei and accounts for 65 percent of total inequality. A comparison between Panels (A) and (B) indicates that the relatively higher within-household inequality in the West Nile region can be attributed to the larger disparity between adults and children. That refugee children in this region are, on average, one year younger than in Kalobeyei or settlements in South West Uganda may partly explain this pattern.

[19] $MLD = \frac{1}{N}\sum_{i=1}^{N}(\ln \overline{y} - \ln y_i)$, where y_i where y_i is individual consumption, $\overline{y}$ is average consumption among all individuals, and N is the total number of individuals.

4.2 Comparison with Surrounding Host Communities

As we have shown above, intra-household consumption inequality is substantial among refugees in our sample, with children's consumption estimated to be significantly lower than per-capita household consumption. One question that takes center stage among policy makers is how refugee settlements fare relative to the surrounding host communities. In the context of our analysis, the question is: How similarly (or differently) do refugee families allocate consumption to their members relative to non-refugee families living in the surrounding areas?

A comparison between Figure 1 and Figure 2 helps answer this question. Three main observations stand out. First, in non-refugee families living in the vicinity of refugee settlements, each child is allocated the smallest share of consumption in all the three regions (25 percent to 30 percent of each adult consumption). Second, among hosts relatively more than refugees, each woman commands a larger share of the budget relative to men.[20] Third, for hosts, the overall consumption inequality (as measured by the MLD) is slighly higher in the South West region: while between household inequality is fairly similar in the non-refugee communities surrounding refugee settlements in the three regions, within-household inequality is significantly larger in South West Uganda. Here within-household disparities in consumption account for 57 percent of total inequality.

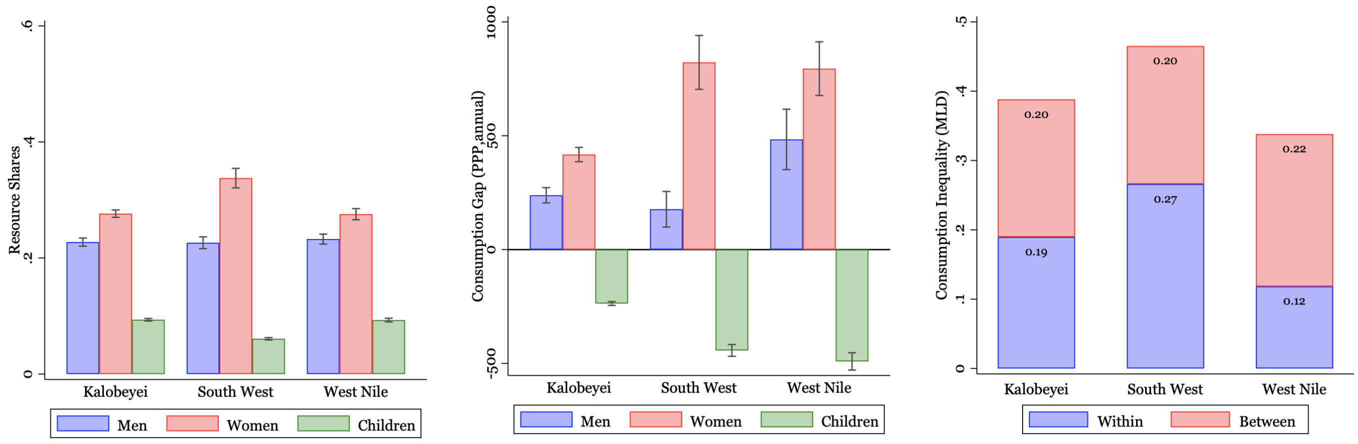
Taken together, our results indicate that intra-household consumption inequality is widespread in both communities and in all areas of study, with children facing a particularly high risk of poverty. Nevertheless, alongside these similarities, differences exist between refugee families and non-refugee families living in the areas around the camps. These differences may call for tailored policies to reduce inequality and poverty among refugee and hosts, and to reach the most vulnerable individuals in the two communities.

4.3 Individual-level Poverty

We now use the model estimates to construct poverty rates that take into account these intra-household disparities in consumption. These are different from standard poverty measures, which by construction assume an equal allocation of household consumption. We focus on the World

[20] This result is consistent with recent estimates of resource shares in the Kenyan context by Cherchye et al. (2021), who estimate resource shares with data from Nairobi and find that women, on average, are allocated 12 percent more than men. The empirical distributions of the estimated resource shares for men, women and children in host areas surrounding refugee camps are provided in Figures A3 in the online Appendix.

Figure 2: Intra-household Inequality in Surrounding Host Communities



(A) Resource Shares

(B) Individual Consumption Gap

(C) Within vs. Between Inequality

NOTES: Panel A shows the average estimated resource shares for women, men, and children with 95 percent confidence intervals. The consumption gap in Panel B is calculated as the difference between individual consumption (estimated) and per-capita household consumption. The mean log deviation (MLD) decomposition of consumption inequality between and within families is provided in Panel C. Only households with men, women and children are included in Panels A and B. Panel C includes both nuclear and extended families with and without children under 18 as well as single-parent families.

Bank's extreme poverty line of US$1.90 per day, which is meant to reflect the amount of resources below which a person's minimum nutritional, clothing, and shelter needs cannot be met. Using the same line for everyone, however, may lead to welfare-inconsistent poverty comparisons if some individuals (such as children) require fewer resources to achieve the same level of welfare as others (Brown et al., 2021b). To account for differences in needs across individuals, we create an equivalence scale based on relative caloric requirements by age and gender. Following Brown et al. (2021a), we assume US$1.90/day to be the average threshold for adults aged 15 to 45. We then scale individual poverty lines up or down based on the USDA Dietary Guidelines (2015-2020). Note that this adjustment relies on relative intakes rather than absolute caloric requirements, which mitigates concerns related to the applicability of the US dietary guidelines to our settings. The guidelines set that the average caloric recommendation for adults aged 15 to 45 to 2,400. So, e.g., the poverty line for an 8-year old girl with recommended intake of 1,600 calories/day would be US$1.27/day; for a 16-year old boy with recommended intake of 2,600 calories/day, it would instead be higher than US$1.90/day (US$2.06/day), and so on.

Table 1 shows head-count ratios (the fraction of individuals below the poverty line) and the daily dollar amount required to bring all children, women or men in our samples up to the poverty line. The latter measure, which is akin to a total poverty gap, reflects both the number of individuals in poverty (which may be higher in bigger settlements), the relative numbers of children, men and women, and the gap between their individual consumption and the poverty line. All poverty measures reported in Table 1 are computed using the model estimates of individual consumption in the refugee settlements of Kalobeyei, South West Uganda and the West Nile region, and in their respective surrounding host communities. While our preferred poverty calculations account for differences in needs by age and gender as discussed above (Panel A), we also report the unadjusted measures for comparison (Panel B). All samples include both nuclear and extended families with and without children under 18 as well as single-parent families.

In all three areas of study and for hosts and refugees alike, children face a substantially higher likelihood to live in poverty relative to adults. According to our child head-count ratio estimates, two out of three refugee children in South West Uganda and non-refugee children in the surrounding communities live in extreme poverty. Even in and around refugee settlements in the West Nile region, where poverty rates are lower overall, our estimated rates of child poverty are large. It is interesting to note that the difference in child and adult poverty rates are much narrower in Kalobeyei; here, the individual-level poverty rates are also close to the official national per-capita poverty rate of 36.1 (2015). As a comparison, in our samples per-capita head count ratios in refugee settlements (adjusted for relative needs) amount to 30.4 percent, 53.1 percent and 32 percent in Kalobeyei, South West Uganda and the West Nile region; in hours hosts samples, they equal 47, 31.3 and 13.8 percent respectively. As expected, our estimates of child poverty are higher when differences in needs by age and gender are ignored. Refugee women face higher poverty risks than non-refugee women. In both communities, however, they tend to face a lower poverty risk relative to men.

While informative, head-count ratios provide little insight on the intensity of individual-level poverty. Our poverty gap estimates help bridge this gap and unveil critical differences between refugee and non-refugee communities. First, child poverty is much more severe among refugees: for instance, we estimate it would take 732US$/day to eliminate child poverty among refugees in our Kalobeyei sample and 379$/day in the surrounding host community; in South West Ugandan samples, bringing each child up to their age and gender adjusted poverty line would require 1,783$/day and 377$/day among refugees and non-refugees, respectively. Second, these differences are not driven by a higher number of children in refugee camps (which could by construction inflate the total poverty gap), but by the combination of low incomes and intra-household inequality. Third, in our samples, the estimated poverty gaps for children are as much as eight times larger

Table 1: Individual Poverty in Refugee Settlements and Host Communities

	Refugees			Hosts		
	Kalobeyei	South West	West Nile	Kalobeyei	South West	West Nile
A) Poverty Line Adjusted for Relative Needs						
Head Count Ratio						
Children	0.391	0.687	0.584	0.540	0.661	0.269
Women	0.306	0.190	0.338	0.104	0.124	0.059
Men	0.366	0.534	0.492	0.329	0.224	0.178
Total Poverty Gap (US$/day)						
Children	731.651	1782.741	489.220	379.080	377.156	135.195
Women	131.174	115.397	59.153	23.659	19.211	8.692
Men	181.432	637.848	111.637	83.530	28.917	29.979
B) Poverty Line Not Adjusted for Relative Needs						
Head Count Ratio						
Children	0.624	0.844	0.733	0.679	0.742	0.438
Women	0.396	0.298	0.464	0.138	0.171	0.092
Men	0.308	0.514	0.454	0.288	0.192	0.119
Total Poverty Gap (US$/day)						
Children	1660.579	3524.797	989.119	748.156	737.363	335.915
Women	211.345	223.532	109.655	40.846	30.959	16.319
Men	132.537	525.681	92.825	61.531	21.475	17.361

NOTE: In Panel A, poverty lines are rescaled using the USDA Dietary Guidelines (2015-2020): we assume US$1.90/day to be the average poverty line for adults aged 15 to 45. We then rescale individual poverty lines based on relative caloric requirements by age and gender. In Panel B, the poverty line equals US$1.90/day for all. All samples include both nuclear and extended families with and without children under 18 as well as single-parent families.

than for adults (women and men combined) in refugee settlements and as much as three times larger in the surrounding host communities.

Even accounting for the estimation error embedded in our calculations, these figures are alarming and may call for more targeted interventions. To help guide these efforts, the next section assesses the extent of poverty mistargeting and identifies several predictors of child poverty.

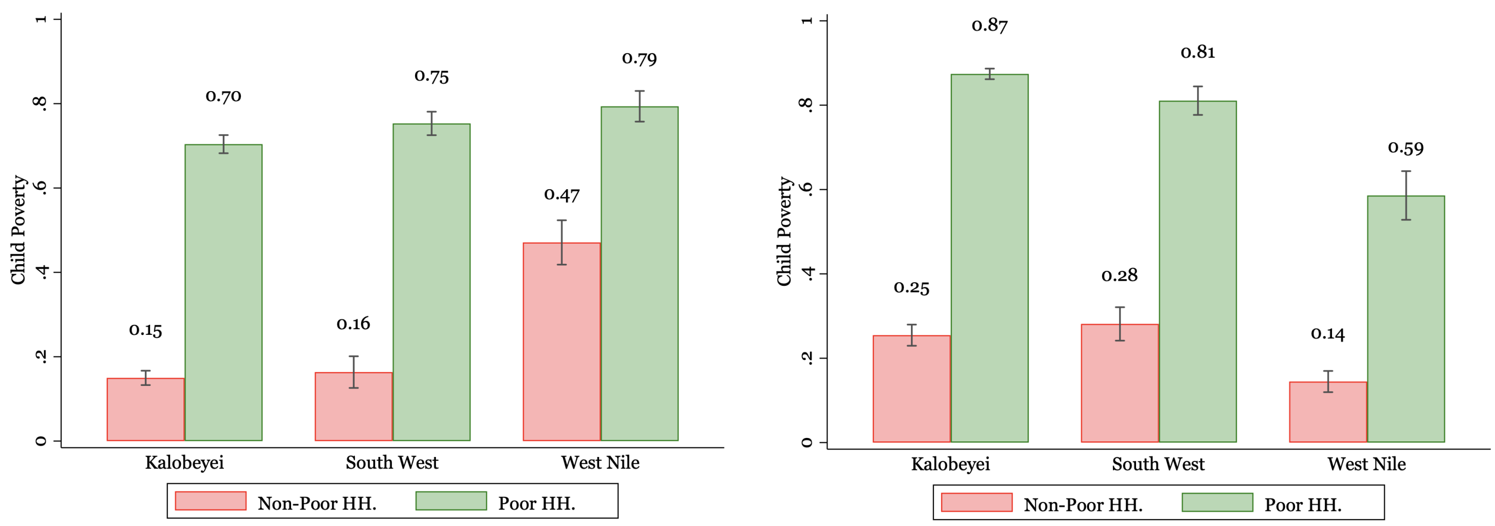
5 Targeting Child Poverty in and around Refugee Camps

5.1 Poor Children in Non-Poor Families?

One advantage of measuring consumption at the individual rather than the household level is the ability to identify poor people living in non-poor families and non-poor people living in poor families. In other words, poverty measures based on individual-level consumption (which account for intra-household consumption inequality) may not coincide with those based on per-capita household consumption. Such measures implicitly assume equal sharing and may lead to a misclassification of poor individuals as non-poor (or vice versa).

In Figure 3, we report the share of children living below their age and gender-adjusted poverty

Figure 3: Child Poverty by Household Poverty



(A) Refugees (B) Hosts

NOTES: The graph shows the shares of poor children living in poor and non-poor households. A poor child is defined as a child with estimated individual consumption below their gender and age adjusted poverty line. A household is defined poor if per-capita consumption is below the 1.90US$/day poverty line. All samples include nuclear and extended families with and without children under 18 as well as single-parent families.

line by their per-capita poverty status: we present the share of poor children living in non-poor families (with per-capita consumption above the poverty line) in red; the share of poor children in poor families is reported in green. Three observations deserve mention. First, and as we expected, the majority of poor children is found in poor households: in 59 to 87 percent of poor families, children also live below the poverty line. Note that these children would be correctly identified as poor by per-capita measures that ignore intra-household consumption inequality. Second, among refugees in Kalobeyei and host communities in the West Nile region, a substantial share of non-poor children (30 and 41 percent, respectively) live in families with per-capita consumption below the poverty line. Critically, many poor children reside in non-poor families. In the refugee settlements under study, up to almost half of poor children live in non-poor families. These children would not by reached by anti-poverty programs that ignore intra-household consumption inequality.

To better assess the accuracy of per-capita household consumption in predicting child poverty and to set the stage for our later analysis of poverty targeting, we follow Brown et al. (2018) and use our model estimates to compute standard measures of predictive performance. First, we compute the *inclusion error rate* associated with per-capita household consumption (which measures the likelihood of counting a child as poor based on per-capita household consumption when she is not based on her individual consumption). In our samples, it ranges between 0.21 and 0.34 in refugee settlements and between 0.13 and 0.43 in the surrounding communities.[21]

[21]The inclusion error rate is computed as the number of non-poor children in poor families, divided by the number of poor and non-poor

We also compute the *exclusion error rate* of per-capita household consumption (that is, the error associated with counting a child who is in fact poor as non-poor). This equals 0.15, 0.10, and 0.31 in refugee communities in Kalobeyei, South West Uganda and the West Nile region.[22] The exclusion error rate is alarming among non-refugees too and equal to 11, 26, and 42 percent in the three communities, respectively.

In summary, our calculations so far have shown that household-level poverty measures based on per-capita expenditure may be weak proxies for child poverty in and around refugee settlements in rural East Africa. The question is: Can poverty-targeting be improved to reach poor children in these settings? If so, can it be done in a parsimonious manner? Below, we provide a practical answer to this question that does not require collecting individual-level nor household-level consumption (which can be costly and time-consuming). Specifically, we apply a supervised machine learning algorithm to identify the most critical predictors of child poverty among refugees and hosts in our areas of study. Based on these predictors, we then develop *proxy-means tests* for child poverty to improve targeting accuracy at low cost. Proxy-means tests (or PMTs) are widespread approaches to targeting poverty under imperfect information (they have been widely used to approximate household consumption or income when such measures are not available or reliable). By combining easily observed measures (such as basic consumer durables or assets, demographic variables and attributes of the household head) into indexes, these approaches can help ameliorate poverty targeting (see e.g. Grosh and Baker (1995), Skoufias et al. (2001), and Brown et al. (2018)).[23] To our knowledge, however, they have not been used to improve poverty targeting at the individual level.

5.2 Proxy-Means Tests for Child Poverty

We now present the design and validation of low-cost, parsimonious targeting models for child poverty in our areas of study. Our approaches are *low-cost* because they rely on easily observable and verifiable household traits (hence avoiding the time and monetary cost of collecting detailed

children living in poor families (with per-capita expenditure below the poverty line). Low inclusion errors may help reduce the cost of social policies that use transfer payments to reduce poverty (Brown et al., 2018). High inclusion errors imply high costs without meaningful reductions in poverty.

[22] The exclusion error rate is computed as the number of poor children in non-poor families, divided by the number of poor children living in poor and non-poor families. It essentially measures the potential undercoverage of a program that targets poor children based on per-capita household consumption.

[23] Brown et al. (2018) assess strengths and weaknesses of standard econometric targeting methods (such as PMTs based on linear or quantile regression). The performance of these methods varies across implementations and contexts (Coady et al., 2004).

consumption expenditure data).[24] They are *parsimonious* because they limit the number of observable traits required to implement them (Jayachandran et al., 2021). Our analysis can guide the design of short-surveys to predict and target child poverty in vulnerable communities in and around refugee camps in rural East Africa. As discussed in Section 2, most anti-poverty programs in these contexts have been either universal or based on *ad hoc* considerations (e.g., the duration of stay in the camp). Importantly, they do not target poor children specifically.

Proxy-means tests or PMTs can be thought of as a weighted function of a vector of observable covariates (Brown et al., 2018). A popular approach consists of using regression coefficients as weights, but the exact set of covariates to be included in the regression model is often arbitrary, mostly driven by data availability rather than their predictive power.[25] As an alternative to standard regression-based methods, we apply a supervised machine learning algorithm (*random forest*; Breiman (2001); Genuer et al. (2010)) to select the most relevant predictors of child poverty. Based on these predictors, we then develop proxy-means tests for child poverty (with varying degrees of parsimoniousness and predictive power) to improve its targeting.[26] We also compare their performance to standard poverty targeting based on per-capita household expenditure.

We start with a comprehensive set of household and individual-level variables that are available in our three surveys. While there are a few minor differences across surveys, our analysis includes variables capturing household composition (household size and number of children in the household), characteristics of the child (such as age and gender) and of the household head (age, gender, employment status, and education level), dwelling characteristics, asset and animal ownership, and measures of food insecurity.[27] From this initial pool of observable traits, we select

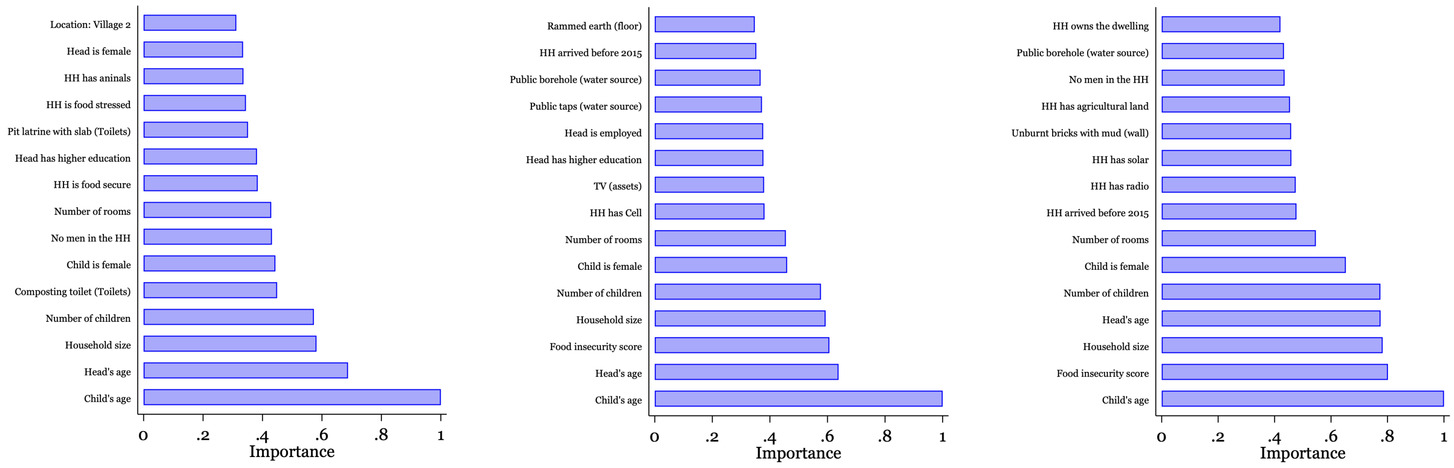
[24] The collection of consumption data can be demanding. Some of the challenges encountered by UNHCR enumerators in Kalobeyei included the fatigue on part of the respondent and the enumerator because of the length of the consumption module as well as the difficulty faced by respondents to remember expenses due to long recall periods in the nonfood items section. The enumerators also reported that some of the households didn't have a conventional consumption budget, which made it made it hard for respondents to tell how much was spent by other members who ate or purchased items outside of the home. Finally, due to lack of measuring scales made it difficult for both enumerators and respondents to give accurate measurements of food items.

[25] McBride and Nichols (2018) and Altındağ et al. (2021) also show that approaches based on in-sample validation, such as standard OLS, are likely to overfit in prediction exercises.

[26] Random forest is an ensemble method that builds decision trees to classify or fit the data (Breiman, 2001; Genuer et al., 2010). At each node of a tree, one of the variables is used to partition the data. Only a random subset of variables is used at each node, and the one that best partitions the data is selected. A random forest combines many trees. For each tree, some observations are left out and the predictions are validated against the testing sample. Relative to alternative variable selection methods like lasso, random forest has more flexibility to fit non-linear relationships in the data (Genuer et al., 2010; Jayachandran et al., 2021). It has been applied for predictions in a wide range of research fields and has also been used for poverty predictions (Altındağ et al., 2021; Niu et al., 2020; Sohnesen and Stender, 2017; Browne et al., 2021).

[27] Overall, we include 67 variables for both refugees and hosts in Uganda. For Kenya, 75 variables are included for both refugees and hosts. For Uganda, we use the Reduced Coping Strategies Index (rCSI) which is a proxy indicator of household food insecurity. It considers both the frequency and severity of five pre-selected coping strategies that the household used in the seven days prior to the survey. It is a simplified version of the full Coping Strategies Index indicator. For refugees in Kalobeyei, we include the Livelihood Coping Strategies Index (LCSI). The LCSI is an indicator to measure the extent of livelihood coping households need to utilize as a response to lack of food or money to purchase food. No measure of food insecurity is available for non-refugee communities in Kenya.

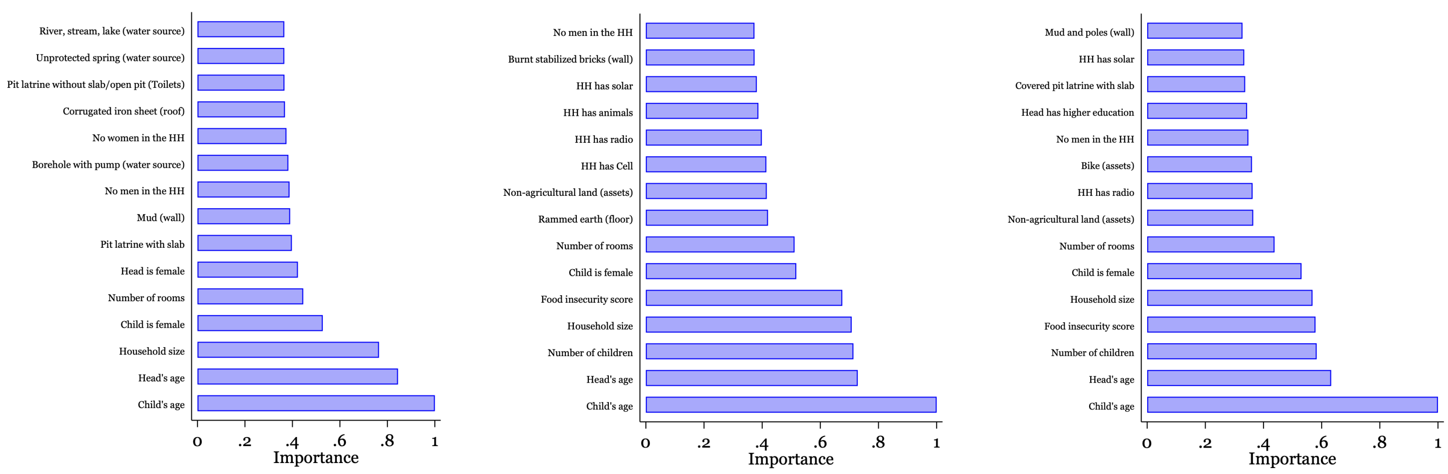
Figure 4: Top-15 Predictors of Child Poverty among Refugees



(A) Kalobeyei (B) South West (C) West Nile

NOTES: Top-15 predictors of child poverty based on their random forest importance score. The values are scaled proportional to the largest value in the set. See footnote 28 for details. All samples include nuclear and extended families with and without children under 18 as well as single-parent families.

Figure 5: Top-15 Predictors of Child Poverty among Hosts



(A) Kalobeyei (B) South West (C) West Nile

NOTES: Top-15 predictors of child poverty based on their random forest importance score. The values are scaled proportional to the largest value in the set. See footnote 28 for details. All samples include nuclear and extended families with and without children under 18 as well as single-parent families.

a limited number of explanatory variables based on their random forests importance score.[28]

Figures 4 and 5 present the top 15 predictors of child poverty among refugees and hosts in our three areas of study. For refugee households, eight out of the top predictors of child poverty are the same across all sites. These include the child's age and gender, the age and education of the household head, household size, the number of children living in the household, and the number of rooms in the household's dwelling. They also include a measure of food security, capturing both the frequency and severity of limited food access (see footnote 27 for details). Four out of

[28] The random forest importance measure of a given variable is the improvement in the split-criterion at each split in each tree, and is accumulated over all the trees in the forest separately for each variable. It is quite similar to the R^2 in regression on the training set for each variable taken separately. In the Stata package `rforest`, the variable importance score is normalized by dividing all scores over the maximum score: the importance of the most important variable is always 100 percent.

the eight common variables are among the top four predictors of child poverty in all sites (the child's age, the age of the household head, household size and the number of children living in the household). Beyond these variables, child poverty is highly predicted by a few additional observable household characteristics, including measures of wealth and assets (access to water sources and sanitation, the materials of walls and roofs in the dwelling, and ownership of animals or appliances).[29] Interestingly, there is substantial overlap between the top predictors of child poverty in refugee communities and the surrounding host areas. To further investigate the link between child poverty and its predictors in our areas of study, we use a linear probability model. Specifically, we regress child poverty on its 15 predictors identified by the random forest algorithm. Figures A4 and A5 in the Appendix plot the estimated coefficients with the associated confidence intervals.

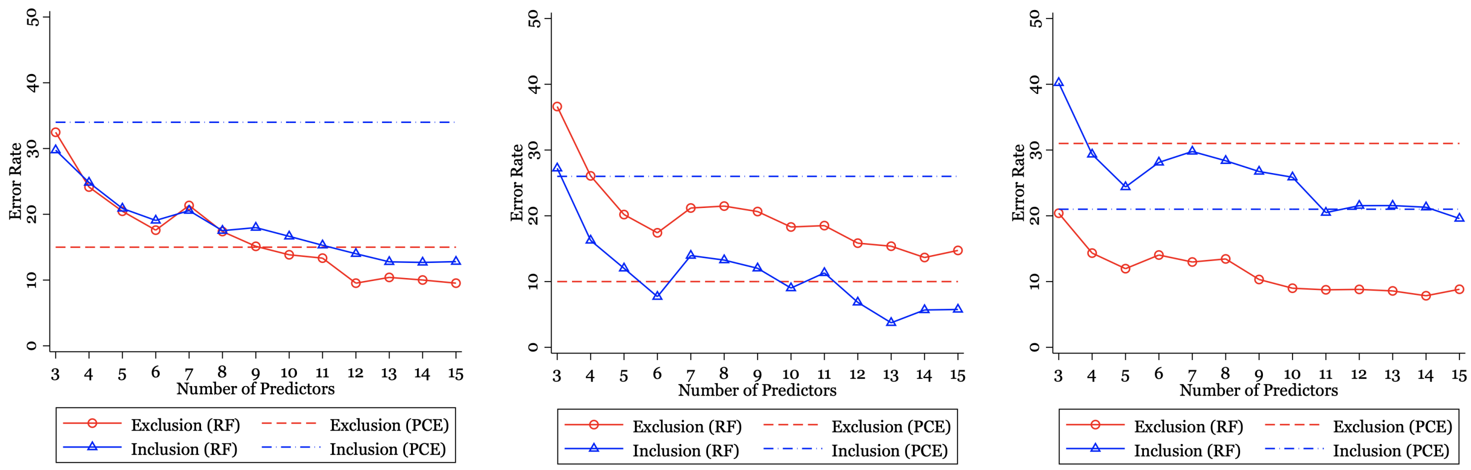
Predicting Child Poverty with Random Forest. We first predict child poverty using random forest classification models. We evaluate the predictive performance of different models with varying degrees of parsimoniousness (depending on the number of variables included for prediction). To ease comparison with our analysis in Section 5.1, we compute the associated inclusion and exclusion error rates. We recall that the inclusion error rate measures the likelihood of counting a child as poor based on the random forest prediction when she is not based on her individual consumption, while the exclusion error rate captures the error associated with counting a child who is in fact poor as non-poor. So, the former is a measure of leakage and the latter of undercoverage.

Figures 6 and 7 summarize the performance of the random forest classification algorithm in predicting child poverty in refugee camps and the surrounding communities in our areas of study. With a few exceptions (e.g., hosts around the Kalobeyei settlement) the random forest classification models (solid lines) out-perform per-capita household expenditure (dashed lines).[30] This is true even when the number of selected predictors is relatively low. In host communities in the West Nile region of Uganda, the random forest classification model based on just three variables achieves a much lower inclusion error rate than per-capita consumption. Similarly, random forest

[29] This finding echoes recent works stressing the important of WASH (i.e., water, sanitation and hygiene indicators). For instance, Brown et al. (2022) decompose the variation in nutritional outcomes between and within families in South Asia, finding that sanitation infrastructure and health facility quality are key correlates of nutritional outcomes. Recent work by Brown et al. (2020) also propose an index of the adequacy of home environments for protection (HEP) from COVID-19: access to water and sanitation are important components of their index. UNHCR has developed a comprehensive approach to WASH service provision in refugee settings around the world (UNHCR, 2020).

[30] Inclusion and exclusion error rates for per-capita household expenditure are computed as described in footnotes 21 and 22.

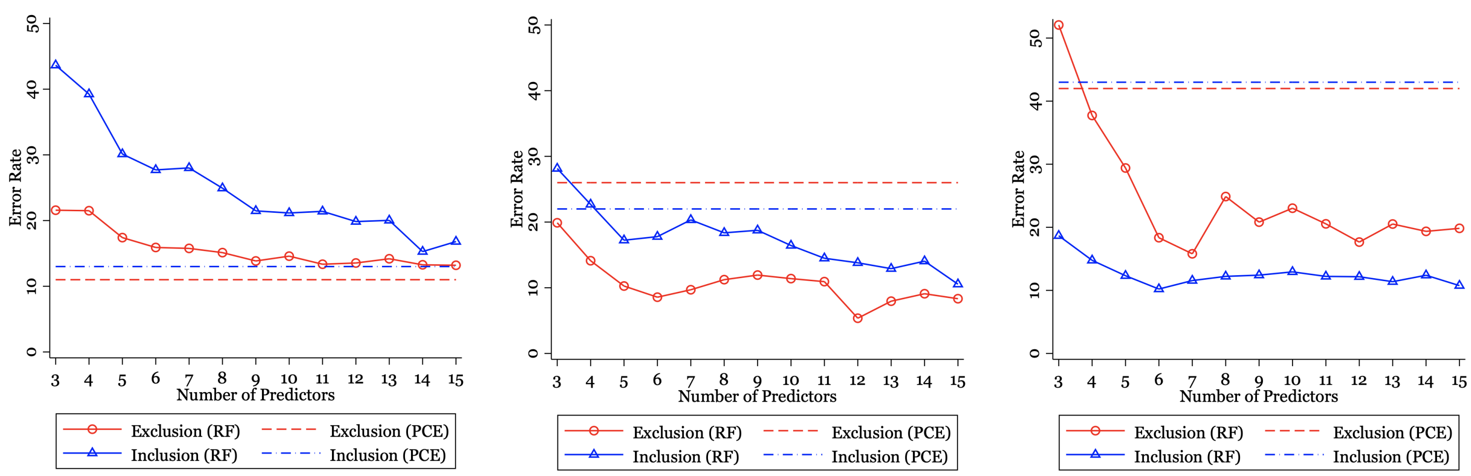
Figure 6: Random Forest Predictive Performance among Refugees



(A) Kalobeyei (B) South West (C) West Nile

NOTES: The graphs show the relative predictive performance of random forest classification models with different number of predictors (RF) and per-capita household expenditure (PCE). The variable to be predicted is whether a child is poor: a poor child is defined as a child with estimated individual consumption below their gender and age adjusted poverty line. Exclusion and Inclusion (RF) are the exclusion and exclusion error rates associated with random-forest classification models that use varying numbers of predictors. Exclusion and Inclusion (PCE) are the exclusion and exclusion error rates associated with per-capita household expenditure (see Section 5.1 and footnotes 21 and 22). All samples include nuclear and extended families with and without children under 18 as well as single-parent families.

Figure 7: Random Forest Predictive Performance among Hosts



(A) Kalobeyei (B) South West (C) West Nile

NOTES: The graphs show the relative predictive performance of random forest classification models with different number of predictors (RF) and per-capita household expenditure (PCE). The variable to be predicted is whether a child is poor: a poor child is defined as a child with estimated individual consumption below their gender and age adjusted poverty line. Exclusion and Inclusion (RF) are the exclusion and exclusion error rates associated with random-forest classification models that use varying numbers of predictors. Exclusion and Inclusion (PCE) are the exclusion and exclusion error rates associated with per-capita household expenditure (see Section 5.1 and footnotes 21 and 22). All samples include nuclear and extended families with and without children under 18 as well as single-parent families.

yields much lower exclusion error rates than standard classifications based on per-capita household expenditure.

When using random forest for prediction, the gains can be substantial. For instance, among refugees in Kalobeyei, the child poverty inclusion and exclusion error rates associated with the random forest classification model can be as much as 63 and 36 percent lower than those associated with per-capita household expenditure. In refugee settlements in South West Uganda, the random-forest inclusion error rates are 85 percent lower than the per-capita expenditure ones; in the West Nile region, the gains in predictive power associated with the random forest model instead of per-capita household consumption equal as much as 22 percentage points. In other words, the error associated with counting a child who is in fact poor as non-poor is 75 percent lower when using the random forest classification model rather than per-capital household expenditure. Turning to the host communities, we show that the targeting of child poverty can be substantially improved upon in South West Uganda and the West Nile regions. By contrast, per-capita household consumption can predict child poverty quite accurately in areas around the Kalobeyei settlement.

Predicting Child Poverty with Regression Models. As an alternative to poverty classification based on random-forest predictions, a much more typical and practitioner-friendly approach to poverty targeting under imperfect information involves computing proxy-means tests based on a regression model of (log) consumption on a vector of covariates (Brown et al., 2018; Altındağ et al., 2021):

$$\ln y_i = \beta_0 + \sum_{j=1}^{k} \beta_k x_{ik} + \epsilon_{ik}, \tag{2}$$

where y_i is consumption and x_{ik} are k independent variables. In our setting, y_i is child consumption (which we estimate in the previous section) and x_{ik} are the top child-poverty predictors selected by the random forest algorithm. We compute child PMT scores equal to $\exp(\hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_k x_{ik})$; we then compare each child PMT scores to her age and gender adjusted poverty line to predict their classification as poor or non-poor. In this case, the inclusion error rate measures the likelihood of counting a child as poor based on their PMT score when she is not based on her individual consumption, while the exclusion error rate captures the error associated with counting a child who is in fact poor as non-poor. The most common method for estimating β_0 and β_k is ordinary least square (OLS). Brown et al. (2018), however, have shown that in some instances it is preferable to calibrate the PMT score to how specific quantiles in the distribution of consumption change with the observable covariates. Below we follow these two approaches to evaluate the performance of various PMTs for child poverty. First, we estimate equation (2) with OLS; we then estimate quantile regressions with the quantile set equal to the child poverty rate in each area of study.

Tables 2 and 3 present the inclusion and exclusion error rates associated with basic PMTs based

Table 2: Linear and Quantile Regression Predictive Performance among Refugees

	Kalobeyei				**South West**				**West Nile**			
	OLS		Quantile		OLS		Quantile		OLS		Quantile	
	IER	EER	IER	EER	IER	EER	IER	EER	IER	EER	IER	EER
↓ *No. Predictors:*												
3	*0.30*	0.43	0.37	0.29	0.40	*0.07*	0.38	0.12	0.33	*0.01*	0.32	*0.05*
4	*0.31*	0.41	0.37	0.27	0.37	*0.08*	0.29	0.19	0.33	*0.03*	0.29	*0.08*
5	*0.29*	0.39	0.35	0.27	0.37	*0.08*	0.29	0.19	0.33	*0.03*	0.28	*0.11*
6	*0.29*	0.39	0.35	0.27	0.37	*0.08*	0.30	0.19	0.33	*0.02*	0.27	*0.10*
7	*0.29*	0.38	*0.34*	0.27	0.34	*0.09*	0.27	0.20	0.33	*0.03*	0.26	*0.10*
8	*0.29*	0.38	0.35	0.25	0.32	*0.07*	0.31	0.11	0.33	*0.02*	0.27	*0.08*
9	*0.28*	0.38	0.35	0.26	0.30	*0.08*	0.28	0.14	0.33	*0.02*	0.27	*0.07*
10	*0.28*	0.37	*0.34*	0.27	0.30	*0.09*	0.29	0.14	0.33	*0.03*	0.27	*0.08*
11	*0.28*	0.36	*0.33*	0.25	0.30	*0.09*	0.28	0.13	0.32	*0.04*	0.27	*0.09*
12	*0.28*	0.36	*0.33*	0.24	0.30	*0.09*	0.28	0.13	0.33	*0.04*	0.26	*0.11*
13	*0.27*	0.36	*0.33*	0.25	0.30	*0.09*	0.27	0.13	0.33	*0.04*	0.25	*0.14*
14	*0.27*	0.36	*0.33*	0.25	0.30	*0.09*	0.28	0.14	0.32	*0.04*	0.26	*0.13*
15	*0.27*	0.36	*0.31*	0.27	0.28	*0.09*	0.27	0.14	0.30	*0.02*	0.25	*0.09*
Per-capita exp.:	0.34	0.15	0.34	0.15	0.26	0.10	0.26	0.10	0.21	0.31	0.21	0.31
Random-forest (min):	0.13	0.10	0.13	0.10	0.04	0.14	0.04	0.14	0.20	0.08	0.20	0.08

NOTES: The table shows the relative predictive performance of PMTs based on linear regressions (OLS) and quantile regressions with varying numbers of predictors. The last two rows also report the predictive performance of random forest classification models with different number of predictors (minimum errors achieved) and per-capita household expenditure. The variable to be predicted is whether a child in refugee settlements is poor: a poor child is defined as a child with estimated individual consumption below their gender and age adjusted poverty line. EER and IER are the exclusion and exclusion error rates associated with the different prediction models. See Section 5.2 for OLS and quantile regressions and random forest classification models. See Sections 5.1 and footnotes 21 and 22 for per-capita expenditure.

on OLS estimation of a linear regression model (which follows the prevailing practice) and poverty-focused PMTs based on quantile regression (which use the child poverty rate as the quantile). We consider PMTs that include varying sets of child-poverty predictors based on their random-forest importance score (see Figures 6 and 7). To ease comparison with our previous analyses, the misclassification errors associated with per-capita household expenditure and the random forest classification model are included at the bottom of the tables. The PMT error rates are reported in italic when lower than those associated with per-capita household expenditure.

A few observations stand out. First, child poverty PMTs based on OLS or quantile estimation are able to achieve lower rates of inclusion and exclusion errors relative to per-capita household expenditure. This holds true both in refugee settlements and the surrounding host communities, and even when the number of predictors is limited. There is however substantial heterogeneity in the relative predictive performance of the various approaches across mistargeting errors, communities and areas of study. For instance, when compared with per-capita household consumption, a basic PMT for child poverty based on OLS estimates achieves the lowest leakage rate among refugees

Table 3: Linear and Quantile Regression Predictive Performance among Hosts

	Kalobeyei				**South West**				**West Nile**			
	OLS		Quantile		OLS		Quantile		OLS		Quantile	
	IER	EER	IER	EER	IER	EER	IER	EER	IER	EER	IER	EER
↓ *No. Predictors:*												
3	0.25	*0.06*	0.20	0.33	0.38	*0.05*	0.33	*0.11*	0.52	0.72	0.63	*0.16*
4	0.25	*0.06*	0.20	0.32	0.36	*0.06*	0.31	*0.10*	0.48	0.67	0.57	*0.16*
5	0.23	*0.08*	0.18	0.28	0.35	*0.08*	0.31	*0.12*	0.43	0.63	0.56	*0.15*
6	0.23	*0.08*	0.17	0.29	0.34	*0.07*	0.31	*0.12*	*0.43*	0.62	0.56	*0.17*
7	0.22	*0.09*	0.17	0.28	0.35	*0.09*	0.29	*0.12*	*0.43*	0.63	0.56	*0.17*
8	0.21	*0.08*	0.16	0.27	0.35	*0.09*	0.31	*0.12*	*0.32*	0.55	0.53	*0.16*
9	0.21	*0.08*	0.16	0.28	0.30	*0.07*	0.24	*0.09*	*0.33*	0.55	0.54	*0.17*
10	0.21	*0.08*	0.16	0.27	0.30	*0.07*	0.24	*0.09*	*0.37*	0.55	0.55	*0.14*
11	0.21	*0.09*	0.15	0.24	0.31	*0.09*	0.27	*0.10*	*0.37*	0.55	0.53	*0.15*
12	0.21	*0.09*	0.15	0.25	0.31	*0.10*	*0.25*	*0.11*	*0.37*	0.53	0.52	*0.15*
13	0.21	*0.08*	0.14	0.23	0.30	*0.10*	0.28	*0.10*	*0.38*	0.54	0.51	*0.15*
14	0.21	*0.08*	0.14	0.23	0.30	*0.09*	0.28	*0.09*	*0.37*	0.54	0.52	*0.17*
15	0.20	*0.09*	0.14	0.23	0.31	*0.10*	0.27	*0.09*	*0.40*	0.54	0.51	*0.16*
Per-capita exp.:	0.13	0.11	0.13	0.11	0.22	0.26	0.22	0.26	0.43	0.42	0.43	0.42
Random-forest (min):	0.15	0.13	0.15	0.13	0.11	0.05	0.11	0.05	0.10	0.16	0.10	0.16

NOTES: The table shows the relative predictive performance of PMTs based on linear regressions (OLS) and quantile regressions with different numbers of predictors. The last two rows also report the predictive performance of random forest classification models with different number of predictors (minimum errors achieved) and per-capita household expenditure. The variable to be predicted is whether a child in host communities around refugee settlements is poor: a poor child is defined as a child with estimated individual consumption below their gender and age adjusted poverty line. EER and IER are the exclusion and exclusion error rates associated with the different prediction models. See Section 5.2 for OLS and quantile regressions and random forest classification models. See Sections 5.1 and footnotes 21 and 22 for per-capita expenditure.

in Kalobeyei (with an associated inclusion error rate equal to 27 percent); in host communities, it yields the lowest risk of undercoverage (with an exclusion error rate as low as 6 percent). In South West Uganda, a PMT for child poverty based on OLS estimates outperforms per-capita household expenditure in terms of exclusion error rates in refugee camps, while PMT based on either OLS estimates or quantile regression yield lower exclusion error rates in the surrounding host communities. Turning to the West Nile region, parsimonious PMTs based on OLS estimates or quantile regression can improve upon the exclusion and inclusion error rates of per-capita household expenditure among refugee and host communities, with rates as low as 1 and 14 percent respectively (as opposed to 31 and 42 percent with per-capita expenditure). Notably but with a few exceptions, standard PMTs can also outperform the random-forest predictive classification model.

Taken together, our findings indicate that a small set of observable traits, such as a child's age, household composition, and access to sanitation and clean water, predict child poverty in refugee settlements and surrounding communities remarkably well, often better than per-capita household expenditure. While we identify such predictors using a supervised machine learning algorithm, we show that practitioner-friendly approaches based on standard econometric techniques (such as linear and quantile regressions) can help improve the targeting of child poverty in refugee camps and surrounding communities in our study. The improvement can be substantial. Collecting data on consumption (at the household level and even more so at the individual level) is costly, challenging, and time consuming. We hope our analysis can help the design of low-cost and parsimonious surveys to predict and target child poverty in vulnerable communities in and around refugee camps in rural East Africa.

6 Concluding Remarks

> "*[...] virtually every aspect of early human development, from the brain's evolving circuitry to the child's capacity for empathy, is affected by the environments and experiences that are encountered in a cumulative fashion, beginning in the prenatal period and extending throughout the early childhood years.*" (Shonkoff and Phillips, 2000)

We provide the first estimates of poverty among children in refugee camps and the surrounding hosts communities in rural East Africa. We find that intra-household consumption inequality is widespread in both communities and in all regions, with children facing a particularly high risk of poverty. Given that children are allocated a remarkably smaller fraction of the household budget, we find that a significant share of poor children reside in non-poor households. Without explicitly accounting for this high incidence of children who are poor in non-poor households, existing humanitarian and development programs targeting assistance to poor households will miss these children, making them especially vulnerable.

The first key takeaway from our analysis is that data exercises measuring poverty inside and around refugee camps in East Africa should be explicitly designed to allow for within household poverty calculations. For the first time, we are able to quantify the differences within household members and highlight the dire situation refugee and host children face. The second takeaway is that there is scope for improving the targeting of child poverty in these contexts. We provide several approaches (with various degrees of parsimony, accuracy and user-friendliness) to predict and target child poverty in the areas under study. While we wish to be cautious about the applicability of our findings to different settings, the fact that there is overlapping across samples in the set of observable household and child characteristics associated with children's likelihood to be poor is promising. Importantly, our overall approach (that combines structural estimation with supervised machine learning) could be applied to various settings, where low-cost improvements in poverty targeting are most needed.

Our machine learning analysis finds that the top predictors of child poverty among refugees are closely linked in all sites and include easily observable traits such as a child's age and gender, the age of the household head, household size and number of children living in the household. Other top predictors of child poverty include measures of household food insecurity, the head of household's education and employment status, the number of rooms in the dwelling, and access to water and sanitation. For UNHCR, who track refugee households using proGRes registration database, many of these household characteristics (gender, age, household composition and education of household head) are already collected. If UNHCR were to add a few survey fields, including basic data on household water and sanitation access, housing stock characteristics and collect employment status universally, our analysis shows that it could greatly improve upon current measures of poverty and inequality within and across households. It could also improve the accuracy of poverty targeting, especially among children in refugee settlements and the surrounding host communities, with potentially substantial gains for children's well-being and poverty alleviation both in the short and long run.

References

Almås, I., C. Ringdal, and I. Hoem Sjursen (2021): "Understanding inequality within households," Tech. rep., GLO Discussion Paper. [8]

Altindağ, O., S. D. O'Connell, A. Şaşmaz, Z. Balcioğlu, P. Cadoni, M. Jerneck, and A. K. Foong (2021): "Targeting humanitarian aid using administrative data: Model design and validation," *Journal of Development Economics*, 148, 102564. [20], [24]

Apps, P. F. and R. Rees (1988): "Taxation and the Household," *Journal of Public Economics*, 35, 355 – 369. [8]

Asiimwe, I. K. (2021): "The impact of emergency education on refugee settlement camps: a case study of Dzaipi Refugee Camp, Adjumani District," Ph.D. thesis, Makerere University. [6]

Atkinson, A. (2016): "Monitoring Global Poverty, Report of the Commission on Global Poverty. Washington, DC: World Bank Group," . [4]

Attanasio, O. P. and V. Lechene (2014): "Efficient Responses to Targeted Cash Transfers," *Journal of Political Economy*, 122, 178 – 222. [8]

Bargain, O., G. Lacroix, and L. Tiberti (2021): "Intrahousehold Resource Allocation and Individual Poverty: Assessing Collective Model Predictions against Direct Evidence on Sharing," Discussion Paper 14406, IZA. [2], [9]

Behrman, J. (1988): "Nutrition, Health, Birth Order and Seasonality," *Journal of Development Economics*, 28, 43–62. [1]

Behrman, J. and P. Tubman (1986): "Birth Order, Schooling, and Earnings," *Journal of Labor Economics*, 3, 5121–5145. [1]

Bicego, G., S. Rutstein, and K. Johnson (2003): "Dimensions of the Emerging Orphan Crisis in Sub-Saharan Africa," *Social Science and Medicine*, 56, 1235–1247. [1]

Black, S., P. Devereux, and K. Salvanes (2005): "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *The Quarterly Journal of Economics*, 120, 669–700. [1]

——— (2011): "Older and Wiser? Birth Order and IQ of Young Men," *CESifo Economic Studies*, 57, 103–120. [1]

BOOTH, A. AND H. J. KEE (2009): "Birth Order Matters: The Effect of Family Size and Birth Order on Educational Attainment," *Journal of Population Economics*, 22, 367–397. [1]

BREIMAN, L. (2001): "Random forests," *Machine learning*, 45, 5–32. [20]

BROWN, C., R. CALVI, AND J. PENGLASE (2021a): "Sharing the pie: An analysis of undernutrition and individual consumption in Bangladesh," *Journal of Public Economics*, 200, 104460. [1], [2], [4], [8], [11], [15], [36]

BROWN, C., R. CALVI, J. PENGLASE, AND D. TOMMASI (2021b): "Measuring Poverty Within The Household," *IZA Word of Labor*, Forthcoming. [2], [4], [8], [15]

BROWN, C., E. KANDPAL, J. LEE, AND A. WILLIAMS (2022): "Unequal Households or Communities?" . [22]

BROWN, C., M. RAVALLION, AND D. VAN DE WALLE (2018): "A Poor Means Test? Econometric Targeting in Africa," *Journal of Development Economics*, 134, 109–124. [18], [19], [20], [24]

——— (2019): "Most of Africa's Nutritionally-Deprived Women and Children are Not Found in Poor Households," *Review of Economics and Statistics*, 101, 631–644. [1]

BROWN, C. S., M. RAVALLION, AND D. VAN DE WALLE (2020): "Can the world's poor protect themselves from the new coronavirus?" *Unpublished Manuscript*. [22]

BROWNE, C., D. S. MATTESON, L. MCBRIDE, L. HU, Y. LIU, Y. SUN, J. WEN, AND C. B. BARRETT (2021): "Multivariate random forest prediction of poverty and malnutrition prevalence," *PloS one*, 16, e0255519. [20]

BROWNING, M., P.-A. CHIAPPORI, AND A. LEWBEL (2013): "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power," *The Review of Economic Studies*, 80, 1267–1303. [8], [36]

CALVI, R. (2020): "Why are older women missing in India? The age profile of bargaining power and poverty," *Journal of Political Economy*, 128, 2453–2501. [2], [11]

CALVI, R., J. PENGLASE, AND D. TOMMASI (2022): “Measuring Women’s Empowerment in Collective Households,” *AEA Papers and Proceedings*, Forthcoming. [8]

CALVI, R., J. PENGLASE, D. TOMMASI, AND A. WOLF (2019): “The More the Poorer? Resource Sharing and Scale Economies in Large Families,” *Unpublished Manuscript*. [36]

CASCO, J. L. (2022): “Intra-household Resource Shares under Poverty Transfers: Evidence from Ecuador,” *Available at SSRN 4104991*. [2]

CASE, A., C. PAXSON, AND J. ABLEIDINGER (2004): “Orphans in Africa: Parental Death, Poverty, and School Enrollment,” *Demography*, 41, 483–508. [1]

CHEN, M. AND J. DRÈZE (1992): “Widows and Health in Rural North India,” *Economic and Political Weekly*, WS81–WS92. [1]

CHERCHYE, L., P.-A. CHIAPPORI, B. DE ROCK, C. RINGDAL, AND F. VERMEULEN (2021): “Feed the children,” . [14]

CHIAPPORI, P.-A. (1988): “Rational Household Labor Supply,” *Econometrica*, 56, pp. 63–90. [2], [8]

——— (1992): “Collective Labor Supply and Welfare,” *Journal of Political Economy*, 100, pp. 437–467. [2], [8]

COADY, D., M. GROSH, AND J. HODDINOTT (2004): *Targeting Transfers in Developing Countries: Review of Lessons and Experience*, World Bank, Washington D.C. [19]

CUNHA, F. AND J. HECKMAN (2007): “The Technology of Skill Formation,” *American Economic Review*, 97, 31–47. [1], [4]

DE HAAN, M. (2010): “Birth Order, Family Size and Educational Attainment,” *Economics of Education Review*, 29, 576–588. [1]

DJUIKOM, M. AND D. VAN DE WALLE (2018): “Marital Shocks and Women’s Welfare in Africa,” World Bank Policy Research Paper 8306, World Bank, Washington DC. [1]

DRÈZE, J. AND P. SRINIVASAN (1997): “Widowhood and Poverty in Rural India: Some Inferences from Household Survey Data,” *Journal of Development Economics*, 54, 217–234. [1]

DUNBAR, G. R., A. LEWBEL, AND K. PENDAKUR (2013): "Children's Resources in Collective Households: Identification, Estimation, and an Application to Child Poverty in Malawi," *American Economic Review*, 103, 438–471. [2], [4], [8], [9], [11], [36]

EVANS, D. AND E. MIGUEL (2007): "Orphans and Schooling in Africa: A Longitudinal Analysis," *Demography*, 44, 35–57. [1]

FIX, J., U. PAPE, F. APPLER, T. BELTRAMO, F. NIMOH, L. R. RIVERA, F. SCHMIEDING, AND N. KARIUKI (2019): "Understanding the Socioeconomic Conditions of Refugees in Kenya: Volume A-Kalobeyei Settlement: Results from the 2018 Kalobeyei Socioeconomic Survey," *Washington, DC: The World Bank Group*. [10]

GENUER, R., J.-M. POGGI, AND C. TULEAU-MALOT (2010): "Variable selection using random forests," *Pattern recognition letters*, 31, 2225–2236. [20]

GRANTHAM-MCGREGOR SM, POWELL CA, W. S. H. J. (1991): "Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican Study," *Lancet*, 338. [1]

GROSH, M. E. AND J. L. BAKER (1995): *Proxy Means Tests for Targeting Social Programs: Simulations and Speculation*, The World Bank. [19]

HERNANDEZ-DE BENITO, M. (2022): "This is a Man's World: Crime and Women's Bargaining Power," *Available at SSRN 4057565*. [2]

HOEHN-VELASCO, L. AND J. PENGLASE (2021): "Changes in Assortative Matching and Educational Inequality: Evidence from Marriage and Birth Records in Mexico," *Available at SSRN 3844840*. [11]

JAYACHANDRAN, S., M. BIRADAVOLU, AND J. COOPER (2021): "Using machine learning and qualitative interviews to design a five-question women's agency index," Tech. rep., National Bureau of Economic Research. [20]

JAYACHANDRAN, S. AND I. KUZIEMKO (2011): "Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India," *The Quarterly Journal of Economics*, 126, 1485–1538. [1]

JAYACHANDRAN, S. AND R. PANDE (2017): "Why Are Indian Children So Short? The Role of Birth Order and Son Preference," *American Economic Review*, 107, 2600–2629. [1]

JENSEN, R. (2005): *Caste, Culture, and the Status and Well-Being of Widows in India*, University of Chicago Press, 357–376. [1]

KNBS (2018): "Kenya Integrated Household Budget Survey (KIHBS) 2015/16," . [10]

KNUDSEN, E. I., J. J. HECKMAN, J. L. CAMERON, AND J. P. SHONKOFF (2006): "Economic, neurobiological, and behavioral perspectives on building America's future workforce," *Proceedings of the National Academy of Sciences*, 103, 10155–10162. [1]

LANCASTER, G., P. MAITRA, AND R. RAY (2008): "Household Expenditure Patterns and Gender Bias: Evidence from Selected Indian States," *Oxford Development Studies*, 36, 133–157. [1]

LECHENE, V., K. PENDAKUR, AND A. WOLF (2020): "OLS estimation of the intra-household distribution of expenditure," *Journal of Political Economy*, Forthcoming. [2], [4]

LEWBEL, A. AND K. PENDAKUR (2008): "Estimation of Collective Household Models with Engel Curves," *Journal of Econometrics*, 147, 350–358. [36]

MARTORELL, R. (2017): "Improved nutrition in the first 1000 days and adult human capital and health," *American Journal of Human Biology*, 29, e22952. [1]

MCALOON, J. (2014): "Complex trauma: How abuse and neglect can have life-long effects," *The Conversation*. [6]

MCBRIDE, L. AND A. NICHOLS (2018): "Retooling poverty targeting using out-of-sample validation and machine learning," *The World Bank Economic Review*, 32, 531–550. [20]

MENON, M., K. PENDAKUR, AND F. PERALI (2012): "On the Expenditure-dependence of Children's Resource Shares," *Economics Letters*, 117, 739–742. [9]

NET, F. (2018): "Uganda Food Security Classification 2018," . [6]

NIU, T., Y. CHEN, AND Y. YUAN (2020): "Measuring urban poverty using multi-source data and a random forest algorithm: A case study in Guangzhou," *Sustainable Cities and Society*, 54, 102014. [20]

Oster, E. (2009): "Does Increased Access Increase Equality? Gender and Child Health Investments in India," *Journal of Development Economics*, 89, 62–76. [1]

Penglase, J. (2020): "Consumption Inequality Among Children: Evidence from Child Fostering in Malawi," *The Economic Journal*, 131, 1000–1025. [2], [11]

Price, J. (2008): "Parent-Child Quality TIme: Does Birth Order Matter?" *Journal of Human Resources*, 43, 240–265. [1]

Rangel, M. and D. Thomas (2019): "Decision-making in Complex Households," *Unpublished Manuscript*. [8]

Ravallion, M. (2016): *The Economics of Poverty: History, Measurement, and Policy*, New York: Oxford University Press. [13]

Shonkoff, J. and D. Phillips (2000): *From Neurons to Neighborhoods: The Science of Early Childhood Development*, Washington (DC): National Academies Press (US). [1], [4], [27]

Skoufias, E., B. Davis, and S. De La Vega (2001): "Targeting the poor in Mexico: an evaluation of the selection of households into PROGRESA," *World development*, 29, 1769–1784. [19]

Sohnesen, T. P. and N. Stender (2017): "Is random forest a superior methodology for predicting poverty? An empirical assessment," *Poverty & Public Policy*, 9, 118–133. [20]

Sokullu, S. and C. Valente (2021): "Individual Consumption in Collective Households: Identification Using Repeated Observations with an Application to PROGRESA," *Journal of Applied Econometrics*. [2]

Sozbir, O. F. (2022): "The Intra-household Effects of Refugee Inflows on Native Families," . [2]

Subramanian, S. and A. Deaton (1990): "Gender Effects In Indian Consumption Patterns," Papers 147, Princeton, Woodrow Wilson School - Development Studies. [1]

Tommasi, D. (2019): "Control of Resources, Bargaining Power and the Demand of Food: Evidence from PROGRESA," *Journal of Economic Behavior & Organization*, 161, 265–286. [2]

Udry, C. (1996): "Gender, Agricultural Production, and the Theory of the Household," *Journal of political Economy*, 1010–1046. [8]

UNHCR (2018): "Kalobeyei Integrated SocioEconomic Development Plan (KISEDP) in Turkana West," . [5], [6]

——— (2019): "Kenya - Standardized Expanded Nutrition Survey (SENS), Kakuma and Kalobeyei Refugee Camps - 2019," . [7]

——— (2020): "UNHCR WASH MANUAL: Programme Guidance," . [22]

——— (2021): "Refugee Data Finder," . [1], [2]

——— (2022a): "UNHCR Kenya Operational Update for April 2022," . [2], [5]

——— (2022b): "UNHCR Uganda Operational Update for March 2022," . [2]

UNHCR, WFP, and G. of Uganda (2019): "Fill the Nutrient Gap: Uganda Refugee Settlements Summary Report," . [6]

van de Walle, D. (2013): "Lasting Welfare Effects of Widowhood in Mali," *World Development*, 51, 1–19. [1]

World Bank (2019): "Informing the Refugee Policy Response in Uganda: Results from the Uganda Refugee and Host Communities 2018 Household Survey," . [6], [10]

Appendix (Not for Publication)

This Appendix contains three main sections. In the first one, we present the theoretical model in details. In the second one, we describe our datasets and sources. Additional figures and tables are included in the third section.

A Collective Households and Resource Sharing

We now set out a collective household model to identify and estimate resource sharing among co-resident family members. Our model builds upon the theoretical framework of Browning et al. (2013) and Dunbar et al. (2013). Our description follows closely Brown et al. (2021a).

Let households consist of J categories of *people* (indexed by j), such as children, men, and women. Denote the number of household members of category j by J, with M, W, and C being the number of men, women, and children in the family. All household members of a specific category are the same and are treated equally (see Section 3.1 for details). Let y denote the household's total expenditure. Each household consumes K types of goods with prices $p = (p^1, ..., p^K)$. Let $z = (z^1, ..., z^K)$ be the vector of observed quantities of goods purchased by each household and $q_j = (q_j^1, ..., q_j^K)$ be the vector of unobserved quantities of goods consumed by individuals of type j (their *private good equivalents*). Economies of scale in consumption are modeled through a Barten type consumption technology: there exists a $K \times K$ matrix A such that $z = A(Mq_m + Wq_w + Cq_c)$. If good k is a private good (i.e., not jointly consumed), then the kth row of A would be equal to 1 in the kth column and zeros elsewhere. If the good is shared, the A_{kk} matrix entry would be less than one, so that the sum of the private good equivalents may be weakly larger than what the household purchases.[31]

Each household member has a monotonically increasing, continuously twice differentiable and strictly quasi-concave utility function over consumption goods. Let $U_j(q_j)$ denote the consumption utility of individuals of type j over the vector of goods q_j.[32] The household chooses

[31] As in Dunbar et al. (2013), while the model allows for scale economies, these are not identified nor estimated. Doing so would require either detailed price variation and/or observability of consumption decisions of children living alone (Browning et al., 2013; Lewbel and Pendakur, 2008), or more demanding assumptions (Calvi et al., 2019).

[32] Each member may also care about other family members' well-being so that her total utility may depend on the utility of other household members. We assume that j's total utility is weakly separable over the consumption utility functions of all household members. So, for instance, member j would have a total utility function given by $\tilde{U}_j = \tilde{U}_j(U_1(q_1), \ldots, U_J(q_J))$. As $\tilde{U}_j$ depends upon $q_{j' \neq j}$ only through the consumption utilities they produce, direct consumption externalities are ruled out.

what to consume solving the following program:

$$\max_{q_1,\dots,q_J} U^H[U_1(q_1),\ \dots.\ ,\ U_J(q_J),\ p, y]$$
$$\text{such that} \tag{A1}$$
$$y = z'p \ \text{ and } \ z = A\sum_{j=1}^{J} Jq_j,$$

where the function U^H describes the social welfare function of the household. U^H exists because we assume that the household reaches a Pareto efficient allocation of goods. Because of this assumption, U^H can be represented as a weighted sum of the individual utilities. The dependency of such weights (known as the *Pareto weights*) on prices and income makes U^H a function of prices and income.

The solution of the problem above yields bundles of private good equivalents that each household member consumes. Pricing these vectors at shadow prices $A'p$ (which may differ from market prices because of the joint consumption of goods within the household) returns the fraction of the household's total resources that are devoted to each household member. These are their resource share η_j, which are our object of interest.

Leveraging the assumption of Pareto efficiency, the household program can be decomposed into two steps: the optimal allocation of resources across members and the individual maximization of their own utility function. Conditional on knowing η_j, household members choose q_j as the bundle maximizing their utility subject to a personal shadow budget constraint. By substituting the indirect utility functions $V_j(A'p, \eta_j y)$ in Equation (A1), the household program simplifies to the choice of optimal resource shares subject to the constraint that total resource shares must sum to one.

While the budget share functions for other goods are more complicated, the ones for private assignable goods are as follows:

$$X_j(y,p) = J\eta_j(y,p)\ x_j(\eta_j(y,p)y,\ A'p), \tag{A2}$$

where x_j is the individual budget share function of member j when facing their shadow budget constraint. With piglog preferences, x_j is linear in log-income (see equation (1) in Section 3.1).

B Data Compendium

Our analysis uses three data sources: i) the Kenya Integrated Household Budget Survey (2015/16 KIHBS), conducted in 2015/16 and covering only Kenyan nationals; ii) the Long Version of Kalobeyei Socio-Economic Assessment (2018/19 SEA-LV), conducted in 2018/19 and covering refugees in Kenya; iii) the Uganda Refugee and Host Communities 2018 Household Survey (2018 RHCS) covering both hosts and refugees in Uganda.

The 2015/16 KIHBS and 2018/19 SEA-LV are combined to study refugees and hosts living in and around Kalobeyei. The 2015/16 KIHBS was designed to be representative for each county, resulting in a national sample of 24,000 households for 21,800 households interviewed during the survey. The sampling for the survey was done in three stages. To measure consumption, the 2015/16 KIHBS collected information on food consumption, including purchases, production, stock and gift over a 7-day period recall. The values of the own output, stock and gift were obtained by using local unit prices. Overall, the 2015/16 KIHBS 2015/16 collected more than 445,300 observations of 217 different food items consumed by about 21,800 households. The 2015/16 KIHBS also collects data on non-food consumption with a recall period varying from one month to one year depending on the importance and frequency of the non-food item. The 2015/16 KIHBS represents the most comprehensive and detailed household budget survey ever collected in Kenya. The analysis use only the subsample of host households living around the camp.

The Long Version of Kalobeyei Socio-Economic Assessment (2019 SEA-LV) is a sample of 1,100 refugee households living in the Kalobeyei camp. Randomly selected households were administered a socio-economic questionnaire based on the World Bank-supported national Kenya Continuous Household Survey (KCHS) and key indicators from the 2016 Kakuma Refugee Vulnerability Study and other sources. These choices make the 2019 SEA-LV comparable to 2015/16 KIHBS and allow us to jointly analyze nationals and refugees. To measure consumption, the 2019 SEA-LV employs a new approach called the "Rapid Consumption Methodology" (RCM) that consists of five steps. First, core items are selected based on their importance for welfare and consumption. Second, the remaining consumption items are partitioned into different optional consumption modules. Third, optional modules are assigned to groups of households. Fourth, after data collection, consumption of optional modules is imputed for all households. Finally, the resulting consumption aggregate is used to estimate poverty.

For both refugees and hosts in the South West and West Nile regions of Uganda, we use cross-sectional household data from the Uganda Refugee and Host Communities 2018 Household Survey (2018 RHCS), which sampled 2,209 residential households, distributed geographically across thirteen districts in the primary refugee-hosting regions in Uganda. As a result, the survey is representative of the refugee and host community populations of Uganda at the national level, as well as in the regions of West Nile, the South West, and the city of Kampala. For West Nile, it includes the districts of Adjumani, Arua, Moyo, Yumbe, Koboko and Lamwo; for Southwest, it includes Hoima, Kamwenge, Isingiro, Kiryandongo and Kyegegwa. To measure consumption, the 2018 RHCS collected information on food consumption, including purchases, production, stock and gift over a 7-day period recall. The 2018 RHCS also gathers data on nonfood consumption with a recall period varying from one month to one year depending on the importance and frequency of the nonfood item. In addition, the data covers other themes such as demographic variables (household size, gender, age, country of origin, and region of residence), education, labor force, food security, access to services, and asset and animal ownership, and variables reflecting coping strategies. The data was collected from May 2018 to July 2018.

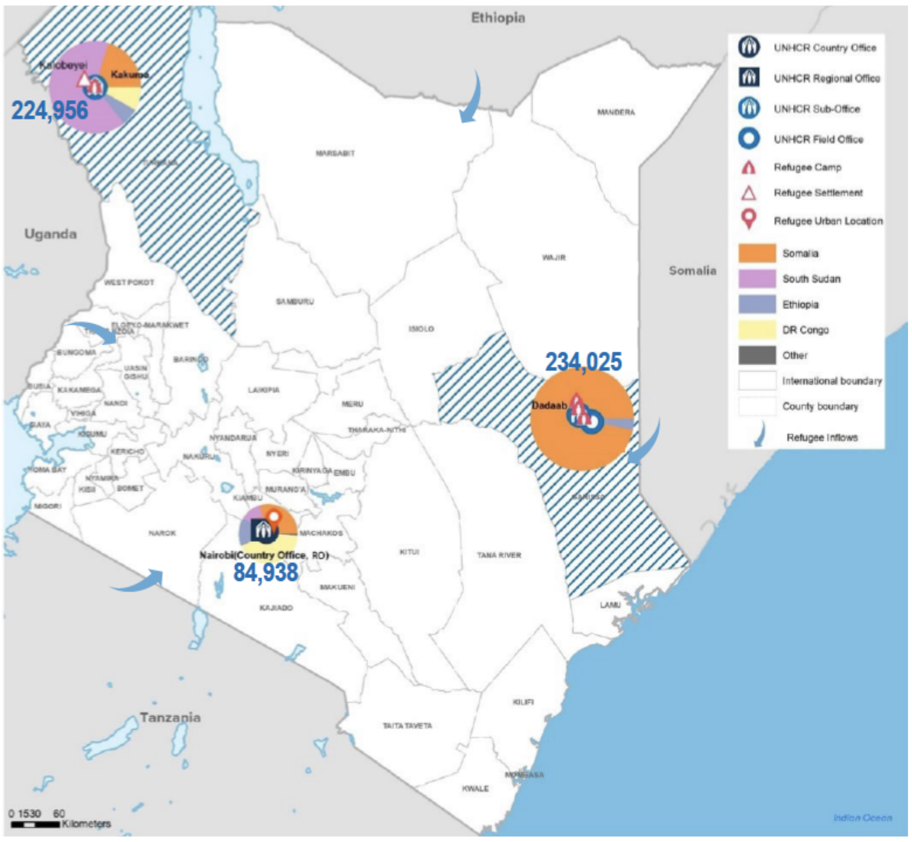
Table A1: Descriptive Statistics

	Mean			Differences between refugees and hosts		
	Kalobeyei	South West	West Nile	Kalobeyei	South West	West Nile
	(1)	(2)	(3)	(4)	(5)	(6)
Household Composition:						
Household Size	6.320	5.686	5.254	-0.818***	-0.490**	0.194
Men (share)	0.132	0.115	0.177	0.067***	0.095***	0.040**
Women (share)	0.199	0.231	0.253	0.043***	0.041***	-0.003
Children (share)	0.668	0.654	0.570	-0.110***	-0.137***	-0.037*
Men, Women, and Children	0.564	0.427	0.653	0.123***	0.232***	0.094*
No Men	0.384	0.509	0.270	-0.157***	-0.277***	-0.119***
No Women	0.052	0.037	0.016	-0.020*	-0.012	0.011
No Children	0.000	0.027	0.060	0.054***	0.057**	0.014
Female Head of Household	0.723	0.625	0.351	-0.358***	-0.314***	-0.179***
Household Expenditure:						
Total Expenditure (US$ PPP)	5,519.403	3,212.946	4,094.597	-1736.282***	1,333.802***	3,035.127***
Per-capita Expenditure (US$ PPP)	973.095	588.206	893.283	-204.128***	369.460***	579.215***
Men's Assignable Clothing Budget Share	0.002	0.003	0.010	0.006***	0.005***	0.000
Women's Assignable Clothing Budget Share	0.001	0.004	0.009	0.008***	0.005***	0.001
Children's Assignable Clothing Budget Share	0.002	0.008	0.013	0.011***	0.003***	-0.001

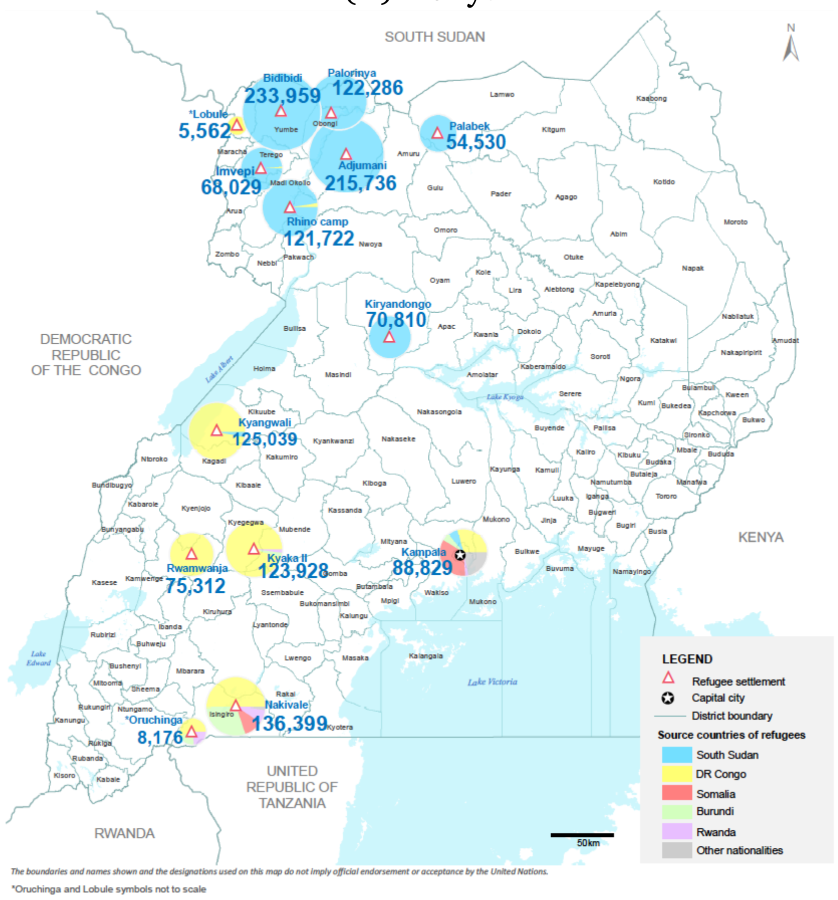
NOTES: Data for Kalobeyei (refugees) are from the Long Version of Kalobeyei Socio-Economic Assessment, for Kalobeyei (hosts) from the Kenya Integrated Household Budget Survey, for Uganda (West Nile and South West, both refugees and hosts) from the Uganda Refugee and Host Communities 2018 Household Survey. Statistical significance for t-test for differences in means (refugees minus hosts) are reported in Columns (4) to (5): * denotes statistical significance at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level.

C Additional Figures and Tables

Figure A1: Refugee Settlements in Kenya and Uganda



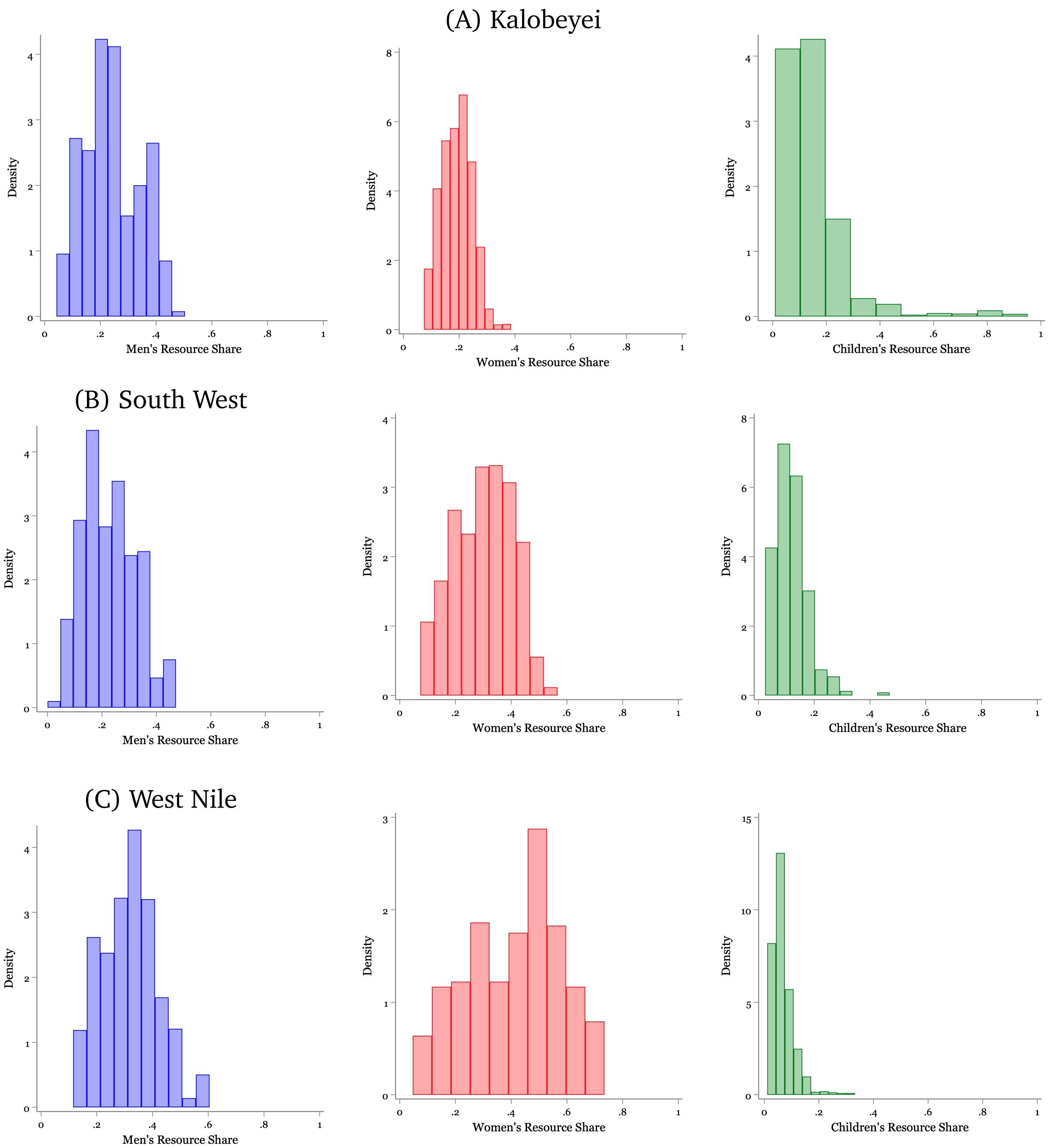
(A) Kenya



(B) Uganda

SOURCE: UNHCR.

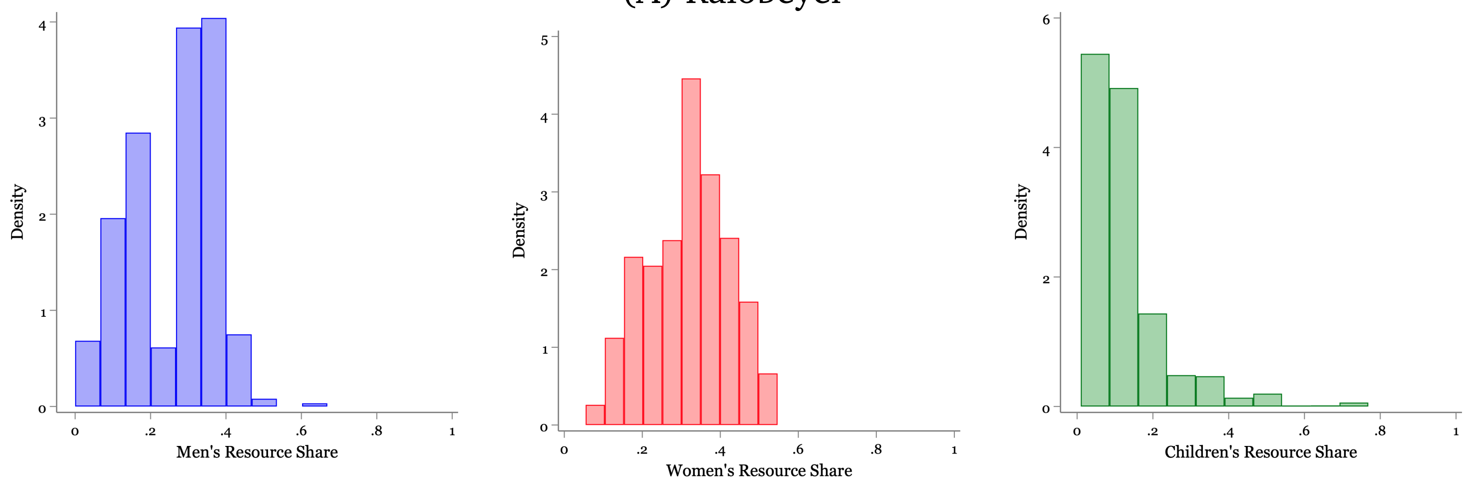
Figure A2: Estimated Resource Shares in Refugee Settlements



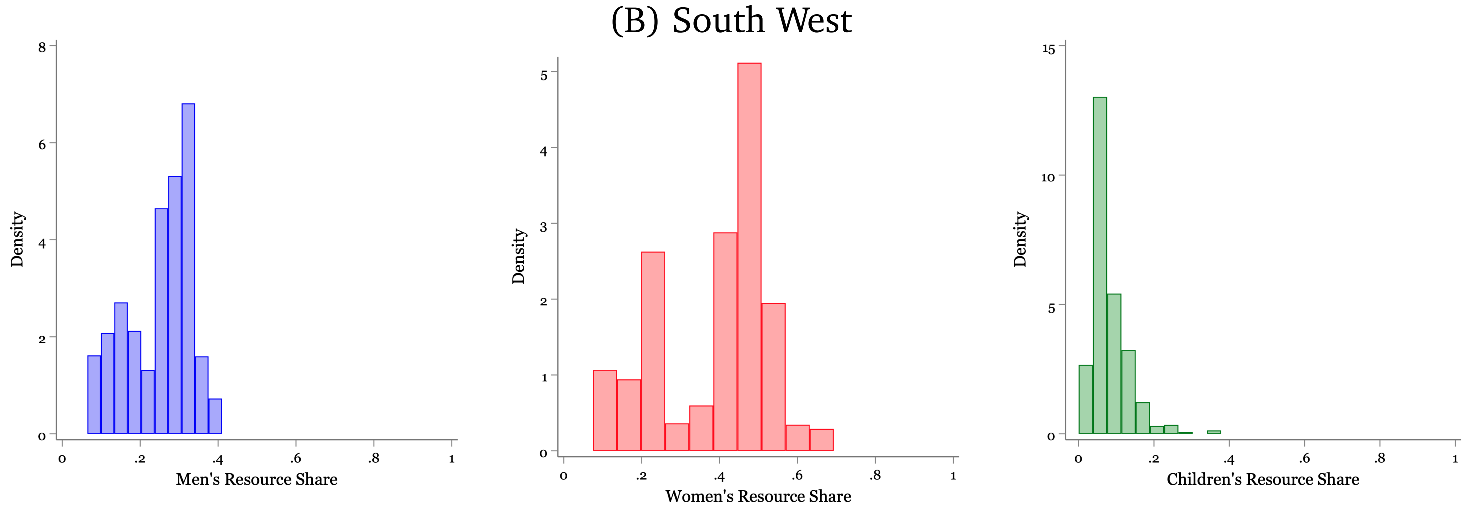
NOTES: The graphs plot the empirical distribution of the estimated resource shares for men, women and children in refugee settlements. All samples include both nuclear and extended families with and without children under 18 as well as single-parent families.

Figure A3: Estimated Resource Shares in Surrounding Host Communities

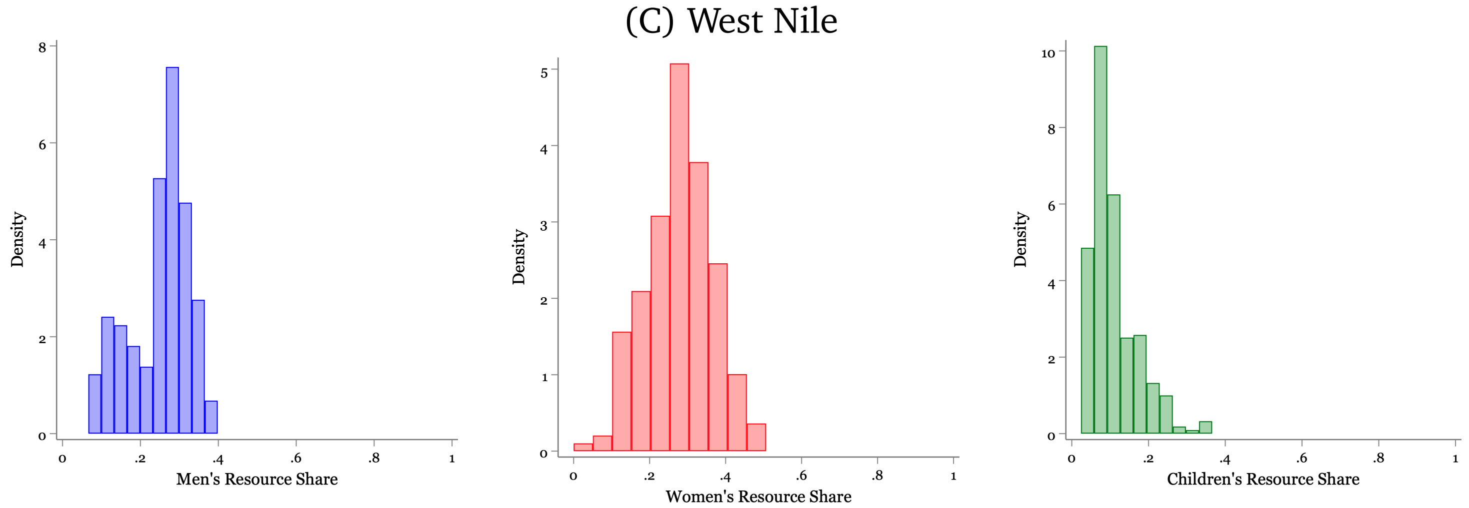
(A) Kalobeyei



(B) South West

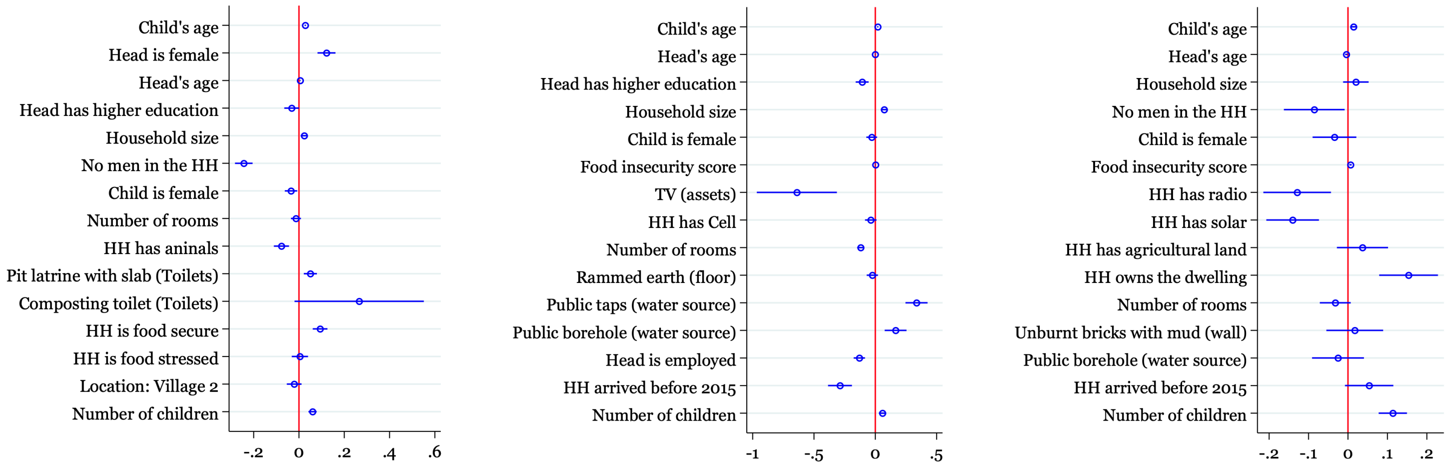


(C) West Nile



NOTES: The graphs plot the empirical distribution of the estimated resource shares for men, women and children in host communities surrounding refugee settlements. All samples include both nuclear and extended families with and without children under 18 as well as single-parent families.

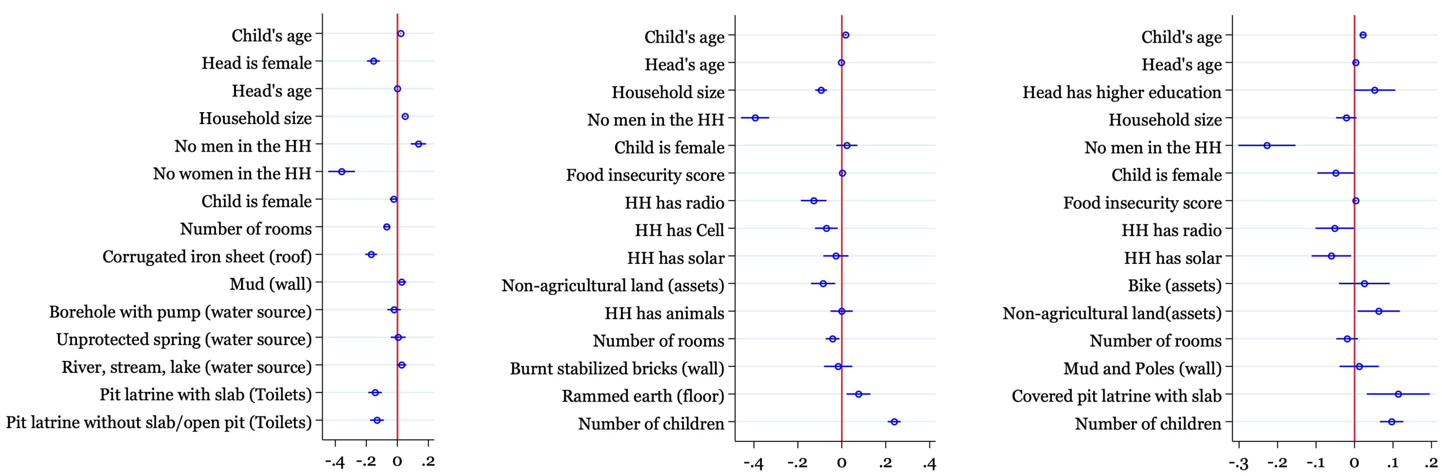
Figure A4: Predictors of Child Poverty in Refugee Settlements



(A) Kalobeyei (B) South West (C) West Nile

NOTES: The graphs plot the estimated coefficients and the associated 95 percent confidence intervals from a linear probability model of child poverty on its top-15 predictors (based on their random-forest importance score).

Figure A5: Predictors of Child Poverty in Surrounding Host Communities



(A) Kalobeyei (B) South West (C) West Nile

NOTES: The graphs plot the estimated coefficients and the associated 95 percent confidence intervals from a linear probability model of child poverty on its top-15 predictors (based on their random-forest importance score).