Hungry hosts? Refugee camps and host community nutritional outcomes in sub-Saharan Africa

Sebastian Anti[1] and Colette Salemi[2]

ABSTRACT: We examine nutritional outcomes in host communities exposed to refugee camps within a multi-country difference-in-differences framework across sub-Saharan Africa (SSA). Our study uses new spatially explicit data on refugee camps in SSA merged with the Demographic and Health Surveys to estimate the cross-country average treatment effect of camps. To test against bias in the coefficients of interest under staggered treatment timing, we use a diagnostic test to evaluate treatment effect homogeneity. We find that being within 10 kilometers of a camp decreases children's weight-for-age z-scores by 10 percent of the sample mean. Children with married household heads experience improved nutrition outcomes near camps. We consider adult loss of employment and worsening child health as explanatory mechanisms: we find no significant evidence of worsening child health or a reduction in employment opportunities. We argue that rising child malnutrition among hosts is due to the changing composition of the host population or to price shocks under localized price dispersal.

Keywords: Refugee impacts, Regional migration, nutrition and food security

JEL Codes: O12, O15, R23

[1] Assistant Professor, Department of Economics, Bryn Mawr College, 204 Dalton, Bryn Mawr, PA 19010

[2] Corresponding author. PhD Student, Department of Applied Economics, University of Minnesota, 213 Ruttan Hall, St. Paul, MN 55108. salem043@umn.edu

1. Introduction

Although the migration of refugees to Europe and the United States has dominated the headlines in recent years, most refugees in the world obtain asylum in developing countries that share a border with the refugee's nation of origin. Such host countries may require that a significant share of the refugee population lives in camps, which are usually located in peripheral and under-developed areas of the country (Alix-Garcia *et al.*, 2018). The reception of a sudden influx of people, and the accompanying shock associated with camp construction and increased humanitarian investment into services and infrastructure, have many potential effects on "host" communities located close to the camps. But our knowledge on how these shocks impact hosts remains limited. Although many countries in sub-Saharan Africa (SSA) have experienced refugee population inflows over the past half century, research quantifying the effects of such inflows has remained limited to a collection of case studies in Tanzania, Kenya, and Uganda, leaving the issue unexplored on a wider geographic scale.

The objective of this paper is to evaluate changes in host community nutritional status in response to camp openings across multiple countries in SSA. By doing so, we derive estimates with stronger external validity for the sub-continent than previously established. We accomplish this by combining 64 country-years of Demographic and Health (DHS) surveys. Based on DHS cluster geo-coordinates, we determine a household's Euclidean (linear) distance to a camp based on a new, spatially explicit dataset on refugee camp locations and years of operation. We use a difference-in-difference approach that exploits variation in proximity to a camp and variation in the years a camp was operating. We also invoke qualitative evidence collected from key informant interviews in East Africa when these accounts provide additional context and clarity.[3]

There are numerous factors to consider when evaluating how a sudden population shock of refugees, as well as camp construction and management, impact hosts. First, camp residents influence the local supply and demand of food. Refugees often sell some of their in-kind aid (Alloush et al., 2017; Callamard, 1994; Montclos and Kagwanja, 2000; Oka, 2011, 2014) and additionally use a combination of savings, earnings, and cash-based assistance to make purchases in local markets (Betts et al., 2017). In doing so, refugee camp populations influence both local supply and demand.

Second, displacement and camp management may indirectly influence host nutritional status by stimulating changes in the local labor market, though the direction of this impact remains ambiguous. When refugees can work, they increase the labor supply, and this may drive host employment and wages down for certain host groups (Maystadt and Verwimp, 2014; Ruiz and Vargas-Silva, 2016, 2015). Conversely, camp management generates economic activity. Humanitarian actors may purchase in-kind assistance from local markets to reduce transportation costs. If the production of relevant goods and services is labor-intensive, we would expect positive gains in hosts' employment and wages as a result. Furthermore, the construction, maintenance, and service provision to the camp, and a camp's contingency of aid workers that relocate to the area, may draw on local labor, resulting in an increase in employment and wages

[3] One of the authors spent the summer of 2019 speaking with government and UN planners involved in refugee assistance in Kenya, Rwanda, Uganda, and Tanzania. For a more detailed account of this qualitative research, see Appendix Section A4.

for those in nearby areas (Alix-Garcia *et al.*, 2018).[4] These factors could result in higher incomes, increased access to food, and improved nutrition for local households.

Finally, in many countries, refugee camp services have become increasingly "inclusive" over time, meaning they are shared with hosts. Living near a refugee camp could therefore mean improved access to potable water, healthcare, schooling, among other benefits. Key informants involved in refugee camp management in East Africa have stated that inclusive services do not include food aid, meaning there may be no direct impact of camp-related humanitarian investment on host nutrition. But increased access to such services could impact household nutrition indirectly by augmenting the time available for productive activities. These services, combined with the potential for increased labor market demand, may also attract certain households from the refugee-hosting country to move towards camps.

Empirical analyses of the effects of refugees on host nutrition have been limited to case studies focused on small regions within countries, or even the immediate locality around one refugee camp. Beyond Chambers' (1986) initial observations of negative price shocks in areas near refugee camps in Africa, there are only a small selection of studies from the economics literature that examine the effects of camps on host nutrition. Most of these studies examine changes in prices and consumption. First, Alix-Garcia and Saah (2010) evaluate food prices and household wealth for hosts in northwest Tanzania. They find that proximity to refugee camps results in large increases in the prices of non-aid food items and a measurable but less dramatic effect on food aid items such as maize and legumes, meaning that the price of certain staple goods did not change substantially for hosts. Next, Alix-Garcia *et al.* (2018) exploit nighttime lights remote sensing data to study how Kakuma refugee camp in northwest Kenya impacted the

[4] The net effect on employment is the subject of a paper we are writing concurrently with this piece. We explore employment in a limited way here to inform our results, but a full exploration is located in this second paper.

economic activity of its immediate locality. Their findings suggest that household consumption is 25 percent higher near the camp than in areas further away. Additionally, Maystadt and Verwimp, (2014) examine changes in consumption per adult equivalent (PAE) in Kagera, Tanzania. They find that, on average, PAE rises among Tanzanian hosts. But they also show that the distribution of benefits from hosting are unequally shared, with agricultural workers seeing relatively smaller increases than other groups and nonagricultural self-employed workers experiencing reductions in their PAEs. Moreover, Kreibaum (2016) looks at the effects of refugee settlements in southwest Uganda. She finds statistically weak evidence of increased host consumption in response to refugee settlement proximity.

To our knowledge, efforts to directly examine the nutritional impacts of refugee camp proximity have been limited. Baez (2011) studies the impacts of refugee camps on the health and nutrition outcomes of host communities in northwestern Tanzania. He shows a decline in children's and adults' anthropometrics, an increase in the prevalence of infectious diseases, and higher childhood mortality.

We contribute to two strands of the literature on the economics of agriculture and food. First, our work corresponds with the discourse on migration and household welfare in developing countries. Past studies have examined how low levels of well-being can stimulate out-migration (Dillon, Mueller and Salau, 2011; Lee, 2017) and how migration impacts welfare in the migrant-receiving area: our study adds to the literature on the latter. Migration, whether seasonal or permanent, can lead to improved consumption for the migrant-sending household (De Brauw and Harigaya, 2007; Zezza *et al.*, 2011; Bryan, Chowdhury and Mobarak, 2014; Chen, Kosec and Mueller, 2019), but it may also constitute an exogenous shock for the migrant-receiving area (German, Gomez Canon and Mueller, 2019). Likewise, our work considers the spillover effects for hosts living in the areas where migrants settle. Given the involuntary nature of the migration episodes we examine, our work aligns with past studies of migration stimulated by climate shocks (Chen and Mueller, 2018; Mueller, Gray and Hopping, 2020).

Second, our study adds to the discourse on nutrition in low- and middle-income countries and how nutritional status responds to an exogenous shock. Previous studies have considered how local and global price changes (D'Souza *et al.*, 2014; Arndt *et al.*, 2016), violent conflict (Akresh and de Walque, 2008; Akresh, Verwimp and Bundervoet, 2011; Minoiu and Shemyakina, 2014), poor harvests (Akresh, Verwimp and Bundervoet, 2011), and humanitarian aid (Quisumbing, 2003; Yamano, Alderman and Christiaensen, 2005) influence nutritional status. We contribute to this discourse by examining a different type of shock: the creation of a refugee camp for a sizable population of people experiencing forced displacement.

We find evidence that the establishment of a refugee camp results in a decline in weight-for-age and height-for-age z-score indicators for children under the age of six within zero to 10 kilometers from the camp. We also find an increase in the Rohrer's index (a measure of body mass) for adult men within close proximity to refugee camps, despite no impact on the Rohrer's index for adult women. The measurable effects for children under the age of six are economically significant, representing a decline in weight-for-age z-scores by ten percent of the mean and a fall in height for age z-scores by eight percent of the mean for children living within 10 kilometers of a refugee camp.

Our results are robust to several checks we examine, including one in which we restrict our sample to only observations from countries with georeferenced DHS data collected before and after camp openings. Our mechanism analysis suggests that the nutritional losses were not due to worsening child health or negative employment outcomes for households. We also find evidence of in-migration in response to camp openings. Limiting our analysis to only those who never moved, the nutritional impacts for children are worse than when we do not account for migration. Based on our analysis, we argue that the worsening of child nutrition is either driven by highly localized changes in food prices or is due to selective out-migration in response to camp opening, which changes the demographic composition of the host community.

The remainder of this paper proceeds as follows. Section 2 describes and discusses the data we use and presents descriptive statistics. Section 3 outlines our empirical strategy and threats to identification. Section 4 presents and discusses estimation results, robustness checks, and potential mechanisms. Section 5 concludes and provides suggestions for future research.

2. Data

2.1 African Refugee Dataset (ARD)

Until recently, continent-wide geo-referenced data on refugee camps in SSA was not publicly available.[5] We constructed the African Refugee Dataset (ARD) to fill this gap. The ARD covers planned refugee camps open at some point between 1999 and 2016 and provides information on refugee camps' geographic locations and years of operation. The data do not include other forms of refugee residence, such as transit centers, informal tented settlements, or other "known refugee locations."[6] The ARD has recently been used in a spatially explicit study of the impact of camps on forest cover in their surrounding landscapes (Salemi, 2021).

[5] The Georefugee Dataset (Fisk, 2014) contains measures of refugee populations by subnational district in SSA from 2000-2010. However, it does not contain precise camp locations, only counts of refugee populations at the first subnational administrative unit for each country. Recently, Zhou and Shaver (2021) released global data on locations of forcibly displaced populations along with operational years for locations such as camps.
[6] "Known refugee locations" in the sources we used to build the ARD primarily represent clusters of refugees living in villages or cities, though other less conventional modes of residence may be lumped into this category. Such

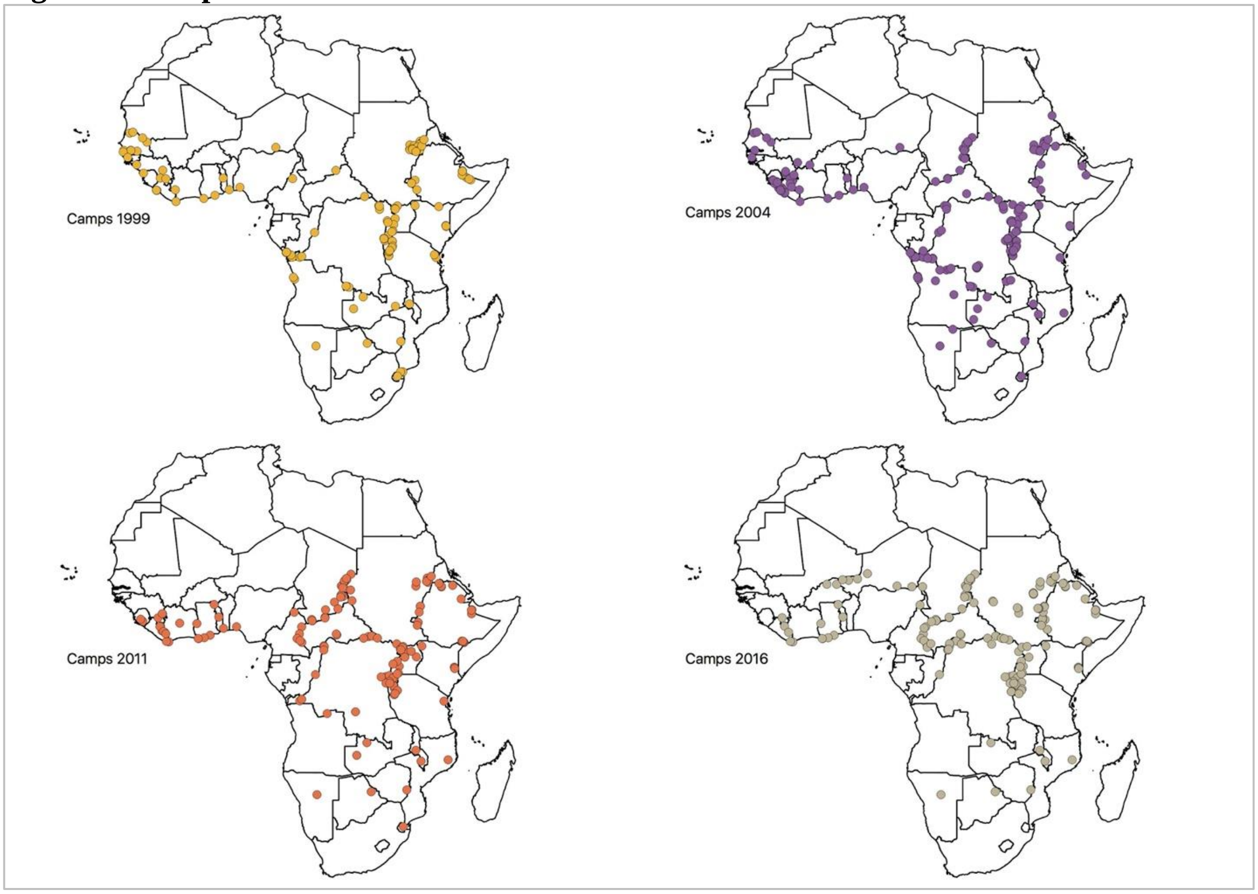
We build the ARD using various sources publicly available from the UNHCR and widely used geospatial databases. The primary resource we use is the list of locations where known "People of Concern" (POC) have been residing as reported in the UNHCR's annual Statistical Yearbooks from 1999 to 2016 (except for the 2001 yearbook),[7] supplemented with the UNHCR's POC map from 2018 (UNHCR, 2018, n.d.). This list is the most complete available set of information on the locations where refugees (and other people of concern, such as internally displaced people) are located over time. We list the various nongovernmental, academic, and media resources used to date, locate, and confirm the geocoordinates of each individual camp and its years of operation in the ARD data itself. We also provide a more complete description of the steps involved in creating the ARD in Appendix Section A1 and in Anti, Salemi and Wilson (2020).

Figure 1 displays the spatial distribution of refugee camps in the region, showing how the locations changed or stayed constant over the period from 1999 to 2016. The figure reflects the spatial trends of conflict over this period. For example, we see the decline in the prevalence of refugee camps in Sierra Leone and Angola as conflicts in those regions stabilized over the period. At the same time, refugee camps in eastern Chad and eastern Cameroon become much more prevalent over the period in response to violence in Darfur and the Central African Republic. Meanwhile, the multitude of camps in the Great Lakes region of eastern Africa remained relatively stable over the period, reflecting the protracted displacement of certain refugee groups (such as the Congolese) due to entrenched conflict and insecurity in the home country.

locations are geo-identified in UNHCR data in various ways, sometimes as the city or village, but at other times as the district.

[7] The 2001 yearbook adopted a different format than all the others and omits the Persons of Concern Locations table.

Figure 1: Camp Locations 1999-2016



Source: authors' calculations based on the ARD

2.1 Demographic and Health Surveys (DHS)

We use the geocoded survey data available from the Demographic and Health Surveys (DHS) to measure host outcomes for people across SSA as they relate to respondent proximity to one or more of the refugee camps contained in the ARD (ICF, no date) The DHS are nationally representative repeated cross-sections that collect information on respondent health, nutritional status, and fertility. They also identify other characteristics such as respondent age, employment, family structure, educational attainment, and (in some years) how long the respondent has lived in their current residence. While there is some variation over the age range of respondents in the

DHS from country to country, the surveys focus on women below the age of 50 and their children.[8]

The DHS data offer other advantages for our current study. Because they omit institutionalized populations, the DHS data do not include any respondents who reside in a refugee camp (Carr-Hill, 2013), which helps us mitigate spillovers of refugee camp residents into the host sample. While the DHS are primarily concerned with tracking the health of women and children, about a third of the surveys collect detailed information on male respondents as well (Pullum *et al.*, 2017), which we exploit for our analysis. Additionally, the DHS program has collected data for countries across SSA. For many countries, several years of cross-sections exist, so our study is not restricted to a particular region or time. Moreover, the DHS data are designed to facilitate harmonization across countries and time, and for this study we merge 62 country-year datasets in SSA. Table A4.1 in the Appendix lists all the DHS country-years used in this study.

For many of the surveys conducted in recent decades, the DHS data contain the longitude and latitude of the survey's primary sampling units, known as "clusters." For the purposes of anonymity, the DHS randomly displaces its clusters' geographic locations.[9] When this random displacement resulted in the cluster being listed as outside of its actual country's borders, we imputed that cluster's subnational administrative unit by assigning it to the closest administrative zone in the country to which it belongs. This random displacement introduces measurement error in our independent variables of interest in our analysis, but this simply attenuates our estimates,

[8] The DHS are widely used in social science research. Within the literature on the effects of refugees and refugee camps, Alix-Garcia and Saah (2009) employ the DHS to examine the impacts of refugee hosting in Tanzania, and Tatah *et al.* (2016) use the DHS to study refugee settlement and changes in host population health service provision in Cameroon.
[9] Urban clusters are randomly displaced 0-2 km. Rural clusters are randomly displaced 0-5 km, with 1% displaced 0-10 km, from the actual geo-coordinate. See Perez-Heydrich *et al.* (2013) for more on geo-located DHS cluster data.

since the error is purely random. With this in mind, we measure the Euclidean distance between the DHS cluster and the nearest refugee camp for each year from 1999 to 2016.

Figures A3.1 and Figure A3.2 in the Appendix show the country representation in both the children's and women's sample residing within five kilometers of a camp. Rwanda has the largest representation, as one would expect from its population density and proximity to conflict in the eastern DRC. But overall, the treated sample contains observations from 17 countries from across different regions of the continent. Table 1 contains the summary statistics for the entire DHS sample of respondents within 100 km of a camp between 1999 and 2000. It also provides descriptive characteristics for respondents living within five kilometers of an operational camp, within 20 kilometers of a camp location (operational or non-operational), and those we use as a comparison group living in clusters beyond 20 and up to 100 kilometers away from a camp location. The descriptive statistics suggest that mean age and years of education, along with the gender distribution, do not vary much by distance to the nearest camp. However, respondents close to the camp do tend to be slightly more likely to reside in an urban area than those further away. The z-score variables are marginally lower for respondents within 20 kilometers of a camp compared to those further away.

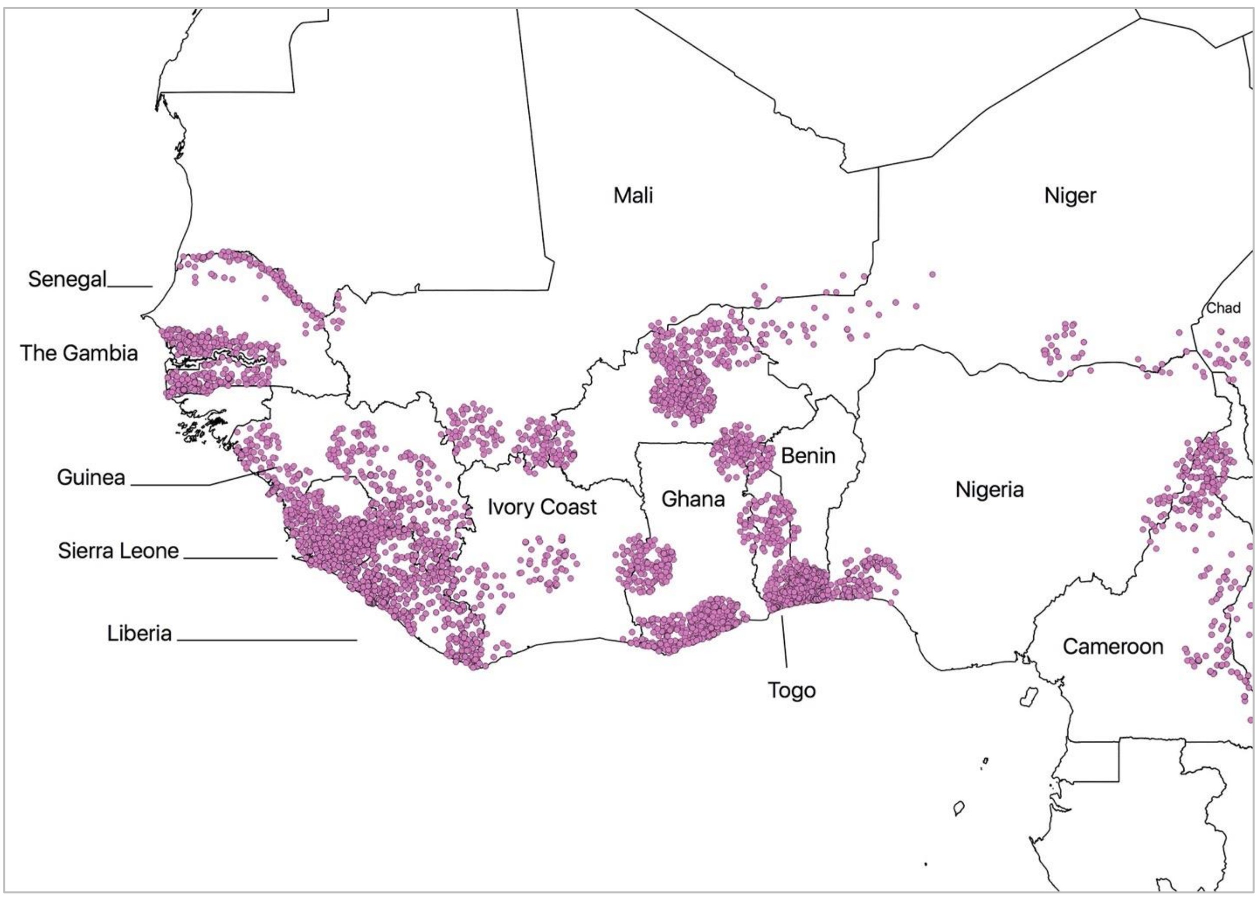
Table 1: Summary Statistics by Euclidean Distance to Nearest Camp and Treatment Status

Variable	Whole Sample		Within 20 km of Camps		Treated by Camps within 5 km		Control Group (>20 km and <=100 km from Camp)	
	Obs.	**Mean**	**Obs.**	**Mean**	**Obs.**	**Mean**	**Obs.**	**Mean**
Age	1,514,921	21.78	262,668	21.65	12,861	21.10	1,252,253	21.81
Sex	1,514,997	0.51	262,671	0.52	12,860	0.51	1,252,326	0.51
Yrs. Education	1,507,562	2.83	261,567	2.79	12,780	2.79	1,245,995	2.83
Rural	1,515,016	0.72	262,674	0.68	12,861	0.54	1,252,342	0.73
HAZ	134,495	-141.68	22,547	-145.52	1,114	-141.41	111,948	-140.91
WHZ	137,133	-30.37	23,104	-32.73	1,123	-31.52	114,029	-29.90
WAZ	134,495	-112.17	22,547	-115.90	1,114	-112.52	111,948	-111.42
Women's Hgb	122,408	122.37	21,281	123.76	886	125.85	101,127	122.07
Men's Hgb	41,538	140.26	6,544	141.56	208	143.20	34,994	140.02
Women's Rohrer's	186,567	1,422.03	32,015	1,411.96	1,564	1,415.98	154,552	1,424.12
Men's Rohrer's	38,470	1,254.16	9,004	1,257.84	303	1,319.37	29,466	1,253.04
Year Between '00 and '05	1,388,403	0.14	233,024	0.09	9,247	0.09	1,155,379	0.16
Year Between '05 and '10	1,515,016	0.19	262,674	0.24	12,861	0.46	1,252,342	0.18
Year Between '10 and '16	1,515,016	0.68	262,674	0.68	12,861	0.48	1,252,342	0.68
Treated within 5 km	1,515,016	0.01						
Treated within 10 km	1,515,016	0.02						
Treated within 15 km	1,515,016	0.04						
Treated within 20 km	1,515,016	0.06						
Camp 5 km	1,515,016	0.03						
Camp 10 km	1,515,016	0.07						
Camp 15 km	1,515,016	0.12						
Camp 20 km	1,515,016	0.17						

Source: authors' calculations based on DHS datasets used for study and ARD

Figure 2 displays the spatial distribution of the DHS clusters within 100 kilometers of a refugee camp between 2000 and 2016 in West Africa.[10] As the figure illustrates, this sampling method captures a large share of the DHS clusters, some of which appear in the repeated cross-section before the nearest camp opening, while others appeared after. For further reference, Figures A3.3 and A3.4 in the Appendix show the spatial distribution for the DHS clusters in the sample for Central and East Africa and Central and Southern Africa.

Figure 2: DHS Cluster Locations within 100 Km of a Refugee Camp, West Africa 2000-2016



Source: authors' calculations using the ARD. Maps for East, and Central/Southern Africa are located in the Appendix.

[10] As we discuss in the next section, we use clusters within 100 km of a refugee camp for our analysis. To improve clarity, we are only showing one region of the continent in this example. We provide additional maps for other regions in the Appendix.

Section 3: Empirical strategy

3.1: Spatial Difference-in-Differences (DID) at 10 and 15 Kilometers

Forced migration patterns and refugee camp placements are not randomly determined (Maystadt and Verwimp, 2014; Ruiz and Vargas-Silva, 2013), so we employ a quasi-experimental estimation approach to account for endogeneity in refugee camp locations. Our empirical strategy exploits the repeated cross-sectional structure of the DHS data and the variation in camp locations over time to implement a multi-period DID estimation as developed in Benshaul-Tolonen (2018) and employed in Knutsen *et al.* (2016), Kotsadam and Tolonen (2016), Kotsadam *et al.* (2018), von der Goltz and Barnwal (2019) and Anti (2021).

In this approach, we compare respondents whose cluster lies within a certain distance to the nearest refugee camp to those outside of that distance over time. We follow Benshaul-Tolonen's (2018) analysis of localized effects of mining in SSA and limit the sample to observations that are within 100 kilometers of a camp between 1999 and 2016, since we think it is more plausible that the parallel trends assumption holds between areas within close proximity to a camp and those close by but not within the immediate proximity. Establishing a binary cutoff that defines the spatial extent of the effect of the camp requires that we consider how host and refugee mobility influences respondent access to markets and institutions. We follow previous studies of mining in SSA (Kotsadam and Tolonen, 2016; Benshaul-Tolonen, 2018) and literature on the commuting distance of people in SSA (Kung *et al.*, 2014), which suggest that localized agglomeration effects are generally within zero to 20 kilometers.[11] For this reason, we

[11] Various studies have suggested different cutoff distances. Shafer (2000) suggests around five kilometers for Tanzania and Ghana. Kung et al. (2013) and Aidoo, Amoh-Gyimah and Ackaah (2013) show findings from Ivory Coast and Ghana that suggest a five to 15 kilometers distance, respectively. Bunte *et al.* (2018) use a 20 kilometers distance. Benshaul-Tolonen (2018) uses 10 kilometers, and Kotsadam and Tolonen (2016) use 20 kilometers.

designate those 20 to 100 km away from camps as the comparison group and evaluate several distance buffers from the camps in separate DID frameworks.

We estimate the following specification using ordinary least squares (OLS),

$$Y_{icdtj}^{k} = \beta_0 + \beta_1 D_{it}^{k} + \beta_2 Camp_{ik} + \varphi X_i + \delta_d + \theta_{d*t} + \gamma_{pt} + \mu_{dj} + \varepsilon_{icdtj} \quad (1)$$

where Y_{icdtj}^{k} is a nutritional outcome of individual i in DHS cluster c and second subnational unit (district) d sampled in year t during month j when treatment exposure is measured based on distance cutoff k.[12] The $Camp_{i,k}$ variable is a binary indicator equal to one if the observation is within distance k of the nearest refugee camp in any year between 1999 and 2016 (regardless of whether the camp has opened yet) and equals zero otherwise. The D_{it}^{k} term is the DID measure of interest. It is equal to one if the individual is in a DHS cluster that is within k kilometers of an operating refugee camp the year the respondent was surveyed and equals zero otherwise. Therefore, the coefficient estimate of β_1 measures the impact of being located within k kilometers of a refugee camp assuming parallel trends in the outcome variable hold for those within and outside of the distance cutoff and controlling for additional covariates.

The $Camp_{ck}$ variable accounts for time-invariant unobservable variables associated with selection into being close to a refugee camp. Additionally, we control for X_i, a vector of individual controls that includes the respondent's sex, age, household size, and whether the respondent's household is in a rural or urban area. The vector also includes the education level, age, and sex of the respondent's household head.

---

[12] The second subnational administrative unit would correspond to a district or *préfecture* in many countries. We henceforth refer to this administration unit as a "district", but please note that not all countries in our sample use this term to refer to this geographic administrative level.

We use a series of fixed effects to account for time-variant and time-invariant heterogeneity across years of the DHS surveys and administrative units. First, we include δ_d, a vector of fixed effects at the district level, and θ_{d*t}, a vector of linear time trends that is also at the district level. Next, γ_{pt} is a vector of interaction terms between the first subnational administrative unit (province)[13] and the survey year dummies, controlling for year- and province-specific shocks. Since provinces are nested within countries, this controls for macroeconomic shocks at the country level in a given year. Finally, μ_{dj} is a vector of interactions between month dummies and district indicator dummies, controlling for regional seasonality (Berazneva and Byker, 2017).

The nutritional outcomes that we examine as dependent variables throughout this paper for children are the height-for-age z-score (HAZ), weight-for-age z-score (WAZ), and weight-for-height z-score (WHZ) for all respondents below the age of six. We use the variables provided by the DHS, which they calculate relative to the distribution of a healthy child population and multiply by 100. The z-score measures for children each represent distinct dimensions of malnutrition and poor health related to food insecurity. Low HAZ, known as "stunting," is the result of inadequate nutrition for an extended period of time, while low WAZ, known as being "underweight," reflects both current and acute, as well as chronic malnutrition. Having a low WHZ, known as "wasting," is a measure of current acute malnutrition (INDEPTH, 2008).

For adults between the ages of 14 and 50, we look at the blood hemoglobin (Hbg) and the Rohrer's index. Hbg is a biomarker recommended by the World Health Organization (WHO) as an indicator of anemia, a condition that results from vitamin and mineral deficiencies, especially

[13] As with the second administrative unit, countries refer to their first administrative unit by different names, such as province or county. We refer to this geographic administrative level as a *province* throughout the paper.

iron deficiency (Pullum et al., 2017). The Rohrer's index is a measurement of leanness similar to BMI and is defined as,

$$Rohrer's\ Index_i = \frac{mass_i}{{height_i}^3}. \tag{2}$$

The Rohrer's Index is preferable to BMI measures because it yields valid comparisons across all body types, even the very tall and very short. Table A4.2 contains all variables, their precise definitions and sources.

For equation (1) to estimate the causal relationship between refugee camps and local nutrition outcomes, the timing and location of refugee camps must be exogenous to local nutrition trends. Based on key informant conversations in East Africa, we have learned that refugee camp placement depends on several factors[14] that are time-invariant prior to camp construction. Hence, our specification should control for these factors through the inclusion of the $Camp_{c,k}$ variable.

The specification also accounts for policy changes and other year-specific shocks by controlling for the survey year fixed effects and unobserved variables that trend smoothly within district units over time, captured by the linear time trends specific to the district. We additionally account for the regional seasonality of nutrition through the inclusion of the month-district fixed effect. Growing seasons and periods of acute hunger can vary at a relatively localized level within SSA (see HarvestChoice (2015) for the variation in the growing season over the

[14] The first consideration key informants listed is the region's proximity to the conflict or natural disaster in the neighboring country that prompted the refugee flow. Host governments tend to place camps close to the point at which asylum-seekers are crossing the border into the host country, but also strive to maintain the UNHCR standard of keeping camps at least 50 kilometers from the border due to protection concerns. Of additional importance is land availability and tenure: camps tend either to be placed on government-owned property, or host governments acquire uninhabited communal land through negotiation with the groups who collectively manage the territory. Host governments and UNHCR site planners also consider multiple factors related to the logistics of camp management, including topography, road networks, access to preexisting services and infrastructure, and the availability of water via boreholes, rivers, or other natural resources. For more on this discussion, see Salemi (2021).

continent), hence our decision to use the district for this fixed effect rather than one at the province level.

Another more serious threat to identification in this approach is treatment-induced migration. We find evidence in our own data of migration towards refugee camp locations (see Section 4.3). This may be due to inclusive services and improved access to water, schools, and medical care for hosts living alongside refugee camps, or because of economic growth in the camp area and resulting job opportunities. It is also possible that particular types of host households move away from refugee camps once they open. With repeated cross-sectional DHS data, we cannot directly identify how camp openings impact a particular household, as we only observe each household once and we have very limited information on households moving. As such, our estimates tell us about the impact on communities or locations, and these estimates may be driven by changes in sample composition.

In an attempt to account for selective inward migration, we follow Kotsadam and Tolonen (2016) and Benshaul-Tolonen (2018) and specify a robustness check in which we restrict the sample to respondents who have not moved since the arrival of a camp within 20 kilometers so that the sample of treated households are only those that have been in the region since before a close-by refugee camp arrives.[15] This approach removes those households induced to move closer to the camp from the exposure group. We acknowledge this robustness check has its limitations because there may still be endogenous selection away from the camp, with certain groups of hosts choosing to relocate, thus influencing the community's composition. For this reason, we include a second robustness check restricting the sample to household heads

[15] The DHS collects migration data for adults but does not directly identify how long a child has lived in a given location. When estimating on the adult sample, we use the adult's migration information. When estimating on the children's sample, we proxy using the child's household head's migration information.

who never moved and respondents who never moved when analyzing the samples of children and adults, respectively.

Following Abadie et al. (2017) we cluster standard errors at the DHS cluster level in all estimations. These DHS clusters correspond to the enumeration areas used in the first stage of DHS sampling.[16] Our analysis looks at several outcome variables, which generates a concern of false positives under multiple hypothesis testing. To avoid this, we provide the Benjamini-Hochberg corrected q-value along with the naïve p-value from the regression, which controls for the false discovery rate (Benjamini and Hochberg, 1995). We prefer it over the Bonferroni correction because it involves less risk of a false negative than the Bonferroni correction, allows for more statistical power (Fink, Mcconnell and Vollmer, 2014) and works well when outcomes of tests are correlated (Benjamini and Yekutieli, 2001; Genovese and Wasserman, 2002).

3.2 Heterogeneous Distance Bands

We extend the DID specification above to examine whether there are heterogeneous impacts by distance to a respondent's closest refugee camp. To do this, we construct four distance bands, each five kilometers wide, around refugee camp locations and estimate the following specification using OLS,

[16] According to Abadie et al. (2017), the enumeration areas (EAs) selected in the first stage are just a small share of all enumeration areas in a particular DHS country. The subsequent lack of representation of many enumeration areas motivates the use of clustered standard errors at the level of the EAs. The DHS does not publicly disseminate the EAs used for sampling. Since they are based on country Census EA tracks, they may be available for some countries. But most countries in our sample do not publicly share their EA tracks out of privacy concerns, so we are unable to identify unique enumeration areas for all of our sampled households. We believe that the DHS sampling clusters better represent the EAs than the larger district that the cluster resides in. Moreover, Abadie et al. (2017) warns against clustering at a higher aggregation level than that used in the sampling strategy. Clustering standard errors at such a high level may result in "standard errors that are unnecessarily conservative" (Abadie et al., 2017).

$$Y_{icdtj} = \eta_0 + \sum_s \eta_s D_{it}^s + \sum_s \kappa_s Camp_s + \rho X_i + \delta_d + \theta_{d*t} + \gamma_{pt} + \mu_{dj} + \varepsilon_{icdtj}, \quad (3)$$

where all variables are defined as before and the set s is composed of the distance bands, measured in kilometers, so that $s = \{[0,5], (5,10], (10,15], (15,20]\}$. This specification yields estimates for each five-kilometer band around a camp, keeping those between 20 and 100 kilometers away as the reference group in the DID setup. The estimated coefficients for each distance band s, $\{\hat{\eta}_{[0,5]}, \hat{\eta}_{(5,10]}, \hat{\eta}_{(10,15]}, \hat{\eta}_{(15,20]}\}$, provide estimates of the impact of the establishment of a refugee camp for each of the distance bands in s under the assumptions necessary for internally valid DID estimates.

We expect the impacts of the camps to be more intense and measurable within closer proximity, so if the effect dissipates over space, it provides confidence in our estimates. Following the previous literature of using 20 kilometers as the cutoff distance for the DID specification leads us to consider the threat of a Stable Unit Treatment Variable Assumption (SUTVA) violation, since we are unsure of the point at which the impact of the refugee camp fully diminishes over space. To address this, we include specifications of equation (3) adding those in the distance bands from 20 to 25 and 25 to 30 kilometers from camps, redefining our treatment group as those from 30 up to 100 kilometers away as the comparison group, in our main results.

3.3 Testing for Homogeneous Treatment Effects

Recent advances in econometrics have highlighted potential sources of bias in DID specifications with staggered treatment timing. In this case, the average treatment effect (ATE) estimate is a weighted average of the ATE's for all 2x2 group-time DID pairs in the data (Goodman-Bacon, 2021). If the treatment effect is homogeneous across treatment groups, this weighted average is unbiased. But when treatment effects are heterogeneous, the weighted average ATE will be biased (Borusyak and Jaravel, 2018; de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021). There are even cases in which the bias is severe enough to flip the sign on the coefficient of interest (Baker, Larcker and Wang, 2021).

Econometricians have offered numerous new estimation approaches when treatment timing is staggered and treatment effects are heterogeneous, but most of these solutions assume a panel data structure (de Chaisemartin and D'Haultfœuille, 2020; Gardner, 2021; Roth and Sant'Anna, 2021). One exception is Callaway & Sant'Anna (2020), who provide methods compatible with repeated cross-sectional data. But their doubly robust estimator cannot yield results when the repeated cross-section is not balanced in calendar time, which is the case when merging DHS data from different country-years. Modifying these estimation strategies to make them compatible with non-balanced repeated cross-sectional data is beyond the scope of the paper.

Jakiela (2021) provides diagnostic tests to evaluate the satisfaction of the homogeneity assumption. Drawing on the Frisch-Waugh-Lovell Theorem, her approach relies on the relationship between residualized outcome and residualized treatment variables. If the homogeneity assumption is satisfied, then the relationship between residualized outcome and treatment variables (the slope) should be linear and the same for the treated and untreated units. Following Jakiela (2021), we estimate the residualized outcome $\tilde{Y}_{it}$ and residualized treatment $\widetilde{D}_{it}$ for regression specifications (1) and (3).[17] We then perform the following regression to see if the slopes of the linear relationship between $\tilde{Y}_{it}$ and $\widetilde{D}_{it}$ vary when we compare treated and untreated observations:

$$\tilde{Y}_{it} = \alpha_0 + \alpha_1\widetilde{D}_{it} + \alpha_2 Treated_{it} + \alpha_3\widetilde{D}_{it} \times Treated_{it} + v_{it} \quad (4)$$

A significant coefficient estimate for α_1 and a null coefficient estimate for α_3 would provide support for the satisfaction of the treatment homogeneity assumption, as this suggests that the relationship between $\tilde{Y}_{it}$ and $\widetilde{D}_{it}$ is linear and is the same for the control and treated observations. If both α_1 and α_3 are significant, then this is evidence that treatment effects are heterogeneous across groups and time. We perform this diagnostic test on all specifications, and in each table of results, we indicate the outcome for each significant coefficient estimate. For these estimates, we identify whether the diagnostic test provides evidence of treatment effect homogeneity based on the following criteria: the corresponding α_1 term is statistically significant at the 10% level or higher, and the α_3 term is statistically insignificant. We report the full results of this test in Appendix Section 9.

[17] To estimate the residualized outcome, we regress the outcome variable on all variables in the specification except the treatment indicator and store the residuals. To estimate the residualized treatment, we regress the treatment indicator on all variables in the specification except the outcome variable and store the residuals. With heterogeneous distance bands, we repeat the diagnostic test for each coefficient of interest in the specification, treating all other variables as covariates. For example, to test the homogeneity assumption for treatment exposure at 0-5 km, we use treatment at 0-5 km to determine the treatment residual, regressing this on all other variables in the regression. We then regress the outcome on all other variables in the regression (omitting the variable indicating treatment at 0-5 km). Consequently, each regression with heterogeneous distance bands requires multiple separate homogeneity tests.

3.3 Event Study Specification

There is reason to suspect that the nutrition impacts of refugee camps are not consistent over time. If there is some positive demand shock, it is possible that over a relatively long time period, local food markets will adjust to accommodate the increased demand. Thus, the negative effects may be limited to an undefined shorter period of time after the establishment of the camp. To examine the impacts of refugee camps on host community nutrition outcomes over time, we amend equation (2) to estimate an event study. This specification is as follows,

$$Y_{icdtj} = \psi_0 + \sum_r \psi_r \, Camp_{i,[0-5]} \times Time_r + \sum_s \nu_s \, Camp_{i,s} \times After_t + \sum_e \lambda_e \, Camp_{c,e} + \pi X_i + \delta_d + \theta_{d*t} + \gamma_{pt} + \mu_{dj} + \varepsilon_{icdtj}, \quad (5)$$

where all variables are defined as in equations (1) and (3). The vector $Time_r$ is a set of binary variables identifying whether the observation is treated within a certain time period relative to the establishment of the camp within zero to five kilometers from its location. We choose aggregate time bands since the treated sample is relatively small and aggregating the time periods provides more power for the estimation. For example, for the child outcomes, we use four-year time periods so that r is defined as $r \in \{[-12,-8],\ [-7,-3],\ [1,4],\ [5,9],\ [10,13],\ [14,20]\}$. Here, the impacts are measured relative to the period from two year to zero years before the establishment of the refugee camp. Note that we adjust the set r for the event study specifications based on data availability. We drop observations from locations that are in areas nearby a disbanded camp to avoid contaminating our estimates from some lingering agglomeration effect. We simply build on the specification in equation (3) to be consistent with its estimate, so the set s is now the set of distance bands other than the zero to five-kilometer band such that $s \in \{(5,10], (10,15], (15,20]\}$, and the set e is the complete set of distance bands such that $e \in \{[0,5],\ (5,10],\ (10,15],\ (15,20]\}$. The OLS estimates of

$\{\hat{\psi}_{[-12,-8]}, \hat{\psi}_{[-7,-3]}, \hat{\psi}_{[1,4]}, \hat{\psi}_{[1,4]}, \hat{\psi}_{[5,9]}, \hat{\psi}_{[10,13]}, \hat{\psi}_{[14,20]}\}$ are our coefficients of interest in this specification. They provide measures of the impact of being located within five kilometers of a refugee camp over time, for periods both before and after the establishment of the refugee camp.

This specification additionally helps us evaluate the validity of the parallel trends assumption necessary for the causal inference of the DID estimator (Angrist and Pischke, 2009). If the estimated effects of proximity to a refugee camp are similar in the pre-treatment and post-treatment periods, the results indicate there is no treatment impact. If there are significant differences between treated and untreated groups in the pre-treatment period, then this is an indication that the parallel trends assumption is violated. However, if estimated impacts in the pre-treatment period are statistically insignificant or close to zero, then it is evidence in support of parallel trends over the pretreatment period.

Recent econometric advances have shown that in the absence of homogeneous treatment effects, event study coefficient estimates will be contaminated by treatment effects from other relative time periods (Sun and Abraham, 2020). This is also the case for coefficient estimates in the pre-treatment period, meaning the seemingly zero differential trend may be nonzero. As we will show in the next section, our evidence in support of treatment effect homogeneity across specifications assuages these concerns. But even with homogeneous treatment effects, Roth (2019) argues that "noise" may conceal a nonzero differential trend in the pretreatment period. We urge our readers to keep these concerns in mind when evaluating our event study results.

We report the results of the event study in Section A5 of the Appendix using WAZ, HAZ, WHZ, and the Rohrer's index for men and women as outcome variables. The figures display confidence intervals reflective of the naïve p-values, although the statistical significance is mostly robust to the Benjamini-Hochberg adjustment. The estimates suggest that there is no significant difference in trends in the pretreatment periods, with a negative estimate in the period immediately following the establishment of a refugee camp.

3.4 Mechanism Analysis

We explore the mechanisms driving our results in several ways. First, we use alternative health data available in the DHS to inform our results. The outcomes we use are the incidence of coughing, fever, or diarrhea among children in the past two weeks. If the results are in the same direction as the nutrition effects, it is suggestive that refugee camps have some broader negative impact on health in the region and that this shock to health may contribute to the worsening nutritional status of children.

We also examine heterogeneity in the impacts of the refugee camp by the respondent's marital status or the marital status of a child's household head. Heterogeneous results by marital status provide insight into whether employment opportunities generated by the camp are improving nutrition outcomes for some populations, since households with married heads are in a better position to seek employment related to the camp if the opportunity arises. For this, we amend equation (1) into a triple difference estimator by interacting $Camp_{c,k} \times After_{ct}$ with a binary indicator for whether the respondent or household head was married at the time of the survey.

Last, we consider the role of changes in employment opportunities. We do this by constructing a binary variable for whether an adult respondent is employed at the time of the survey. We use this indicator as an outcome variable in our heterogeneous treatment effects specification (Equation 3). Following Kotsadam and Tolonen (2016), we estimate this specification omitting the month-district fixed effect.[18]

4. Results

4.1 DID Main Results

All mentions of statistical significance in this discussion refer to the naïve measures. We discuss the implications of the Benjamini-Hochberg procedure on statistical inference in Section 4.5. Table 2 contains the results of estimating equation (1) using the cutoffs of 10 and 15 kilometers around a camp and the z-score variables.[19] We see statistically significant negative coefficients for the WAZ and the HAZ outcomes at both thresholds. For all four of these results, we find evidence in support of homogeneous treatment effects using the diagnostic test outlined in Jakiela (2021).

The estimates indicate that being within 10 kilometers of an active refugee camp results in a 0.11 standard deviation decrease in child weight-for-age. We consider this result to be economically significant, as it is about 10 percent of the mean WAZ for the sample. We also see that being within 10 kilometers of a camp results in a decrease in child height-for-age by about 0.11 standard deviations, although this estimate is statistically significant at only the 10 percent level. This is about 8 percent of the magnitude of the sample mean for HAZ. Results for the adult sample with equation (1) show null results for women, but strong positive gains for men's Rohrer's index within 10 kilometers of a camp (and evidence of the satisfaction of the

[18] We believe the inclusion of the month-district fixed effect will only influence our estimation if the within-district difference in employment outcomes is highly related to the month of DHS data collection. As a double-check, we estimate Equation 3 with adult employment as our outcome variable and including the month-district fixed effect. The results of this specification are available upon request. We find no difference in terms of sign or significance when we include the month-district fixed effect. The coefficient magnitudes are in a similar range as those in Table 6.
[19] The five-kilometer threshold results in a negative coefficient with statistical significance at the ten percent confidence level for the WAZ outcome.

homogeneity assumption for this result). We report these outcomes in Table A5.1 in the Appendix.

Table 3 contains the results of estimating equation (3) for the sample of children. All statistically significant coefficient estimates pass the homogeneity assumption test as defined in Section 3.3. The estimates are consistent with our expectations, as the impact is strongest closer to the camp and dissipates monotonically over space out to 20 kilometers. Within five kilometers of a refugee camp, children's WAZ declines by 15.41, 13.7 percent of the sample mean. The HAZ declines by 16.93 in areas within five kilometers of an active refugee camp, about 12 percent of the sample mean. These results are significant at the 5-percent and 10-percent level respectively.

The estimates for the adult sample are in Table 4 and show strong statistically significant coefficients for men. The results indicate adult men living within five kilometers of a refugee camp have higher Rohrer's index levels than those farther away. Our diagnostic testing provides evidence of homogeneous treatment effects for this result. For women, we find no evidence of a change in the Rohrer's index, but we do estimate a significant, positive impact on Hbg for those in the 5-25 km buffers. Again, these statistically significant results are supported by evidence of the satisfaction of the homogeneous treatment assumption.

Table 2: Coefficient estimates for regressions using anthropometric measures as outcome variables and treatment groups at 10 and 15 km thresholds

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	WAZ	HAZ	WHZ	WAZ	HAZ	WHZ
$Camp_{c,10} \times After_{ct}$	-11.330**, ††	-11.201*, †	-6.020			
	(4.620)	(6.000)	(4.225)			
B-H q-values	[0.042]	[0.059]	[0.154]			
$Camp_{c,10}$	0.408	4.796	-1.709			
	(2.899)	(3.687)	(2.480)			
Homog. T.E.?	YES	YES				
$Camp_{c,15} \times After_{ct}$				-7.698**, †	-10.767**, †	-0.954
				(3.731)	(4.911)	(3.478)
B-H q-values				[0.059]	[0.059]	[0.784]
$Camp_{c,15}$				2.756	7.995**	-1.721
				(2.549)	(3.277)	(2.338)
Homogen. T.E.?				YES	YES	
Dep. Var. Mean	-112.166	-141.593	-30.420	-112.166	-141.593	-30.420
Observations	134,060	134,060	136,692	134,060	134,060	136,692
R-squared	0.127	0.146	0.124	0.127	0.146	0.124

Source: authors' calculations using DHS surveys and ARD. Robust standard errors clustered at the DHS cluster level in parentheses. All regressions include controls as outlined in equation (1). Their estimated coefficients are omitted. *** p<0.01, ** p<0.05, * p<0.1. Benjamini-Hochberg q-values in brackets††† q<0.01, †† q<0.05, † q<0.1. "Homog. T.E.?" indicates the result of the test for evidence of the satisfaction of the homogeneous treatment effect assumption based on Jakiela (2021) as described in Section 3.3.

Table 3: Coefficient estimates for regressions using anthropometric measures as outcome variables and heterogeneous distance bands

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	WAZ	HAZ	WHZ	WAZ	HAZ	WHZ
$Camp_{c,[0,5]} \times After_{ct}$	-15.407**, †	-16.931*, †	-6.966	-16.639**, †	-17.471*, †	-8.181
	(6.969)	(9.028)	(6.416)	(7.168)	(9.327)	(6.572)
B-H q-values	[0.081]	[0.092]	[0.278]	[0.06]	[0.092]	[0.213]
Homogen. T.E.?	YES	YES		YES	YES	
$Camp_{c,(5,10]} \times After_{ct}$	-11.871**, †	-14.775*, †	-2.767	-13.182**, †	-15.495*, †	-3.985
	(5.668)	(7.923)	(5.254)	(5.868)	(8.247)	(5.422)
B-H q-values	[0.093]	[0.093]	[0.598]	[0.075]	[0.09]	[0.462]
Homogen. T.E.?	YES	YES		YES	YES	
$Camp_{c,(10,15]} \times After_{ct}$	-3.581	-8.085	2.950	-4.893	-8.796	1.740
	(4.650)	(5.695)	(4.373)	(4.944)	(6.142)	(4.628)
$Camp_{c,(15,20]} \times After_{ct}$	-0.419	0.169	0.260	-1.888	-0.622	-1.091
	(3.895)	(5.211)	(3.633)	(4.288)	(5.850)	(3.984)
$Camp_{c,(20,25]} \times After_{ct}$				-2.989	-0.990	-3.743
				(3.995)	(5.047)	(3.810)
$Camp_{c,(25,30]} \times After_{ct}$				-2.500	-1.848	-1.141
				(3.375)	(4.342)	(3.140)
Dep. Var. Mean	-112.166	-141.593	-30.420	-112.166	-141.593	-30.420
Observations	134,060	134,060	136,692	134,060	134,060	136,692
R-squared	0.127	0.147	0.124	0.127	0.147	0.124

Source: authors' calculations using DHS surveys and ARD. Robust standard errors clustered at the DHS cluster level in parentheses. All regressions include controls as outlined in equation (3). Their estimated coefficients are omitted. *** p<0.01, ** p<0.05, * p<0.1. Benjamini-Hochberg q-values in brackets††† q<0.01, †† q<0.05, † q<0.1. "Homog. T.E.?" indicates the result of the test for evidence of the satisfaction of the homogeneous treatment effect assumption based on Jakiela (2021) as described in Section 3.3.

Table 4: Coefficient estimates for regressions using adult nutrition measures as outcome variables and heterogeneous distance bands

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All		Men		Women	
VARIABLES	Hgb	Rohrer's	Hgb	Rohrer's	Hgb	Rohrer's
$Camp_{c,[0,5]} \times After_{ct}$	0.492	-8.449	3.040	54.365***, †††	0.067	-22.479
	(1.178)	(13.478)	(1.988)	(16.894)	(1.223)	(15.377)
B-H q-values	[0.676]	[0.676]	[0.26]	[0.002]	[0.956]	[0.288]
Homgen. T.E.?				YES		
$Camp_{c,(5,10]} \times After_{ct}$	2.747***, †††	2.376	0.705	14.935	3.292***, †††	-1.777
	(0.878)	(8.977)	(1.876)	(12.240)	(0.910)	(10.377)
B-H q-values	[0.004]	[0.791]	[0.707]	[0.444]	[0.000]	[0.846]
Homogen. T.E.?	YES				YES	
$Camp_{c,(10,15]} \times After_{ct}$	1.385*	6.970	0.027	11.013	1.888**, ††	3.382
	(0.738)	(7.552)	(1.787)	(10.033)	(0.778)	(8.837)
B-H q-values	[0.122]	[0.356]	[0.988]	[0.544]	[0.03]	[0.7]
Homogen. T.E.?	YES				YES	
$Camp_{c,(15,20]} \times After_{ct}$	1.840***	-0.871	1.656	5.798	2.060***	-4.266
	(0.618)	(8.091)	(1.412)	(9.517)	(0.656)	(9.642)
Homogen. T.E.?	YES				YES	
$Camp_{c,(20,25]} \times After_{ct}$	1.436**	3.219	2.004	9.676	1.360**	1.196
	(0.577)	(6.881)	(1.342)	(8.093)	(0.604)	(8.243)
	YES				YES	
$Camp_{c,(25,30]} \times After_{ct}$	0.691	-0.059	1.534	-3.243	0.422	0.605
	(0.495)	(5.444)	(1.279)	(7.048)	(0.514)	(6.388)
	YES					
Dep. Var. Mean	126.660	1,395.018	140.642	1,253.298	122.367	1,421.905
Observations	159,385	221,081	37,411	35,233	121,925	185,816
R-squared	0.289	0.213	0.186	0.177	0.176	0.178

Source: authors' calculations using DHS surveys and ARD. Robust standard errors clustered at the DHS cluster level in parentheses. All regressions include controls as outlined in equation (3). Their estimated coefficients are omitted. *** p<0.01, ** p<0.05, * p<0.1. Benjamini-Hochberg q-values in brackets††† q<0.01, †† q<0.05, † q<0.1. "Homog. T.E.?" indicates the result of the test for evidence of the satisfaction of the homogeneous treatment effect assumption based on Jakiela (2021) as described in Section 3.3.

4.3 Migration and Additional Robustness Checks

Treatment-induced migration based on unobservable variables is the primary threat to identification that would violate the assumption of parallel trends in the post-treatment period because we are using repeated cross-sectional data. While we do not know of any pre-existing studies that claim to find increased migration towards camps, we must acknowledge that if refugee camps result in economic growth (Alix-Garcia *et al.*, 2018) or if refugee camps result in greater access to services for hosts, that there could be selective migration of those intending to take advantage of opportunities provided by the camp. Also, it is possible that refugee populations not in camps would settle close to the camp in an effort to be close to family or friends. Finally, camps may stimulate out-migration among those who have the means to relocate. This would mean that the remaining population is poorer on average.

In the absence of panel data, we cannot determine whether camps stimulate out-migration for a particular type of host household. But we can examine in-migration to some extent. Our first step is to determine whether such migration is taking place at all. To do so, we construct a binary variable equal to one if an individual has moved to their village and zero otherwise. We use this as the dependent variable in a regression specification identical to equation (3) but with the month-district fixed effects removed. Table A6.1 contains the results, which indicate that in-migration to the five-kilometer zone around refugee camps is occurring.[20]

[20] We have evidence to suggest this in-migration is not driven by non-camp refugees moving closer to refugee camps to be near friends and family. For one, countries that have pursued a policy of encampment have attempted to keep refugees in camps with few exceptions (often made on the basis of health, protection, or education needs). In countries where refugees have free mobility in and out of camps, those with the skills and resources needed to live outside of the camp (and forego a considerable amount of humanitarian assistance) will likely seek greater opportunities in the more prosperous urban areas of the country, instead of staying in the peripheral areas of the country near camps. We investigate this possibility formally using UNHCR's 2018 POC map data. We find that refugee camps are quite remote relative to areas where non-camp refugees reside. The average distance from each camp to its closest non-camp settlement is 153.3 kilometers, and it is rare for a camp to have a non-camp refugee settlement nearby. Only 5.6 percent of camps have non-camp refugee settlements within five kilometers, and 9.7 percent of camps have non-camp refugee

To see whether our results change when accounting for in-migration with our repeated cross-sectional data, we first follow Benshaul-Tolonen (2018) and Kotsadam and Tolonen (2016) and limit the sample to non-migrants. There is no migration information for child respondents, so we use the child's household head as a proxy of a child's migration history. We restrict to non-migrants using two strategies: focusing only on those adults and household heads who have never moved (in both the exposed and comparison groups) and restricting the exposure sample to those who did not move after the opening of a camp 20 km or closer. Tables A6.2 to A6.5 contain the results of these estimations using our preferred specification outlined in equations (1) and (3) for the child sample and equation (3) for the adult sample. We find that the results for children hold in sign, and although statistical significance is weaker, it is still present for several of the estimates, especially the WAZ outcome. The magnitude of the coefficients is actually larger, as one would expect if the inward migration is driven by new camp-generated work opportunities for hosts.[21] The coefficients for the men's Rohrer's index hold despite the reduction in sample size, at least when we use the 20-kilometer threshold for the cutoff.

To further check against refugee mobility within host country contaminating our results, we also use data available in the ARD on non-camp refugee locations. We measure whether a DHS cluster is within 20 kilometers of a non-camp refugee location, and if they are, we remove respondents in that cluster from the sample. Table A6.7 contains the results: the main findings hold in magnitude and sign. The same pattern of a declining size of the estimated impact is

settlements within 10 kilometers. Over 97 percent of non-camp refugee locations are located beyond 10 kilometers from a refugee camp. Figure A3.9 in the Appendix contains a map displaying these relationships in Central Africa.

[21] Our intuition is that those moving towards camps in pursuit of employment opportunities obtain a steady source of household income that will help them maintain their household nutritional status or that those moving for such opportunities are higher-skilled than the never-movers living alongside camps.

present, although we only see weak statistical significance on the closest coefficient, and this is not robust to the Benjamini-Hochberg correction.

As a robustness check on our estimation strategy, we limit the sample to countries for which we have DHS data collected before and after the opening date of least one camp opened in the country. This restricted sample includes observations from Cameroon, Benin, the DRC, Ethiopia, Ghana, Kenya, Liberia, Malawi, Nigeria, Rwanda, and Uganda, which still provides considerable geographic breadth across the subcontinent. We report these estimates in Appendix Section A8. The results from this reduced sample do not alter our findings using the larger sample from the broader continent.

4.4 Mechanism Analysis Results

As shown in Table 5, we do not find evidence that child health is worsening after camp opening.[22] Instead, children near camps are 7 percentage points less likely to have a fever in the two weeks prior to DHS sampling. The diagnostic test provides evidence of a homogeneous treatment effect, and this result is robust to our additional migration checks (see Appendix A7.1-7.2). We examine this in an event study framework and illustrate the results in Figure A5.6, which shows camp-induced reductions in the likelihood of a fever appear to be delayed and show up only in the later years in the sample. We interpret this to be the impact of improved health service provision for hosts residing near the camp and humanitarian investment in the welfare of hosts.

Our naïve estimates of camp-stimulated changes in employment indicate significant increases in the probability of being employed for men and women, but we fail to find evidence

[22] We report our full mechanism analysis in Section 7 of the Appendix.

of treatment effect homogeneity for most of the coefficients of interest (Table 6). This may be due to variation across treatment groups with respect to national labor restrictions imposed on refugees. This evidence of heterogeneous effects is inconsistent with the hypothesis that camp-related nutritional losses for children are mediated by parental loss of work. Were that the case, we would expect the treatment impact on child nutrition to also be heterogeneous. But with child nutrition as the outcome variable, diagnostic testing suggests that treatment effects are homogeneous.

Table 7 examines the impacts on children by marital status of the child's household head. Using the diagnostic test, we find evidence of the satisfaction of homogeneous treatment effects for the coefficient of interest in this triple difference estimation. We see that children with married household heads do significantly better than those that do not. Across all three measures, the impact of camps for children with unmarried household heads is negative and statistically significant. The impact of the camp with a married household head is positive across all three and statistically significant at the one-percent level for the WAZ and the HAZ measures. This is consistent with the notion that camps are providing income-generating opportunities, since it is likely that households with married household heads are in a better position to take advantage of those opportunities. Table A7.3 contains the results for the adult nutrition outcomes, which are generally not statistically significant. It appears that marital status is not a factor in differential outcomes across these measures.

Table 5: Coefficient estimates for regressions using alternative child health measures as outcome variables and heterogeneous distance bands

VARIABLES	(1) Cough	(2) Fever	(3) Diarrhea	(4) Cough	(5) Fever	(6) Diarrhea
$Camp_{c,[0,5]} \times After_{ct}$	-0.016	-0.071***, ††	-0.028	-0.021	-0.068**, ††	-0.029
	(0.026)	(0.027)	(0.023)	(0.027)	(0.027)	(0.023)
B-H q-values	[0.533]	[0.027]	[0.336]	[0.442]	[0.036]	[0.315]
Homogen. T.E.?		YES			YES	
$Camp_{c,(5,10]} \times After_{ct}$	-0.013	-0.008	0.019	-0.017	-0.006	0.019
	(0.023)	(0.023)	(0.019)	(0.024)	(0.024)	(0.020)
$Camp_{c,(10,15]} \times After_{ct}$	-0.015	-0.021	-0.012	-0.019	-0.019	-0.013
	(0.018)	(0.018)	(0.015)	(0.019)	(0.019)	(0.016)
$Camp_{c,(15,20]} \times After_{ct}$	-0.004	0.003	-0.010	-0.010	0.004	-0.011
	(0.015)	(0.016)	(0.013)	(0.017)	(0.018)	(0.014)
$Camp_{c,(20,25]} \times After_{ct}$				-0.009	0.018	-0.004
				(0.015)	(0.016)	(0.013)
$Camp_{c,(25,30]} \times After_{ct}$				-0.006	-0.012	0.004
				(0.014)	(0.013)	(0.012)
Dep. Var. Mean	0.278	0.281	0.179	0.278	0.281	0.179
Observations	120,368	120,406	120,428	120,368	120,406	120,428
R-squared	0.151	0.139	0.101	0.151	0.139	0.101

Source: authors' calculations using DHS surveys and ARD. Robust standard errors clustered at the DHS cluster level in parentheses. All regressions include controls as outlined in equation (3). Their estimated coefficients are omitted. *** p<0.01, ** p<0.05, * p<0.1. Benjamini-Hochberg q-values in brackets††† q<0.01, †† q<0.05, † q<0.1. "Homog. T.E.?" indicates the result of the test for evidence of the satisfaction of the homogeneous treatment effect assumption based on Jakiela (2021) as described in Section 3.3.

Table 6: Coefficient estimates for regressions using regressions using men's and women's employment as the outcome variable and heterogeneous distance bands

Sample VARIABLES	(1) All Employed	(2) Men	(3) Women	(4) All	(5) Men	(6) Women	(7) All	(8) Men	(9) Women
$Camp_{c,[0,5]} \times After_{ct}$	0.052**	0.064**	0.041*	0.058***	0.070***	0.047*	0.087***	0.081**	0.080**
	(0.021)	(0.026)	(0.024)	(0.021)	(0.027)	(0.025)	(0.026)	(0.033)	(0.031)
Homogen. T.E.?	NO	NO	NO	NO	NO	NO	NO	NO	NO
$Camp_{c,(5,10]} \times After_{ct}$	0.008	-0.004	0.012	0.014	0.003	0.018	0.046***	0.034	0.052**
	(0.014)	(0.021)	(0.017)	(0.015)	(0.022)	(0.018)	(0.018)	(0.030)	(0.021)
Homogen. T.E.?							NO		YES
$Camp_{c,(10,15]} \times After_{ct}$	0.018	-0.002	0.022*	0.024**	0.005	0.028**	0.040***	0.022	0.044***
	(0.012)	(0.018)	(0.013)	(0.012)	(0.019)	(0.014)	(0.015)	(0.026)	(0.017)
Homogen. T.E.?			NO	NO		YES	YES		YES
$Camp_{c,(15,20]} \times After_{ct}$	0.003	-0.030*	0.014	0.010	-0.022	0.021	0.021	0.002	0.031**
	(0.011)	(0.016)	(0.012)	(0.012)	(0.017)	(0.013)	(0.013)	(0.024)	(0.015)
Homogen. T.E.?		YES							YES
$Camp_{c,(20,25]} \times After_{ct}$				0.015	0.025*	0.011	0.019*	0.035**	0.016
				(0.010)	(0.015)	(0.011)	(0.012)	(0.018)	(0.013)
Homogen. T.E.?					NO		NO		
$Camp_{c,(25,30]} \times After_{ct}$				0.010	-0.002	0.015	0.016*	0.009	0.022*
				(0.008)	(0.012)	(0.010)	(0.010)	(0.014)	(0.012)
Homogen. T.E.?							NO	NO	YES
Sample Restricted to Those with Rohrer's Index	No	No	No	No	No	No	Yes	Yes	Yes
Dep. Var. Mean	0.738	0.814	0.708	0.738	0.814	0.708	0.724	0.854	0.699
Observations	415,076	116,023	299,051	415,076	116,023	299,051	209,849	33,827	176,022

R-squared	0.227	0.295	0.246	0.227	0.295	0.246	0.254	0.265	0.254

Note: Robust standard errors clustered at the DHS cluster level in parentheses.
Although not shown, all estimations include controls discussed related to equation (3) except for the seasonality fixed effects.
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Coefficient estimates for regressions using child anthropometric measures as outcome variables, 10 km treatment band, and accounting for the marital status of the child's household's head

VARIABLES	(1) WAZ	(2) HAZ	(3) WHZ
$Camp_{c,[0,10]} \times After_{ct} \times Married\ HH\ Head_i$	26.016***, †††	29.870***, †††	11.141
	(8.242)	(10.402)	(8.159)
	[0.006]	[0.006]	[0.172]
Homogen. T.E.?	YES	YES	
$Camp_{c,[0,10]} \times Married\ HH\ Head_i$	-12.381**	-11.022	-7.929
	(5.566)	(6.786)	(4.942)
$Camp_{c,[0,10]} \times After_{ct}$	-32.627***, †††	-35.040***, †††	-16.317**, ††
	(8.011)	(9.826)	(8.249)
	[0.000]	[0.000]	[0.048]
$Camp_{c,[0,10]}$	10.590**	14.718**	4.656
	(5.269)	(6.743)	(4.857)
Dep. Var. Mean	-108.393	-134.927	-30.271
Observations	97,240	97,240	98,977
R-squared	0.132	0.143	0.120

Note: Robust standard errors clustered at the DHS cluster level in parentheses
All regressions include controls as outlined in equation (7). Seasonality fixed effects are included. Their estimated coefficients are omitted.
*** p<0.01, ** p<0.05, * p<0.1
Benjamini-Hochberg q-values in brackets
††† q<0.01, †† q<0.05, † q<0.1

4.5 Discussion

The results consistently estimate negative nutritional outcomes for children and positive nutritional outcomes for adult men in close proximity to camps. We have included the Benjamini-Hochberg q-values to account for the possibility of Type I error, though this conservative estimation check renders some of our significant results insignificant or diminish confidence in the estimates. But even with the Benjamini-Hochberg correction, much of the statistical significance holds, including in our preferred specification outlined in equation (3). Moreover, our main results are robust to diagnostic testing defined in Jakiela (2021) that evaluates treatment effect homogeneity, which suggests that are findings are not the product of biased estimates under staggered treatment timing.

Why are nutritional outcomes falling for host children living near camps? Our results suggest children living closer to camps are more likely to be underweight after camp opening, but that this change in anthropometrics does not occur in the context of parental loss of employment or higher prevalence of general illness after camp opening. Instead, we believe that our results are driven by either highly localized demand shocks in the food market, selective migration, or a combination of these factors. Regarding the former, it is possible our results are due to demand-driven price shocks, as Chambers (1986) observed. But given that we only see an impact within only 10 or 15 km exposure buffers, the price shocks would have to be highly localized. Concerning the latter, the analysis provides evidence of inward migration to the area immediately around the camp. Our robustness checks suggest that the results for exposed children are worse when we omit those households that are moving towards the camps, meaning that in-migration is not driving down average nutritional status for children. It is also possible that the poor are less likely to migrate away from the camp after camp opening. By consequence, the community we observe after camp openings may be poorer on average, and consequently, their nutritional outcomes are worse on average. Such a result would suggest that camp openings do not actually *cause* host children to experience worsening nutritional status, but that the host children who live nearby after a camp opens tend to be poorer.

The result motivates policy consideration with respect to additional food assistance for the host households present after camp creation. Whether the mechanism is selective migration or localized food market disruptions, refugee camps lead to an agglomeration effect with a profound spatial reordering of the local economy. Policymakers should be aware that while servicing a camp and its residents, they are inadvertently inducing shifts in where people live, where they find income-generating opportunities, and what goods and prices they encounter in the market. Based on this, the analysis suggests future policy planning by host governments, the UNHCR, and its partner agencies carefully consider the changing demographic nature of the host community, as well as changes in the purchasing power of hosts. Our results suggest that by doing so, stakeholders will identify cases in which extending nutritional programming to host children may help prevent the malnourishment that our study identifies.

5. Conclusion

The organization of refugee populations into camps is a phenomenon that takes place in many low- and middle-income countries. To the best of our knowledge, this study is the first attempt to quantify the economic impacts of refugee camps on host populations on a scale beyond highly localized case studies. In the process of carrying out this research, we have developed a novel dataset on all the known locations of refugee camps in SSA between 1999 and 2016, which allows for a new level of broader continent-wide investigation of the effects of camps on hosts without sacrificing the advantage of previous work focused on examining highly localized effects. We draw on previous work on the proximity to resource extraction to examine the impacts of refugee camps on the nutritional outcomes in refugee camp host communities using a DID empirical strategy and investigate the mechanisms behind these results.

The results indicate that refugee camps drive a decrease in child WAZ and HAZ measures in host communities for the camps. For many host countries, including hosts in service provision in the proximity of refugee camps is an important component of policy planning, but this support often does not include food assistance for host households. Our results suggest that policymakers should carefully consider increasing food security programming for children in refugee camp host communities in their planning.[23] We argue that our outcomes are due to either local price changes or selective migration. The remaining ambiguity concerning migration in response to camps motivates future work determining the extent to which camp creation stimulates selective in- and out-migration among hosts.

Our newly created ARD allows for further exploration of the refugee phenomenon across SSA. This study does not explore outcomes by variations in the size of a camp in terms of land area or population counts, preferring to look at the dichotomous state of there being a camp in some proximity to an individual or not. Future research could take the size and scale of the camp into account to refine these estimates. Although only used tangentially in this paper, future studies of refugee impacts can also take advantage of the information in the ARD on refugee settlements outside of camps. There is reason to suspect that the impacts of these settlements are distinct from those associated with refugee camps since the infrastructure of the camp is not

[23] Key informant interviews with humanitarian stakeholders in East Africa in 2019 indicate that host households are generally excluded from food assistance programs but are included in education, health, and WASH programming. For more information on this qualitative work, please see Appendix Section A2.

present to segregate refugees from the local community. And because out-of-camp refugees generally do not receive as much (or any) humanitarian support, spillovers from the activities of international aid institutions are likely for hosts. Additional opportunities exist in employing remote sensing imagery on forest cover, cropping patterns, and nighttime lights to examine resource use changes and economic growth patterns related to the settlement of forcibly displaced people on a broad spatial scale.

6. Works Cited

Abadie, A. *et al.* (2017) 'When should you adjust standard errors for clustering?', *NBER Working Paper No. 24003*.

Aidoo, E. N., Amoh-Gyimah, R. and Ackaah, W. (2013) 'The effect of road and environmental characteristics on pedestrian hit-and-run accidents in Ghana', *Accident Analysis and Prevention*, 53.

Akresh, R., Verwimp, P. and Bundervoet, T. (2011) 'Civil war, crop failure, and child stunting in Rwanda', *Economic Development and Cultural Change*, 59(4).

Akresh, R. and de Walque, D. (2008) 'Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide', *World Bank Policy Research Working Paper*, 4606(April). doi: 10.1111/j.0042-7092.2007.00700.x.

Alix-Garcia, J. *et al.* (2018) 'Do refugee camps help or hurt hosts? The case of Kakuma, Kenya', *Journal of Development Economics*, 130. doi: 10.1016/j.jdeveco.2017.09.005.

Alix-Garcia, J. and Saah, D. (2010) 'The effect of refugee inflows on host communities: Evidence from Tanzania', *World Bank Economic Review*, 24(1). doi: 10.1093/wber/lhp014.

Alloush, M., Taylor, J. E. and Gupta, A. (2017) 'Economic Life in Refugee Camps', *World Development*, 95, pp. 334–347. doi: 10.1016/j.worlddev.2017.02.030.

Anti, S. (2021) 'Land grabs and labor in Cambodia', *Journal of Development Economics*, 149(January).

Anti, S., Salemi, C. and Wilson, K. (2020) 'Introducing the African Refugee Dataset 1999-2016'.

Arndt, C. *et al.* (2016) 'Economics and Human Biology Effects of food price shocks on child malnutrition : The Mozambican experience 2008 / 2009', *Economics and Human Biology*, 22(September 2008), pp. 1–13.

Baez, J. E. (2011) 'Civil wars beyond their borders: The human capital and health consequences of hosting refugees', *Journal of Development Economics*, 96(2), pp. 391–408. doi: 10.1016/j.jdeveco.2010.08.011.

Baker, A., Larcker, D. F. and Wang, C. C. Y. (2021) 'How Much Should We Trust Staggered Difference-In-Differences Estimates?', *ECGI Finance Working Paper No. 736*. doi: 10.2139/ssrn.3794018.

Benjamini, Y. and Hochberg, Y. (1995) 'Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing', *Journal of the Royal Statistical Society, Series B (Methodological)*, 57(1).

Benjamini, Y. and Yekutieli, D. (2001) 'The control of the false discovery rate in multiple testing under dependency', *The Annals of Statistics*, 29(4).

Benshaul-Tolonen, A. (2018) 'Local Industrial Shocks and Infant Mortality', *The Economic Journal*, pp. 1–32. doi: 10.1111/ecoj.12625.

Berazneva, J. and Byker, T. S. (2017) 'Does forest loss increase human disease? Evidence from Nigeria', *American Economic Review Papers and Proceedings*, 107.

Betts, A. *et al.* (2017) *Refugee economies: Forced displacement and development*. Oxford University Press.

Borusyak, K. and Jaravel, X. (2018) 'with an Application to the Estimation of the Marginal Propensity to Consume ∗', pp. 1–35.

De Brauw, A. and Harigaya, T. (2007) 'Seasonal Migration and Improving Living Standards in Vietnam', *American Journal of Agricultural Economics*, 89(2).

Bryan, G., Chowdhury, S. and Mobarak, A. M. (2014) 'Underinvestment in a profitable technology: the case of seasonal migration in Bangladesh', *Econometrica*, 82(5).
Bunte, J. B. *et al.* (2018) 'Natural resource sector FDI, government policy , and economic growth : Quasi-experimental evidence from Liberia', *World Development*, 107. doi: 10.1016/j.worlddev.2018.02.034.
Callamard, A. (1994) 'Refugees and Local Hosts: A Study of the Trading Interactions between Mozambican Refugees and Malawian Villagers in the District of Mwanza *', *Journal of Refugee Studies*, 7(1).
Callaway, B. and Sant'Anna, P. H. C. (2020) 'Difference-in-Differences with multiple time periods', *Journal of Econometrics*, (xxxx), pp. 1–31. doi: 10.1016/j.jeconom.2020.12.001.
Carr-Hill, R. (2013) 'Missing Millions and Measuring Development Progress', *World Development*, 46. doi: 10.1016/j.worlddev.2012.12.017.
de Chaisemartin, C. and D'Haultfœuille, X. (2020) 'Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects', *American Economic Review*, 110(9), pp. 2964–2996. doi: 10.1257/aer.20181169.
Chambers, R. (1986) 'Hidden losers? The impact of rural refugee programs on poorer hosts', *The International Migration Review*, 20(2).
Chen, J., Kosec, K. and Mueller, V. (2019) 'Moving to despair? Migration and well-being in Pakistan', *World Development*, 113.
Chen, J. and Mueller, V. (2018) 'Coastal climate change , soil salinity and human migration in Bangladesh', *Nature Climate Change*, 8(November).
D'Souza, A. *et al.* (2014) 'Food Insecurity in Vulnerable Populations: Coping with Food Price Shocks in Afghanistan', *American Journal of Agricultural Economics*, 96(3).
Dillon, A., Mueller, V. and Salau, S. (2011) 'Migratory Responses to Agricultural Risk in Northern Nigeria', *American Journal of Agricultural Economics*, 93(4).
Fink, G., Mcconnell, M. and Vollmer, S. (2014) 'Testing for heterogeneous treatment effects in experimental data : false discovery risks and correction procedures', *Journal of Development Effectiveness*, 6(1).
Fisk, K. (2014) *Geo-Refugee: A refugee location dataset, Harvard Dataverse Network*.
Gardner, J. (2021) 'Two-stage differences in differences', (April).
Genovese, C. and Wasserman, L. (2002) 'Operating characteristics and extensions of the false discovery rate procedure', *Journal of the Royal Statistical Society, Series B (Methodological)*, 64(3).
German, B., Gomez Canon, C. and Mueller, V. (2019) 'Spillover effects of the Venezuelan crisis: migration impacts in Colombia'.
von der Goltz, J. and Barnwal, P. (2019) 'Mines: The local wealth and health effects of mineral mining in developing countries', *Journal of Development Economics*, 139. doi: 10.1016/j.jdeveco.2018.05.005.
Goodman-Bacon, A. (2021) 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics*, (xxxx). doi: 10.1016/j.jeconom.2021.03.014.
HarvestChoice (2015) *Travel time to nearest town over 20k (mean, hours, 2000), International Food Polocy Research Institute and University of Minnesota*.
ICF (no date) 'Demographic and Health Surveys (various). Funded by USAID'. Rockville, Maryland: ICF.
INDEPTH (2008) *Anthropometric status indicators: INDEPTH Resource kit for demographic surveillance systems*.

Jakiela, P. (2021) ‘Simple Diagnostics for Two-Way Fixed Effects’. Available at: http://arxiv.org/abs/2103.13229.
Knutsen, C. H. *et al.* (2016) ‘Mining and local corruption in Africa’, *American Journal of Political Science*, 61(2).
Kotsadam, A. *et al.* (2018) ‘Development aid and infant mortality. Micro-level evidence from Nigeria’, *World Development*, 105.
Kotsadam, A. and Tolonen, A. (2016) ‘African Mining, Gender, and Local Employment’, *World Development*, 83(June 2013), pp. 325–339. doi: 10.1016/j.worlddev.2016.01.007.
Kreibaum, M. (2016) ‘Their Suffering , Our Burden? How Congolese Refugees Affect the Ugandan Population’, *World Development*, 78, pp. 262–287. doi: 10.1016/j.worlddev.2015.10.019.
Kung, K. S. *et al.* (2014) ‘Exploring Universal Patterns in Human Home-Work Commuting from Mobile Phone Data’, *PLOS One*, 9(6).
Lee, Y. N. (2017) ‘Does Commodity Price Volatility Drive Migration? Evidence from Rural Does Commodity Price Volatility Drive Migration ? Evidence from Rural Ethiopia’, *Working paper*.
Maystadt, J.-F. and Verwimp, P. (2014) ‘Winners and Losers among a Refugee-Hosting Population’, *Economic Development and Cultural Change*, 62(4), pp. 769–809. doi: 10.1086/676458.
Minoiu, C. and Shemyakina, O. N. (2014) ‘Armed conflict, household victimization, and child health in CÔte d’Ivoire’, *Journal of Development Economics*, 108, pp. 237–255. doi: 10.1016/j.jdeveco.2014.03.003.
Montclos, M.-A. P. De and Kagwanja, P. M. (2000) ‘Refugee Camps or Cities? The Socio-economic Dynamics of the Dadaab and Kakuma Camps in Northern Kenya’, *Journal of Refugee Studies*, 13(2).
Mueller, V., Gray, C. and Hopping, D. (2020) ‘Climate-Induced migration and unemployment in middle-income Africa’, *Global Environmental Change*, 65.
Oka, R. (2011) ‘Unlikely Cities in the Desert: The Informal Economy as Causal Agent for Permanent “Urban” Sustainability in Kakuma Refugee Camp, Kenya’, *Urban Anthropology and Studies of Cultural Systems and World Economic Development*, 40(3/4).
Oka, R. C. (2014) ‘Coping with the Refugee Wait: The Role of Consumption, Normalcy, and Dignity in Refugee Lives at Kakuma Refugee Camp, Kenya’, *American Anthropologist*, 116(1), pp. 23–37. doi: 10.1111/aman.12076.
Perez-Heydrich, C. *et al.* (2013) ‘Guidelines on the use of DHS GPS data’, *DHS Spatial Analysis Reports 8*.
Pullum, T. *et al.* (2017) *Hemoglobin data in DHS surveys: Intrinsic Variation and Measurement Error*. Rockville, Maryland, USA.
Quisumbing, A. R. (2003) ‘Food Aid and Child Nutrition in Rural Ethiopia’, 31(7), pp. 1309–1324. doi: 10.1016/S0305-750X(03)00067-6.
Roth, J. (2019) ‘Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends’, *Working Paper*, (April 2018), pp. 1–81.
Roth, J. and Sant’Anna, P. H. C. (2021) ‘Efficient Estimation for Staggered Rollout Designs’. Available at: http://arxiv.org/abs/2102.01291.
Ruiz, I. and Vargas-silva, C. (2016) ‘The labour market consequences of hosting refugees’, *Journal of Economic Geography*, 16. doi: 10.1093/jeg/lbv019.
Ruiz, I. and Vargas-Silva, C. (2013) ‘The Economics of Forced Migration’, *Journal of*

Development Studies, 49(6). doi: 10.1080/00220388.2013.777707.
Ruiz, I. and Vargas-Silva, C. (2015) 'The Labor Market Impacts of Forced Migration', *American Economic Review Papers and Proceedings*, 105(5), pp. 581–586.
Salemi, C. (2021) 'Refugee camps and deforestation in Sub-Saharan Africa', *Journal of Development Economics,* 152. doi: 10.1016/j.jdeveco.2021.102682.
Sun, L. and Abraham, S. (2020) 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics*, (xxxx), pp. 1–25. doi: 10.1016/j.jeconom.2020.09.006.
Tatah, L. *et al*. (2016) 'Impact of Refugees on Local Health Systems: A Difference-in-Differences Analysis in Cameroon', *PLOS One*.
UNHCR (2018) *'wrl ppl poc p unhcr' - People of Concern Map, ArcGIS Hub*.
UNHCR (no date) *UNHCR Statistical Yearbooks*.
Yamano, T., Alderman, H. and Christiaensen, L. (2005) 'Child growth, shocks, and food aid in rural Ethiopia', *American Journal of Agricultural Economics*, 87(2).
Zezza, A. *et al*. (2011) 'Assessing the impact of migration on food and nutrition security', *Food Policy*, 36(1).
Zhou, Y.-Y. and Shaver, A. (2021) 'Reexamining the Effect of Refugees on Civil Conflict: a Global Subnational Analysis', *American Political Science Review*.