Refugee Camps and Deforestation in Sub-Saharan Africa

Revised draft

April 24th, 2021

Abstract: To date, there have been few quasi-experimental efforts to evaluate the impact of refugee camps on host landscapes. Yet many stakeholders believe refugee camps lead to deforestation in nearby areas. I use data on camp locations and years of operation as well as secondary geospatial data to produce a high-resolution panel dataset of 0.01° tiles. My difference-in-difference specification with tile fixed effects exploits variation in camp openings and tile proximity to camps. F-tests on event study pre-trends provide support for the satisfaction of parallel trends prior to camp exposure. I find that within the rainforest biome, camps are associated with a small *reduction* in extensive margin forest loss (i.e., land clearing) and a small increase in intensive margin forest loss (i.e., gradual reductions in canopy cover). In the grasslands biome, camps lead to small increases in forest loss at the intensive margin but have no impact on the extensive margin.

JEL: Q23, Q34, O13,

Keywords: Africa, refugee camps, forced displacement, deforestation

1. Introduction

In 2019, there were an estimated 29.4 million refugees in the world, and one-quarter of this population resided in sub-Saharan Africa (SSA) (UNHCR, 2019). As of 2013, about 40% of refugees in SSA were required to live in planned refugee camps, where they receive humanitarian assistance (Verwimp and Maystadt, 2015). Academic attention towards refugees has increased considerably in recent years, with many scholars seeking to understand whether refugee camps or out-of-camp refugee populations have any impact on the hosting country (Maystadt *et al.*, 2019; Verme and Schuettler, 2021). Although the impact of refugee populations on host communities has been a common topic of inquiry, their impact on host landscapes remains understudied. Despite limited evidence, many stakeholders in refugee hosting believe that refugee camp residents fuel deforestation in the area around the camp. For example, UNHCR documentation claims that refugee reliance on firewood is "a main driver of forest degradation and deforestation" in displacement settings (UNHCR and FAO, 2018).

This paper takes a spatially explicit, quasi-experimental approach to study deforestation in response to camp openings in sub-Saharan Africa.[1] My objective is to evaluate whether tree canopy

[1] I use the Food and Agriculture Organization's definition of deforestation and define the phenomenon as "the conversion of forest to another land use or the long-term reduction of the tree canopy cover below the minimum 10% threshold." This is not the same as forest degradation, which the FAO characterizes as "changes within the forest which negatively affect the structure or function of the stand or site, and thereby lower the capacity to supply products and/or services" (FAO, 2000).

cover declines in response to camp openings and to determine the distances from the camps at which these losses manifest.

Past economic research suggests that camps could stimulate deforestation through several channels. Population-driven productivity growth (Alix-Garcia *et al.*, 2018) may result in an upward movement along the environmental Kuznets curve (Foster and Rosenzweig, 2003), especially if higher food demand leads to land clearing for agricultural expansion. The resulting environmental damages may be highly concentrated around camp areas in the absence of sufficient transportation infrastructure (Alix-Garcia *et al.*, 2013). But population growth does not always result in forest loss (Cropper and Griffiths, 1994), as these processes will vary based on differences in country policy (Scrieciu, 2007).

Moreover, demand for fuelwood in the camp may raise the returns to harvesting and selling firewood (for hosts and refugees), resulting in more forest products extracted. Evidence suggests that households are more likely to harvest and sell firewood when (1) forests are far enough from the village to drive up the market price, and (2) a market is accessible to the household (Albers and Robinson, 2013; Miteva *et al.*, 2017; Bošković *et al.*, 2018). Since there are markets within each refugee camp (Betts *et al.*, 2017), these results suggest that income-generating forest extraction could manifest several kilometers from the camp itself. Income transfers may reduce demand for forest products in the event that substitution between forest and non-forest goods is feasible (Ferraro and Simonrangkir, 2020). But refugees face severe income constraints, and substitutes for cooking fuel, such as liquid propane gas cookstoves, may not be available in the markets they can access.

My analysis focuses exclusively on the impacts of planned refugee camps, and the findings are not generalizable to other types of refugee locations, such as settlements, transit centers, etc. Countries in my sample also vary with respect to government hosting policies, with some placing strict limitations on refugee labor and mobility outside of camps (Blair, Grossman and Weinstein., 2020). These policies may have important implications for environmental outcomes, but in the absence of available data on country encampment policies, examining this important heterogeneity is beyond the scope of the present study.

The effect of refugee population influxes and refugee camps on host landscapes remains an understudied topic in the economics literature. The growing body of quasi-experimental studies that have explored the impacts of refugee influxes in SSA tend to examine changes for host communities, focusing on outcomes such as employment, wages, health, consumption, wellbeing, prices, or nutrition (see for example Alix-Garcia and Saah, 2010; Baez, 2011; Alix-Garcia, Bartlett and Saah, 2013; Maystadt and Verwimp, 2014; Ruiz and Vargas-Silva, 2015, 2016; Kreibaum, 2016; Alix-Garcia *et al.*, 2018; Maystadt and Duranton, 2019; Maystadt *et al.*, 2019).

To my knowledge, Maystadt et al. (2020) is the only previous quasi-experimental study to examine the impacts of camps on the landscape, and they also focus on SSA. Their analysis uses an instrumental variable approach that regresses land cover changes on predicted refugee camp population size and other covariates. They find that increases in the refugee camp population within a 111 by 111 km area results in a small *increase* in vegetation density for that area, but they also find that tiles with larger camp populations were more likely to experience both forest loss and agricultural expansion between 2001 and 2012. The relatively coarse resolution of analysis in Maystadt *et al.* (2020) might obscure significant relationships between camp proximity and deforestation. I therefore extend their analysis by processing all spatial data at a much finer spatial resolution (1.1 km by 1.1 km*)* so that I can examine *where*, in relationship to the camp, forest losses occur. I also build on this previous work by removing areas within the camps from the analysis and by dropping refugee settlements from the sample. By doing so, I can more accurately measure how a planned refugee camp, as a spatial unit, impacts the land cover around it.

I use a dataset of refugee camp locations and years of operation across the subcontinent as well as numerous spatially explicit, gridded secondary datasets. I generate 499,586 sample tiles at 0.01° resolution selected from areas near camps. For each sample tile, I calculate the zonal statistics for each of the spatially explicit dependent variables, along with the annual number of camps present at different distances from each tile. I estimate a difference-in-difference specification that exploits variation in camp openings and closings, as well as variation in tile proximity to camps, to observe the magnitude of camp-attributed forest losses and the distances from the camps these losses manifest. To ensure the comparability of areas with similar ecological endowments, I estimate my regressions separately for two different biome groups: grasslands ($N = 359{,}284$) and rainforests ($N = 112{,}134$).

I find evidence of statistically significant but extremely small changes to forest cover in response to camp exposure. Regression outcomes suggest that camps are not statistically associated with extensive margin forest losses in grasslands, and in rainforests, I find statistically significant, but modest, reductions in extensive margin forest loss in response to camp exposure. In both biomes, camps appear to drive statistically significant, but very small, reductions in forest cover along the intensive margin. These results are robust to several checks. I also estimate first differences over relative time, and the outcomes suggest that camp impacts may take up to a decade to manifest.

The remainder of this paper is structured as follows. Section 2 discusses the data sources used in my analysis and provides descriptive statistics on the geographic and economic characteristics of the camp locations. Section 3 outlines the analytical approach, and in Section 4 I report the results. I conduct robustness checks and describe the outcomes in Section 5. My concluding marks appear in Section 6.

2. Data

2.1 Datasets

I use the African Refugee Dataset (ARD, see Anti, Salemi, and Wilson 2020), which identifies the locations of refugee camps in sub-Saharan Africa that operated between 1999 and 2016. The data includes camp name, year of creation, year of dissolution (if relevant), and the geographic coordinates of the location. This data is based on various UNHCR sources that were cross-checked. The ARD covers 424 refugee camps active between 1999 and 2016, which are distributed across 35 countries. Data collection, cleaning, and validation procedures are all recorded in Anti, Salemi, and Wilson (2020).

The ARD dataset may be more accurate than the UNHCR's internal data, which Maystadt *et al.* (2020) use and which is not publicly available. UNHCR's internal data informs the Statistical Yearbook and People of Concern Maps that the ARD is built on. Cross-checks of these sources against other secondary accounts revealed frequent data entry errors, such as the misclassification of non-camp refugee locations as camps or incorrect data on the first year of camp operation. We can see these errors in the mapping of camp locations in Maystadt et al. (2020). For example, that paper reports several refugee camps in Libya, but since there are no planned refugee camps in Libya, I suspect these points actually represent detention centers.[2] Based on the aforementioned errors, I believe that ARD is a relatively more accurate account of camp locations and years of operation.

For this study, I restrict the ARD sample to only camps operating for at least one year between 2001 and 2012 with no missing information and drop Ugandan refugee settlements. This is another difference between my study and Maystadt *et al.* (2020), which keeps the refugee settlements in Uganda in the analysis, even though their geographic size and land use policies differ considerably from planned refugee camps.[3] The resulting sample is made up of 300 camps. The majority (35%) of these camps are located in east Africa, with sizable shares in central (33%) and west (25%) Africa as well. Very few (7%) of the camps were in southern Africa.

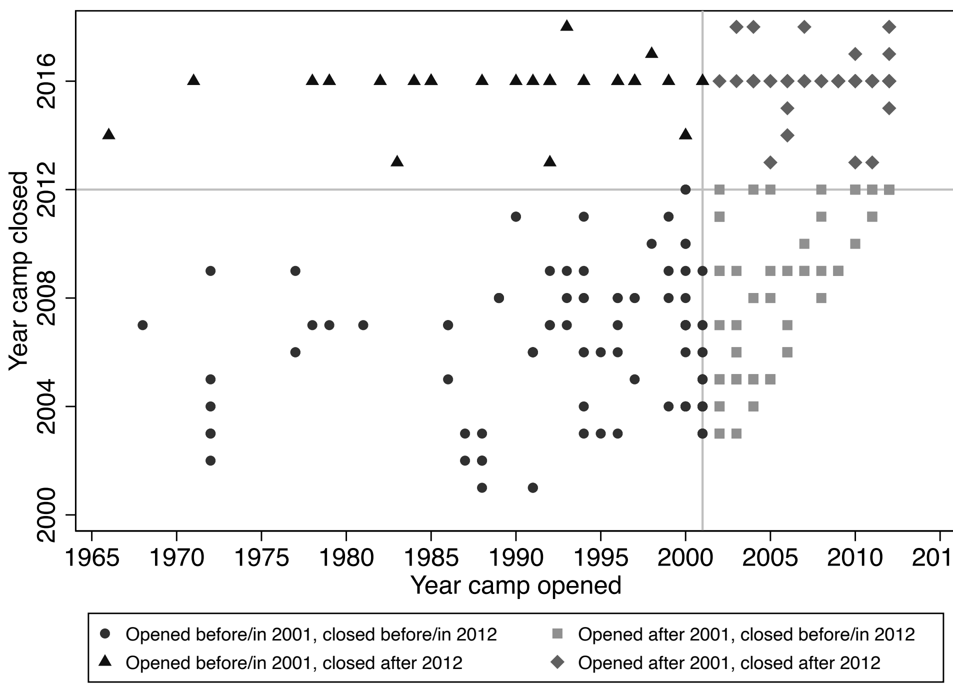
Figure 1illustrates the opening and closing years of all camps in the study and classifies camps into four categories based on their years of operation. The oldest camps are in the upper-left-hand corner of the graph, and the youngest in the lower right-hand corner. About one-third (33.7%) of the camps opened and closed during the study period. One-quarter (27%) opened before 2001 and closed during the study period. Additionally, one-quarter of the camps (28%) opened after 2001 but

[2] This map is included in the supplemental information document of

[3] Following the ARD's data validation protocol, refugee settlements in Uganda were included in the ARD data (see Anti, Salemi and Wilson (2020)). But these settlements are not comparable to camps because they are significantly larger. Some of the settlements are as large as 50 km in length. Moreover, settlement residents are allocated land to cultivate, which will result in different land clearing patterns as compared to planned camps.

closed in the years after the study period. For a small number (11%) of cases, the camps were open the entire study period. Figure 1 also shows that some camps in the sample were only open for a short period of time. Indeed, one-fifth (19%) of sample camps were only active for two years (see Appendix A2). Camps that operated for a short period of time may not trigger as much forest loss as camps that operated for years, which motivates one of my robustness checks (Section 3.4).

Figure 1: camp opening and closing years, ARD camps operating 2001-2012



Source: author's calculations using ARD data. The vertical reference line is at 2001, the first year of the study period. The horizontal reference line is at 2012, the final year of the study period.

Throughout this paper, I distinguish between extensive margin deforestation and intensive margin deforestation. The former refers to the case where an area transitions from non-zero forest cover to zero forest cover in a given year. The latter refers to gradual reductions in the area's percent forest cover over time. For cleaned and consistent geospatial data on extensive margin deforestation, I use the Global Forest Change (GFC) data for the years 2000-2012 (Hansen *et al.*, 2013). GFC grid-cells have a high resolution of 30 meters, and the data only classify forest cover as vegetation growth of at least five meters in height. The data consist of a 2000 baseline measure of forest coverage and indicate the year in which a grid-cell transitioned to zero forest cover, if relevant. The GFC is widely used in economic studies of deforestation,[4] but the data have some flaws. For example, Tropek *et al.* (2014) show that the GFC often misclassify plantations for products such as oil palm and rubber as natural forests.

Another weakness of the GFC is the data's high threshold for change. Because the GFC can only indicate when a 30-meter grid-cell transitions to zero forest cover, smaller-scale losses are not detected (Burivalova *et al.*, 2015). To examine land degradation at the intensive margin, I use the Global Forest Cover Change (GFCC) data produced by NASA (Sexton *et al.*, 2013). The GFCC is at the same resolution as the GFC (30 meters) and also uses the 5-meter tree growth criteria to classify forests. At five-year intervals (2000 to 2015), the GFCC data reports each grid cell's percentage forest cover. Hence, the GFCC can capture marginal reductions in percent forest cover that do not result in the grid cell transitioning to zero forest cover.

I additionally use RESOLVE's 2017 ecoregions data (Dinerstein *et al.*, 2017). I use this data

[4] see for example: Alix-Garcia, Sims and Yañez-Pagans (2015), Berazneva and Byker (2017) and Abman (2018).

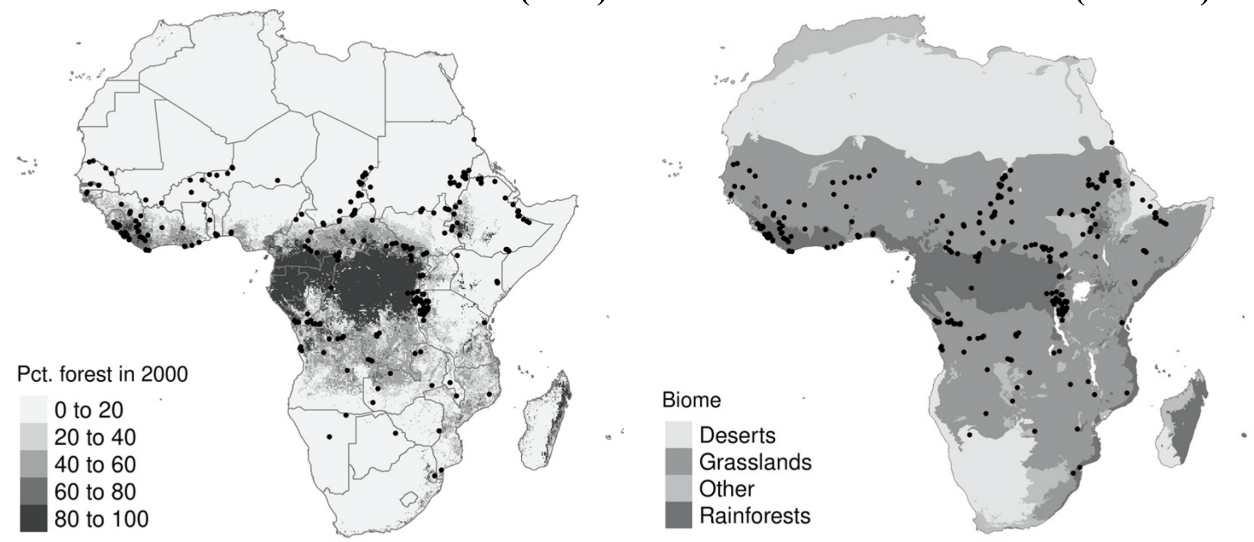
to separately examine outcomes for two biomes of interest:[5] tropical and subtropical moist broadleaf forests ("rainforests" hereafter) and tropical and subtropical grasslands, savannahs, and shrublands ("grasslands" hereafter). I disaggregate the analysis by biome following the assumption that any camp-stimulated changes to the landscape are likely a function of baseline ecological characteristics, which are a function of bioclimatic attributes. Consequently, if outcomes vary by biome but the specification does not account for biome, then the estimates may fail to accurately convey the impacts of camp exposure.

Figure 2 maps GFC 2000 forest cover data and REACH biome data for the continent and illustrates the spatial distribution of the ARD camps used in this study. The camps are often located very close to the border that refugees crossed to enter the host country. Panel A shows that while some of these camps are in the more densely forested regions of central and west Africa, many are also located in regions with very little forest cover in 2000. Panel B shows that camps predominantly fall into the grasslands and rainforest biomes.

[5] Biomes are large areas of earth characterized by similar climate and ecological conditions with specific communities of flora and fauna that thrive in these conditions. There are seven terrestrial biomes across the planet.

Figure 2: camp locations mapped onto spatial data used for the study

Panel A: Percent forest cover in 2000 (GFC) **Panel B: Biomes of Africa (REACH)**



Source: author’s calculations using ARD data for camps operating 2001-2012, GFC and RESOLVE data. Camp locations displayed as black points. Panel A displays percent forest cover in 2000 based on the GFC (aggregated at a 1 km resolution). Panel B shows the two biomes used in the study (grasslands/savannahs and rainforests) as well as areas characterized as desert biome. “Other” includes flooded grasslands and savannahs, dry broadleaf forests, mangroves, and montane grasslands/shrublands: it is very uncommon to find ARD camps in these biomes.

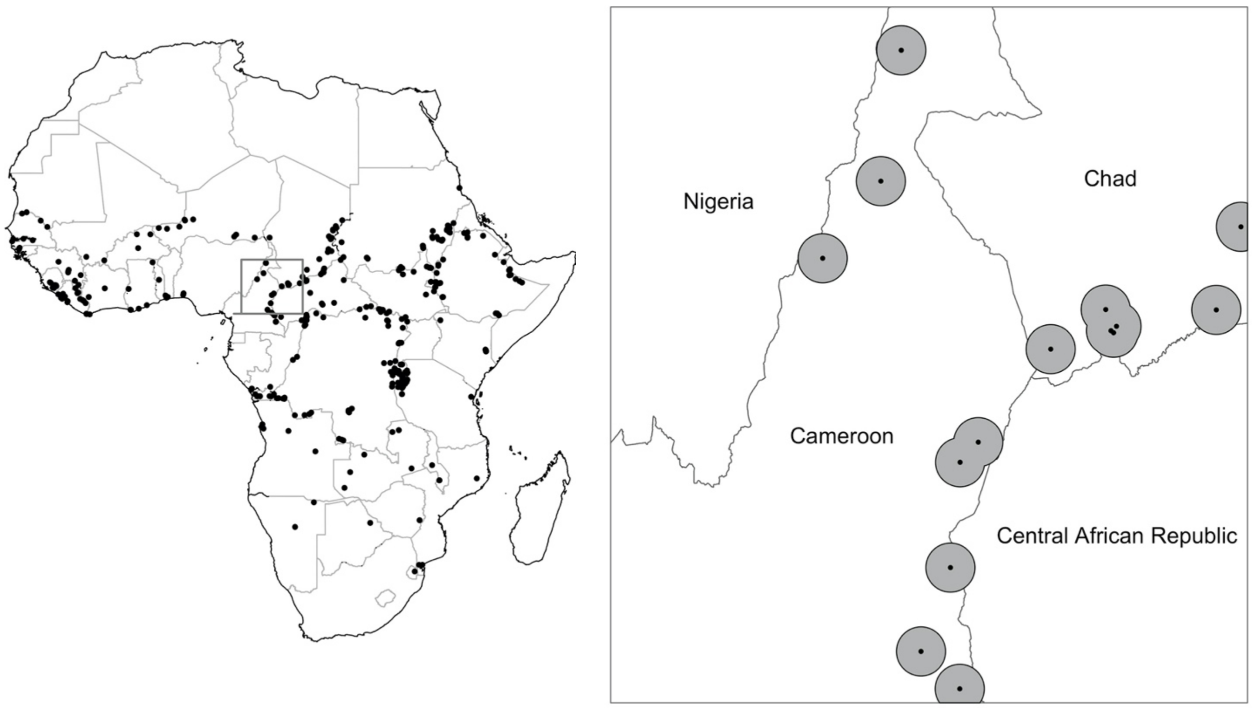
2.2 Data preparation

Figure 3 illustrates the sampling technique I use in this study. Based on the geographic locations of camps (Panel A), I determine the 30 km buffer areas around each camp, shown in Panel B. I merge overlapping buffer areas together to prevent over-counting. Within these buffer areas, I generate 0.01° tiles, which are roughly 1.1 km by 1.1 km each (Panel C). I use the GFC, GFCC, and RESOLVE data to determine zonal statistics for each tile. For example, Panel D shows the 30-meter GFC grid-cell data on 2000 percent forest cover over one sample tile. I calculate zonal statistics using grid-cell data that intersects the tile. I also define each tile's centroid and use this point to produce annual estimates of the number of camps 1, 2, …, 20 km from the tile every year 1996-2012, based on Euclidean distance.[6]

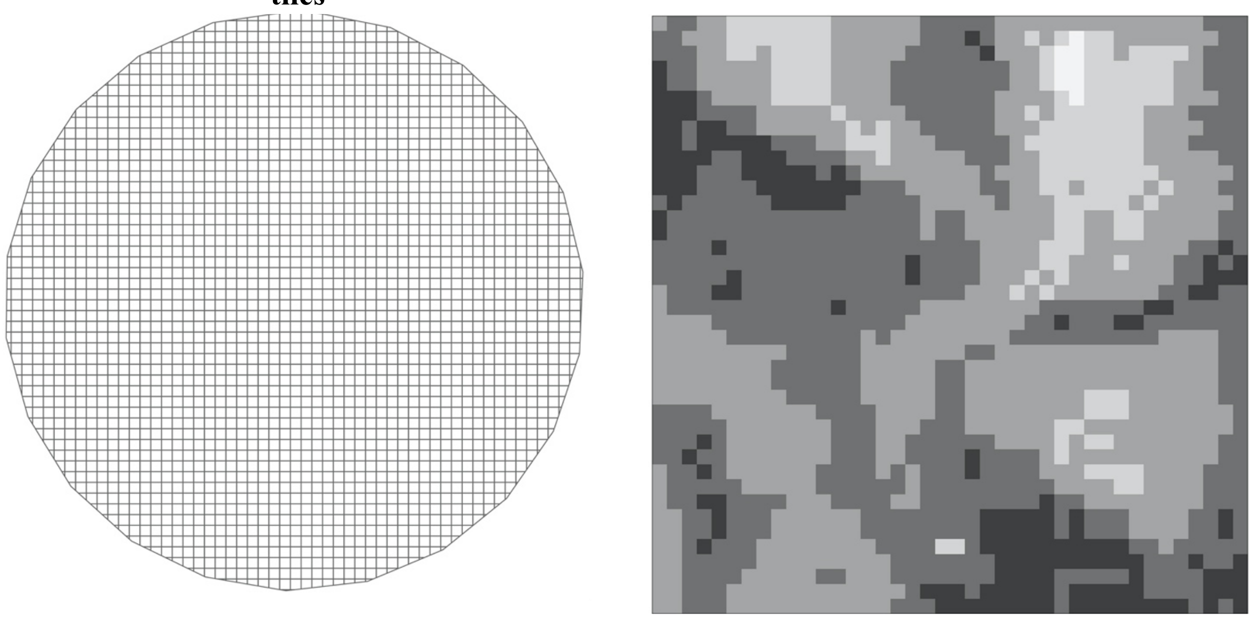
[6] As I describe in Section 3, I use camp proximity data 1996-2000 in order to determine the number of years a tile was exposed to a camp over the previous five years. I use this an explanatory variable in my DID regression with percent forest cover (GFCC) serving as the outcome variable.

Figure 3: Illustration of sampling strategy used for this study

Panel A: Geographic locations of camps **Panel B: 30 km buffers, west-central Africa**



Panel C: 30 km buffer divided into 0.01° tiles **Panel D: 30m GFC grid-cells within 0.01° tile**



Source: author's calculations using ARD 2001-2012 and GFC 2000 baseline forest cover data. Panel A shows all camps open at some point 2001-2012 as dots and locates the area of the continent shown in Panel 2 in a grey box. Panel B illustrates 30 km buffer areas around camps in west-central Africa. Panel C shows the 0.01° tiles within one of these buffers. Panel D illustrates higher-resolution grid cells within one sample tile, which are used to calculate zonal statistics.

If the land allocated for a camp has vegetation cover, land clearing may be obligatory in order to make room for housing and necessary infrastructure. In order to focus on how camps impact the landscapes *around* them, I drop all observations that are less than 1 km from the nearest camp at any point, as in this case, the camp and the tile likely overlap. The assumption that camps tend to have a radius of 1 km is justifiable,[7] though there are some camps that are larger in surface area. I address the possibility that camps are larger than assumed in a robustness check (Section 3.4).

To ensure that tiles are large enough to be comparable regardless of their relative position within a buffer area, I drop all tiles from my 0.01° sample with fewer than 1,000 30-meter grid-cells within them. In doing so, I omit tiles around the periphery of the buffer that are considerably smaller than other tiles (See Figure 3 Panel C). I also drop all observations from Liberia, where early results revealed considerable noise in the years leading up to camp openings.[8]

2.3 Descriptive statistics

Table 1 provides descriptive statistics: I report means for the entire sample and for sub-

[7] To my knowledge, there is no publicly available information on the average geographic size of refugee camps in SSA. But UNHCR protocols can provide an idea of what a suggested camp size would be for a certain size population. UNHCR recommends a camp size of 45 square meters per resident. Assume this standard is maintained and consider a refugee camp with a radius of 1 kilometer (and an area of roughly 3,141,592 square meters). By the UNHCR standard, such a camp area could accommodate roughly 70,000 residents. My sense is that many camps in SSA have populations this size or smaller, though there are some exceptions. For example, there are approximately 150,000 refugees living in Nyarugusu Camp in Tanzania, which is one of the largest camps on the continent.

[8] For more discussion on the omission of Liberia from the study, see Appendix A1.

samples defined by the minimum distance between each tile and the nearest camp open at any point 2001-2012. About half (56.4%) of sample tiles were not within 20 km of any refugee camps at any point in the study period: these tiles serve as a comparison group in my econometric specification. Over one-third of the tiles were within 20 km of one camp at some point over the study period (34.5%), though a small share (9.2%) were within 20 km of two or more camps.

Tiles 1-10 km from the nearest camp had slightly lower forest cover in 2000 compared to tiles further away based on the GFC and GFCC data. These relationships hold even when restricting to only tiles that had not yet been exposed to a camp as of 2000 (Appendix A2). This slightly lower baseline forest cover and forest quality cover may be due to the following factors. First, when looking for land to build a camp on, host governments tend to prefer land available at the lowest cost. Landowners may not be incentivized to allocate land with valuable natural capital, such as more forested land, and may instead allocate land that is already cleared, has poorer soil quality, etc. Second, in order to minimize the costs of preparing the area for settlement, stakeholders may try to select plots that are already relatively cleared. Third, in countries that have experienced numerous waves of refugee arrivals over time, camps may sit dormant for years and come back into use when needed. These observations may appear in the ARD as "new" camps, even though they were first built years prior.[9]

[9] For example, the Tanzanian government re-opened Nduta camp for Burundian refugees in 2015 (Oxfam International, 2015).

Both the GFC and GFCC data provide evidence of modest reductions in forest cover over the study period. Table 1 shows the average number of 30-meter GFC grid-cells that transitioned to zero forest cover between 2001 and 2012 for sample tiles. On average, tiles lost 1.8 30-meter grid-cells, less than one percent (0.13%) of the tile's surface area. It appears that tiles closer to camps lost slightly more 30-meter grid-cells than those further from camps. The naïve estimates suggest that over the study period (2001-2012) tiles 1-5 km and 6-10 from camps experienced an additional 720 and 360 square meters of forest loss (respectively), relative to comparison tiles 21-30 km away. These differences are not especially large (they are smaller in surface area than a baseball diamond). The GFCC data suggests that on average, only tiles 1-5 km from a camp (at some point over the study period) experienced intensive margin forest loss 2000-2010, but this reduction is very small, 0.02 percentage points (ppt).

Table 1: Descriptive statistics

	Full sample	Nearest camp 1-5 km away	Nearest camp 6-10 km away	Nearest camp 11-15 km away	Nearest camp 16-20 km away	Nearest camp 21+ km away
Pct. forest cover '00 (GFC)	26.08	23.58	25.32	25.73	26.22	26.38
Num. 30m grid-cells transition zero forest cover '01-'12 (GFC)	1.82	2.51	2.10	1.92	1.79	1.71
Mean forest cover '00 (GFCC)	15.13	13.53	14.47	14.63	15.00	15.49
Ppt. forest cover change '00-'10 (GFCC)	0.18	-0.02	0.03	0.09	0.16	0.24
Pct. with 1 camp 1-20 km away	34.47	60.86	67.83	76.63	89.14	0.00
Pct. with 2+ camps 1-20 km away	9.18	39.14	32.17	23.37	10.86	0.00
Pct. in grassland biome	74.90	77.85	76.72	76.80	75.74	73.74
Pct. in rainforest biome	19.35	19.13	19.71	18.97	19.26	19.43
Pct. in other biome	5.75	3.02	3.56	4.23	5.00	6.84
N	479,686	15,117	36,915	50,727	85,265	270,316
Number of buffers	149	149	149	149	149	149

Source: author's calculations using ARD, GFC, GFCC, and RESOLVE, data. Columns are organized based on the distance between the tile's centroid and the nearest camp open at any point between 2001 and 2012.

3. Method

The difference-in-difference econometric approach used in this study exploits variation in the timing of camp openings and closings to determine camp impacts. The sample tiles are much smaller than those specified in Maystadt et al. (2020), facilitating a spatially explicit examination of whether camps influence their surrounding areas.

3.1 Endogenous camp site selection

Refugee camp locations are not exogenous and will likely have characteristics that are different from other nearby areas. Information gathered from key informant interviews suggests that camp site selection operates on the basis of several rather time-invariant factors.[10] In the event that a new camp is required, government stakeholders generally aim to acquire land in the area close to the refugee point of border entry. The objective of this strategy is to minimize transportation costs, increase the likelihood of voluntary return in the future, minimize security risks in major urban areas, and maximize cultural proximity between refugees and hosts. Given a preferred region for the camp, the host government then seeks to obtain a land allocation of sufficient size. In some cases, they re-designate government land for the camp, which often means allocating parcels of forest reserves, government ranches, or former military sites. In many other cases, host governments negotiate with communal or individual landholders, seeking to obtain plots on the basis of "good will." Such agreements are most attractive to villages that

[10] My understanding of the political economy of camp site selection comes from fieldwork conducted in July and August 2019 in Kenya, Rwanda, Uganda, and Tanzania. During this time, I spoke with government officials, UNHCR site planners, FAO specialists, and other nonprofit stakeholders to learn more about the decision-making process behind refugee camp land designation. This fieldwork was funded by the University of [XXX]. For a full description of this work, see the Appendix section of [XXX and XXX (XXXX)].

both have unused land to offer and see the benefits of hosting as outweighing the costs.[11] Whether the land is heavily forested or not will depend on the site's tenure regime and the characteristics of the land available in the area.[12]

The factors that determine camp selection—proximity to the border with the refugee-sending country and the availability of free or low-cost land—are relatively time-invariant. But this selection process may mean that land allocated for camps is often not desired by others. It is possible that the land's undesirability stems from ecological characteristics that are also related its forest cover (Maystadt *et al.*, 2020).

3.2 Primary estimation approach

To estimate the impact of camp exposure on extensive margin forest loss, I perform the following regression:

$$y_{it} = \alpha + \sum_{d=1}^{D} \alpha_d camps_{it}^{d} + \gamma_i + \Phi + e_{ibt} \tag{1}$$

I define y_{it} as the number of 30-meter grid-cells within 0.01° sample tile i that transitioned to zero forest cover in year t according to the GFC for years 2001-2012.[13] The explanatory variables of interest indicate the number of camps that are in distance bin d, where $d \in D$ and

[11] That is, the value of the services that the camp-related humanitarian assistance brings to the area – in terms of improved schools, healthcare, safe water access, etc. – outweighs the losses from donating land indefinitely. The net value is positive for these communities partly because they were peripherally located and poor prior to the camp's establishment.
[12] Generally speaking, when there is an abundance of unused land in a fertile region, camps will be within close proximity to forest resources. But if unused land in the region tends to be the least suitable for vegetation, then refugees may not be in close proximity to forest resources.
[13] The outcome variable is measured in terms of 30-meter grid-cells lost, instead of percent of grid-cells lost, in order to obtain coefficients that do not need to be re-scaled. Using the latter measure, I obtain the same sign and significance for my estimates, but the coefficients are extremely small.

$D = \{(1,5], (5,10], (10,15], (15,20]\}$ kilometers. For example, the variable $camps_{it}^{[1,5]}$ reports the number of camps that were 1-5 km from the tile's centroid in a given year. The tile fixed effect γ_i controls for numerous time-invariant characteristics of the tile that are likely endogenous to refugee camp site selection and land cover changes, such as soil fertility, terrain slope and elevation, baseline forest cover, agroecological zone, land tenure status, and time-invariant heterogeneity across countries and regions. The sampling approach my lead to serial autocorrelation for tiles sampled from the same buffer. Following Cameron and Miller (2014) I cluster my standard errors at the buffer level to address serial autocorrelation arising from the grouped structure of the data. The clustering of standard errors is also important because of the DID framework: Bertrand, Duflo and Mullainathan (2004) argue that naïve DID specifications under-state the standard deviation of the estimated treatment impacts and that clustering standard errors is one way to avoid mischaracterizing statistical significance.

The coefficient ϕ represents one of two approaches to control for time. In the first approach, $\Phi = \rho_t$, a time fixed effect that captures time-varying, location-invariant factors relevant to camp openings and forest change. But recent scholarship has shown that two-way fixed effects can lead to biased estimates in the event that the timing of treatment varies across observations and treatment impacts are heterogeneous. An important identifying assumption for this two-way fixed effects specification is that each "cohort"[14] of exposure tiles is exposed to the same treatment, a planned refugee camp. Because of potential bias associated with two-way fixed effects, I also estimate Equation 1 where Φ is a linear trend.

Without more information on the populations residing in this area, I cannot determine whether any changes can absolutely be attributed to the activities of hosts versus refugees. But

[14] "Cohort" refers to groups of tiles exposed to camps in the same year.

because the explanatory variables are based on the distance to camps each year, the coefficient estimates will reveal whether or not changes in forest cover manifest at distances from camps that are accessible to refugees. Qualitative accounts suggest that some refugees allocate several hours per day to fetching firewood (Whitaker, 2002; Mulumba, 2011; Rivoal and Haselip, 2017). But to my knowledge, there have been no studies that track the distances people traverse as they extract forest products. A back-of-the-envelope calculation leads me to suspect that any direct refugee forest extraction will likely take place 10 kilometers or less from the camp boundary.[15] Since my distance measures are based on the camp centroid, and since some camps may be larger than 2 km in diameter, I also consider the 11-15 km area as a potential site of refugee-driven forest change. If the results indicate forest losses only within this 1-15 km distance, that will support the theory that camp residents are driving land cover changes.

I estimate Equation 1 separately for tiles in the grasslands biome and tiles in the rainforest ecoregion. If forest loss was driven by large numbers of refugees extracting firewood for consumption, then one would expect that grasslands, with their very limited tree cover, would quickly transition to zero forest cover near camps. The impact at the extensive margin in rainforests is initially ambiguous: refugee foraging may not be sufficient to lead to full land clearing, but the population influx could influence the marginal returns to harvesting from forests and may also stimulate agricultural expansion that requires land clearing.

I estimate the impact of camp exposure on intensive margin forest loss using the GFCC

[15] Suppose it takes on average 15-20 minutes to walk 1 km over natural terrain (when factoring in the weight of firewood for at least half the trip). If a refugee traveled in a straight line away from the camp for 10 km fetching firewood, this would require about 5-6 hours of roundtrip walking. In reality, it seems unlikely that someone would follow such a direct path and continue to move without taking any breaks. This simple calculation leads me to suspect that any forest losses beyond 10 km are not likely attributable to camp residents.

with the following regression:

$$y_{it} = \eta_0 + \sum_{d=1}^{D} \eta_d camp.years_{it}^{d} + \kappa_i + \Phi \; + \epsilon_{ibt} \tag{2}$$

Here, y_{it} is the percent of forest cover over grid cell i in year t based on the GFCC. The GFCC only includes data at five-year intervals, so $t = 2000, 2005, 2010$. Because I do not have annual GFCC data, the variable of interest $camp.years_{i\tau}^{d}$ represents how many years (over the previous five years) the tile was exposed to *at least* one camp at distance d. For example, $campyears_{i,2005}^{[1,5]} = 2$ means that between 2001 and 2005, there was at least one camp 1-5 km away from tile i for two years. Like Equation 1, Equation 2 includes a tile fixed effect, robust standard errors clustered at the buffer level, and time controls Φ. Again, I estimate this specification for tiles in grasslands and rainforest biomes separately.

One limitation of the estimation approach is that it does not account for variations in camp sizes over time. Camp population data in secondary sources is largely incomplete, and Maystadt *et al.* (2020) had to drop many camps from their sample for their estimations because of insufficient population data. The identification strategy rests on the assumption that camp population sizes are generally constant over time, meaning camps do not oscillate between very large and very small populations.

3.3 Evaluation of pre-trends

In the identification strategy, I assume that tiles over 20 km from camps, as well as tiles not yet exposed to a camp, serve as a valid counterfactual. The true counterfactual is unknowable, but an evaluation of trends prior to camp exposure can provide support for the comparison group as a valid proxy counterfactual. Given the staggered timing of camp exposure, I use an event study to evaluate pre-trends. The specification is as follows:

$$y_{it} = \beta_0 + \sum_{d=1}^{D} \sum_{\tau=-\infty}^{\infty} \beta^{d\tau}(time_{it}^{d,\tau}) + \lambda_i + \Phi + \varepsilon_{ibt} \tag{3}$$

Here, $treated_i$ is a binary indicator equal to 1 if t corresponds with relative exposure year τ for *at least* one camp at distance d from tile i. For example, $time_{ti}^{[1,5],-2} = 1$ means that in year t, tile i is not yet exposed to a camp 1-5 km away, but a camp will open at this distance after two years. Tiles 21-30 km from a camp serve as the comparison group, and the period right before camp opening serves as the placebo "pre-treatment" period. I perform this event study using GFC (extensive margin) and GFCC (intensive margin) outcome variables. When using the GFC, I organize relative time into one-year bins. Because I have fewer years of GFCC data, I organize relative time into two-year bins.

Null estimates of $\beta^{d\tau}$ when $\tau < 0$ suggest that trends prior to camp openings are parallel. But scholars have recently begun to show that this traditional event study evaluation of pre-trends may lead to Type II errors, with researchers inferring zero pre-trends despite the presence of a nonzero pre-trend (Borusyak and Jaravel, 2017; Kahn-Lang and Lang, 2018; Roth, 2020). To more rigorously test for the presence of nonzero pre-trends,[16] I follow Borusyak and Jaravel

[16] The topic of DID pre-trends represents an active frontier in economics. Recent working papers have also examined methods to account for non-zero pre-trends in DID frameworks. For example, Freyaldenhoven, Hansen and Shapiro (2019) present an instrumental variables approach using some observable that proxies the confound driving the divergence in trends pre-treatment. Rambachan and Roth (2020) recently developed a technique that only requires that the

(2017) and perform an F-test comparing the unrestricted event study to a restricted specification. The authors recommend dropping two pre-treatment variables for an event study with one treatment group. Because I estimate the impact for four treatment groups simultaneously, I omit eight relative time variables at $\tau = -1$ and $\tau = -4$ when regressing extensive margin forest loss on relative time. For the intensive margin event study I omit eight relative time variables at $\tau = [-3, -2]$ and $\tau = [-9, -8]$. If the restricted and unrestricted specifications are not statistically different, this lends additional support to the assumption that the difference in trends prior to camp opening is statistically zero.

3.4 Robustness checks

As mentioned, some of the camps are only active for a few years, which may not be enough time for them to impact the landscape. Consequently, I repeat my estimates restricting the sample to only tiles around camps that were active for more than two years.

Moreover, the impacts on the landscape may only manifest after a number of years of camp exposure. To examine average impacts by year of exposure, I use a first difference specification to estimate differences in forest cover each relative year. This specification closely resembles the event study (Equation 3), but it does not include a placebo pre-treatment year.

Additionally, it is possible that camp-stimulated increases in firewood demand could lead to harvesting in areas with the highest marginal value (i.e., denser forests). Consequently, I estimate the impact of camp exposure on forest cover change at the intensive and extensive margin using a sub-sample of tiles with 50 percent forest cover or greater in 2000 (according to

researcher impose restrictions on the difference in pre-trends to estimate "honest" DID coefficient estimates.

the GFC data).

My identification strategy largely rests on the assumption that camps tend to be about 2 km in diameter. While this makes sense for many camps, there are some large camps that exceed this size. If the camp is larger than expected, then the comparison group may be contaminated, as refugees may not be as far from the area 21 km from the camp centroid as expected. To ensure that the results are not driven by contamination of the comparison tiles, I build a new sample consisting of 596,722 0.01° tiles extracted from 30 km buffer areas on the subcontinent that are outside of the study areas used for the primary analysis.[17]

For this final robustness test, I drop all tiles from my primary data that are 21-30 km from the nearest camp (the comparison group in my main analysis). I then randomly select 5% from the new sample of comparison tiles, append these to my primary dataset, and estimate the main specifications. I do this 100 times for each specification and evaluate the distribution of coefficient estimates relative to the main results.

4. Results

4.1 Pre-trends testing

The event study results are reported in Appendix A4. Across outcome variables and biomes, the estimates are statistically insignificant over pre-treatment years leading up to camp exposure for grasslands and rainforest tiles. Table 2 reports the outcomes of the F-test on the pre-treatment trends. The sample of tiles from the grasslands biome performs well across all specifications and outcome variables, with relatively low F-values ($p \geq 0.27$). For rainforest tiles, the results offer the strongest support for zero pre-trends for the two-way fixed effects specifications when the

[17]For more information on this sampling strategy, see Appendix A3.

extensive margin forest loss variable is the dependent variable. Rainforest tiles perform poorly in the pre-trends test when percent forest cover serves as the dependent variable, which could be related to the low number of observations over relative time when restricting to the rainforest biome and only three years of data.

Table 2: Results of F-test of pre-trends in event study specifications

Dep. var. - number of 30m grid cells that transition to zero forest cover (GFC)				
	Grasslands		**Rainforests**	
	F-value	**P-value**	**F-value**	**P-value**
Twoway fixed effects	1.2521	0.2749	1.2048	0.3191
Linear trend	1.0674	0.3903	1.7918	0.1052
Dep. var. - percent forest cover (GFCC)				
	Grasslands		**Rainforests**	
	F-value	**P-value**	**F-value**	**P-value**
Twoway fixed effects	0.9313	0.4933	6.2826	0.0000
Linear trend	0.7860	0.6158	11.6054	0.0000

Notes: table reports results of pre-trends test proposed by Borusyak and Jaravel (2017). To perform the test, I estimated Equation 3, then dropped eight pre-treatment variables and tested to see if the specification fit of the restricted regression was significantly different than that of the unrestricted regression. Using the GFC outcomes variable, I drop explanatory variables at relative time $\tau = -1$ and $\tau = -4$. When the GFCC data provides the outcome variable, there are fewer years of data, so relative time is organized into two-year bins. I drop $\tau = [-3, -2]$ and $\tau = [-9, -8]$ for the F-test. Regressions use sample of tiles measured at 0.01° resolution.

4.2 Main results

With respect to extensive margin forest loss (Table 3), my results suggest that camp exposure does not lead to economically or statistically significant impacts when the camp is located in the grasslands biome. The results suggest that in rainforests, camp exposure results in a small reduction in extensive margin forest loss. Each year of camp exposure is associated with a small fraction of one 30-meter grid-cell within the 0.01° tile *not* transitioning to zero forest cover. This avoided extensive margin forest loss manifests at all distances 1-20 km from the camp. But the magnitudes are small. For example, the result of the two-way fixed effects model for tiles 1-5 km from a camp suggest that camp exposure results in approximately 128 square meters[18] of avoided forest loss per year for that tile. To put this result in perspective, that surface area is roughly the size of a medium-sized swimming pool.

Table 3: Coefficient estimates for Equation 1 with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable

	Grasslands		Rainforests	
	(1)	(2)	(3)	(4)
N camps 1-5 km	0.011	0.010	-0.142***	-0.163***
	(0.007)	(0.007)	(0.017)	(0.017)
N camps 6-10 km	0.002	0.001	-0.067***	-0.088***
	(0.003)	(0.003)	(0.009)	(0.009)
N camps 10-15 km	0.004	0.003	-0.102***	-0.126***
	(0.003)	(0.003)	(0.007)	(0.007)
N camps 15-20 km	-0.005**	-0.006***	-0.120***	-0.146***
	(0.002)	(0.002)	(0.005)	(0.005)
Two-way FE?	Yes	No	Yes	No
Linear trend?	No	Yes	No	Yes
Obs.	4,343,568	4,343,568	1,349,436	1,349,436
N. buffers	124	124	44	44
R^2	0.212	0.211	0.197	0.190

Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level and are reported in parentheses. *p<0.05 ** p<0.01 *** p<0.001

At the intensive margin, I find evidence of statistically significant, but modest, reductions in tile percent forest cover in response to camp exposure (Table 4). In the grasslands biome, the impacts at the intensive margin are negative and significant, but small: for example, one year of

[18] Each 30-meter grid-cell has a surface area of 900 square meters. 14.2% of this surface area is roughly 127.8 square meters.

exposure to at least one camp results in the reduction in the tile's tree canopy cover by 0.13-0.15 ppt. for tiles 1-5 km from the camp. This represents a 1.2-1.4 percent reduction of the sample mean in 2000. Intensive margin losses in rainforests are slightly larger in magnitude and statistically significant: among rainforest tiles 1-5 km from a camp, one year of camp exposure is associated with a reduction in tile forest cover by 0.21-0.23 ppt., a 0.6 percent reduction relative to the 2000 sample mean. But the results of the pre-trends analysis suggest that these estimates may be biased.

Table 4: Coefficient estimates for Regression 2 with tile percent forest cover (based on GFCC data) as outcome variable

	Grasslands		**Rainforests**	
	(1)	(2)	(3)	(4)
Yrs. exposure to camp 1-5 km	-0.133***	-0.154***	-0.216***	-0.226***
	(0.008)	(0.008)	(0.023)	(0.023)
Yrs. exposure to camp 6-10 km	-0.091***	-0.114***	-0.213***	-0.224***
	(0.005)	(0.005)	(0.013)	(0.013)
Yrs. exposure to camp 11-15 km	-0.115***	-0.140***	-0.187***	-0.199***
	(0.004)	(0.004)	(0.010)	(0.010)
Yrs. exposure to camp 16-20 km	-0.117***	-0.144***	-0.176***	-0.189***
	(0.003)	(0.003)	(0.008)	(0.008)
Two-way FE?	Yes	No	Yes	No
Linear trend?	No	Yes	No	Yes
GFCC mean 2000	10.82	10.82	34.13	34.13
Obs.	1,077,852	1,077,852	336,402	336,402
N. buffers	124	124	44	44
R^2	0.983	0.983	0.983	0.983

Source: author's calculations based on the ARD, GFCC, and RESOLVE data. Regressions use sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level and are reported in parentheses. *p<0.05 ** p<0.01 *** p<0.001

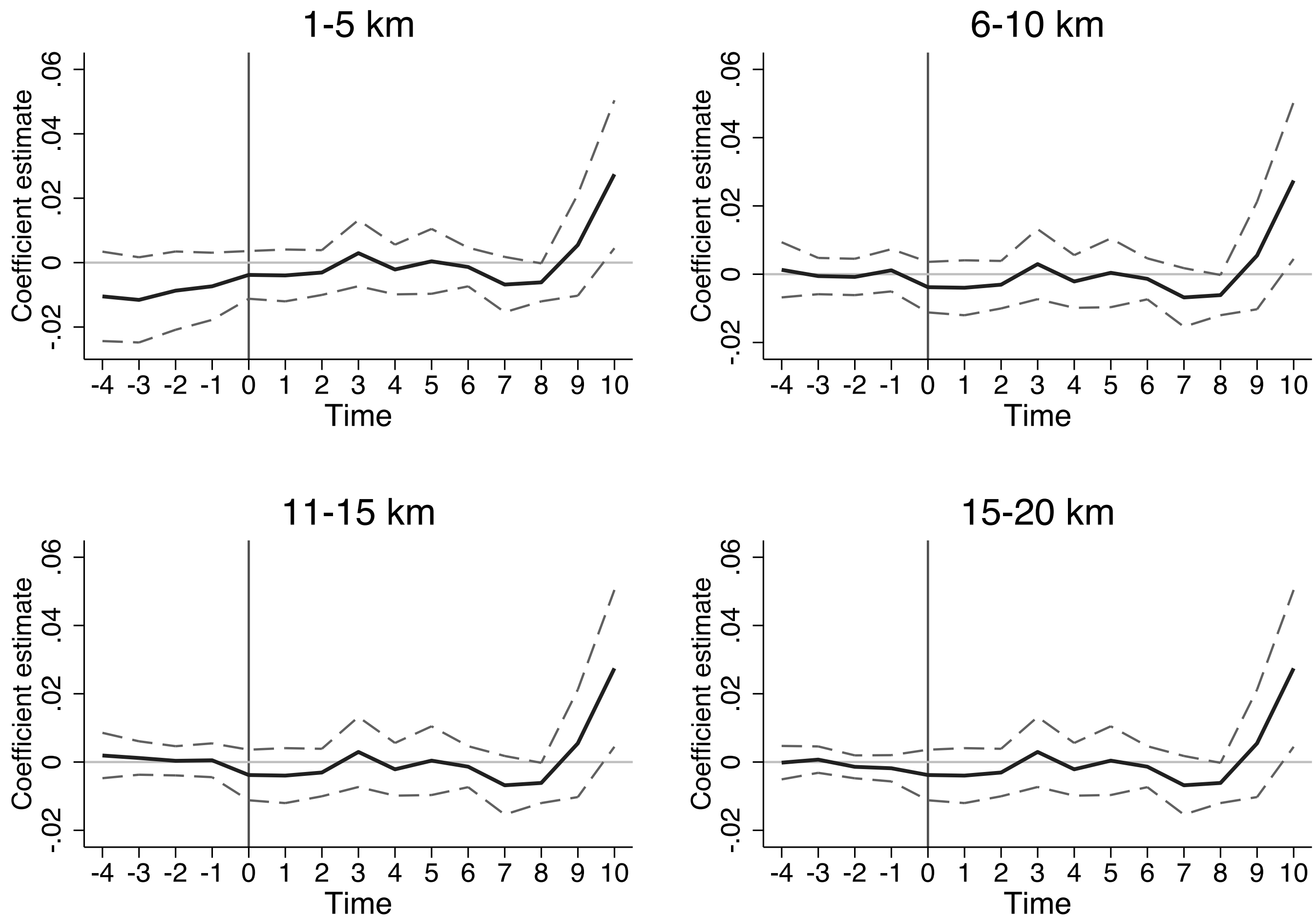
Perhaps refugees living in the camp are driving these intensive-margin losses through their forest harvesting activities. But impacts manifest at distances so far from camps (15-20 km) that refugee encroachment seems unlikely. These results highlight the complexity of camp-associated forest loss. It is very possible that the activities of camp residents lead to some of these modest intensive margin losses. But the losses 15 km from the camp or further suggest that the activities of other groups may also change in response to camp openings in a manner that influences forest cover.

4.3 Robustness checks

In Appendix A5, I report the results of the robustness check in which I drop any tiles that were exposed to camps for two years or fewer. The extensive margin outcomes are the same for grasslands, but the rainforest estimates become weaker in terms of statistical significance. Intensive margin results for grasslands are similar to the main results with respect to coefficient magnitudes, though these estimates have lower statistical significance than the main results. The rainforest estimates for intensive margin change are slightly larger in magnitude than the main estimates, though they also tend to exhibit lower statistical significance.

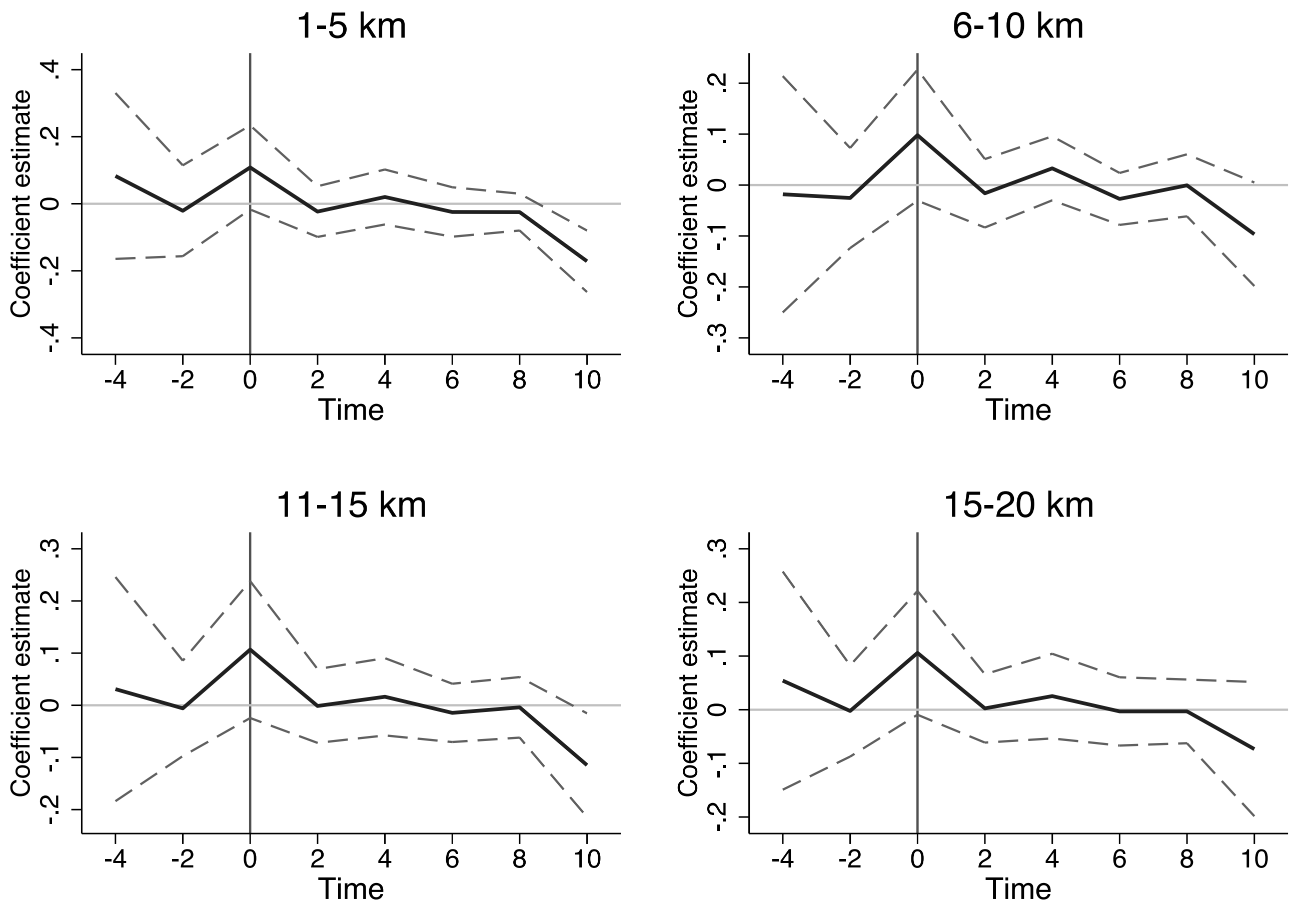
The first difference results (using two-way fixed effects) are shown in Figure 4 and Figure 5 for grasslands (graphs for rainforests reported in Appendix A6). The extensive margin first difference trend exhibits a spike in losses manifesting after ten years for grasslands (Figure 4). But the magnitude of the difference at $\tau = 10$ is very small. Likewise, the intensive margin results suggest that losses manifest about a decade after camp opening (Figure 5). The trends for rainforests closely resemble the trends for grasslands. These results suggest that camp-stimulated changes may take considerable time to manifest. My estimates may not capture high levels of forest loss because it takes so long for the changes to occur and because my study period consists of only twelve years.

Figure 4: Coefficient estimates for first difference specification with number of 30-meter grid-cells in the tile that transitioned to zero forest cover (based on GFC data) as outcome variable, tiles in grasslands biome



Source: author’s calculations based on the ARD, GFC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level. 95% confidence intervals displayed with dashed lines.

Figure 5: Coefficient estimates for first difference specification with percent forest cover (based on GFCC data) as outcome variable, tiles in grasslands biome

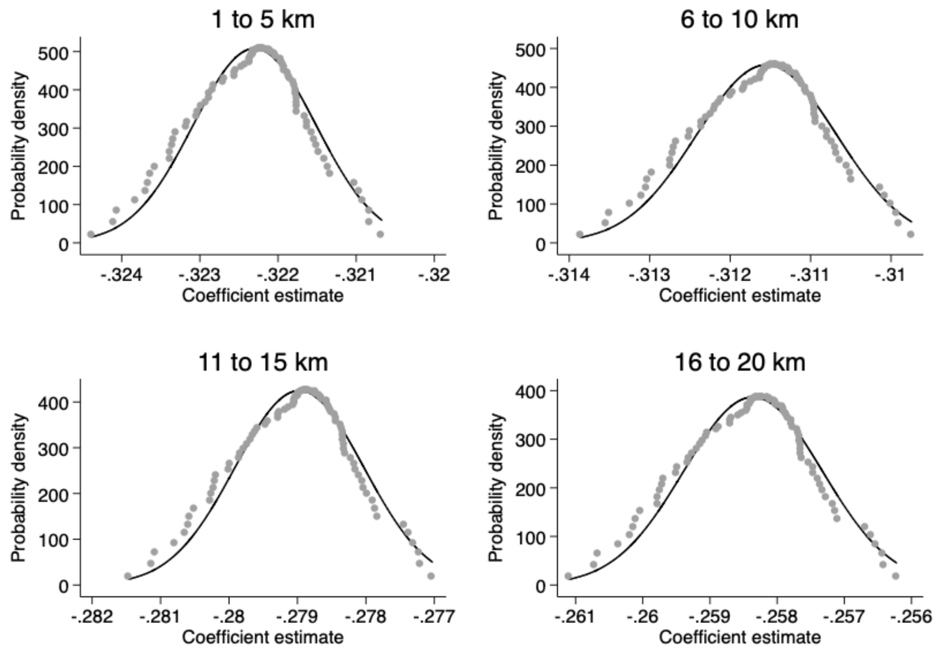


Source: author's calculations based on the ARD, GFCC, and RESOLVE data. Regression uses sample of tiles measured at 0.01° resolution. Robust standard errors are clustered at the buffer level. 95% confidence intervals displayed with dashed lines

Appendix A7 reports the results of the placebo test in which I estimate the main results using only sample tiles with 50 percent forest cover or more in 2000. The extensive margin results of this robustness check closely resemble the estimates in Table 3. Intensive margin estimates are higher, but mean forest cover at baseline is also higher, and as in the main results, the estimates represent an 0.8-1.4 percent reduction relative to the sample mean for each year of refugee camp exposure.

Appendix A8 reports the coefficient estimates collected from iteratively drawing a random subset of the new comparison group sample and estimating the main results. Across almost all specifications and outcome variables, the results of this robustness check closely mirror the main results. This suggests that the 21-30 km area from camps served as a valid comparison area, given the similar performance of the two comparison groups. The one exception is with intensive margin changes in rainforests, shown in Figure 6: the coefficient magnitudes are slightly larger than those derived in the primary analysis (Table 4 4), though they still constitute reductions of less than one ppt. per year.

Figure 6: Distribution of coefficient estimates for Equation 2 with percent forest cover as the outcome variable (based on GFCC data) using comparison tiles from buffers with centroids 100 km from the nearest camp, tiles in rainforest



Source: author's calculations based on the ARD, GFC, and RESOLVE data. Regressions use primary sample of tiles measured at 0.01° resolution and drop all tiles 21-30 km from a refugee camp. Comparison tiles are randomly drawn from 30 km buffers with centroids at least 100 km from an ARD camp location. Regression is estimated 100 times to obtain distribution of coefficient estimates. Robust standard errors are clustered at the buffer level.

4.4 Discussion

The statistically significant association between camps and avoided extensive margin forest losses in rainforests may correspond with the increase in vegetation density observed by Maystadt *et al.* (2020). In both cases, these gains are very small in magnitude and may not be economically significant. But it remains possible that under certain circumstances, camps may reduce the returns to harvesting forest products, either because they impact the agricultural wage (and by extension, the opportunity costs of harvesting from forests) or because an augmented security presence around camps increases the risks of illegal logging. Testing these theories would require more information about agricultural prices and the spatial distribution of illegal forest harvesting activities.

Building on the findings in Maystadt *et al.* (2020), I show that camps are associated with small intensive margin forest cover losses. On average, I find that camp exposure leads to annual forest cover losses amounting to less than one percentage point. These losses typically represent a 0.6-1.4 percent loss relative to the sample mean. Perhaps refugee forest extraction explains a considerable share of these modest losses. But I find significant impacts 15-20 km from camps, distances that refugees are unlikely to reach. And given the small magnitudes of the changes, the economic significance of these findings is questionable. These results challenge the preconception that refugees fuel extensive deforestation in camp areas.

The modest evidence of forest loss stands in contrast to the main finding in Maystadt et al. (2020): they argue that increases in a tile's refugee camp population greatly augments the likelihood that the sample tile experiences both a reduction in forest cover and an increase in crop cover over time. The discrepancy is likely due to the different approaches of the two studies. As mentioned, the underlying camp data differs, and this could explain difference to some extent. But without access to UNHCR internal data, I cannot test for this directly. The different spatial resolutions and econometric approaches of our analyses likely contribute to discrepancies in our findings. Maystadt *et al.* (2020) use very large spatial tiles (111x111 km) to estimate impacts, meaning they may pick up spurious changes unrelated to camps that coincide with their growth. Moreover, some of the forest loss that they estimate may be due to the clearing of the camp area itself, since their study does not try to omit camp areas. My study attempts to omit these areas and removes refugee settlements entirely to ensure that the estimates reflect the impact of camps on their surrounding landscape.

It remains possible that camp populations degrade the landscape in other ways that do not lead to deforestation. Refugee extraction of forest products such as firewood may not require the felling of living trees, but more likely involves the collection of already dead wood. The resulting clearing of the canopy floor may not impact how the canopy appears in remote sensing imagery, but it will impact biodiversity as habitats become increasingly degraded.

5. Conclusion

This study examined the impact of refugee camps on the landscapes surrounding them. Exploiting variation in camp operational years and defining camp exposure based on Euclidean distance to the camp, my difference-in-difference estimates suggest that camps have a very small negative impact on intensive margin forest cover 1-20 km from camps. Future work can expand on this finding using a spatially explicit examination of intermediate outcomes driven by camp exposure, such as host population growth. Such intermediate outcomes may help explain why camps lead to modest forest losses at such distances from the camp.

The current study was unable to explore heterogeneity by country-level policies that influence refugee mobility in and out of camps and refugees' rights to work. Our understanding of how camps influence the environment would be enhanced by estimations that account for country policies. This would also provide additional information on whether illiberal hosting policies actually protect landscapes.

Refugees often serve as scapegoats for host governments and host communities (Baylouny, 2020). Pejorative misnomers about refugee impacts have the potential to fuel conflict between refugees and hosts (Martin, 2005), and they also provide bad faith host governments with pretexts to deny entry to asylum-seekers (Black, 1998). Future studies of refugee impacts on the environment are needed, as this work provides policy-relevant evidence to the stakeholders who seek to protect both refugees and the environment. Moreover, such research performs the important function of testing beliefs that too often make some of the world's most vulnerable people even worse off.

References

Abman, R. (2018) 'Rule of law and avoided deforestation from protected areas', *Ecological Economics*, 146.

Albers, H. J. and Robinson, E. J. Z. (2013) 'A review of the spatial economics of non-timber forest product extraction : Implications for policy', *Ecological Economics*. Elsevier B.V., 92, pp. 87–95. doi: 10.1016/j.ecolecon.2012.01.021.

Alix-Garcia, J. *et al.* (2013) 'The ecological footprint of poverty alleviation: evidence from Mexico's Opportunidades program', *The Review of Economics and Statistics*, 95(2).

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

Alix-Garcia, J., Bartlett, A. and Saah, D. (2013) 'The landscape of conflict: IDPs, aid and land-use change in Darfur', *Journal of Economic Geography*, 13, pp. 589–617.

Alix-Garcia, J. M., Sims, K. R. and Yañez-Pagans, P. (2015) 'Only one tree from each seed? Environmental effectiveness and poverty alleviation in Mexico's Payments for Ecosystem Services Program', *American Economic Journal: Economic Policy*, 7(4).

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

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

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

Baylouny, A. M. (2020) *When blame backfires: Syrian refugees and citizen grievances in Jordan and Lebanon*. Cornell University Press.

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

Bertrand, M., Duflo, E. and Mullainathan, S. (2004) 'How much should we trust differences-in-differences estimates?', *The Quarterly Journal of Economics*, 119(1).

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

Black, R. (1998) *Refugees, environment, and development*. Essex, United Kingdom: Addison Wesley Longman Limited.

Blair, C. W., Grossman, G. and Weinstein., J. (2020) 'Introducing the Developing World Refugee and Asylum Policy Dataset'. Copenhagen.

Borusyak, K. and Jaravel, X. (2017) 'Revisiting event study designs, with an application to the estimation of the marginal propensity to consume'.

Bošković, B. *et al.* (2018) 'The effect of forest access on the market for fuelwood in India', *Fondazione eni Enrico Mattei: Working paper*.

Burivalova, Z. *et al.* (2015) 'Relevance of global forest change data set to local conservation: Case study of forest degradation in Masoala National Park, Madagascar', *Biotropica*, 47(2).
Cameron, A. C. and Miller, D. L. (2014) 'A practitioner's guide to cluster-robust inference', *Journal of Human Resources*, 50(2).

Cropper, M. and Griffiths, C. (1994) 'The interaction of population growth and environmental quality', *American Economic Review*, 84(2).

Dinerstein, E. *et al.* (2017) 'An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm', *BioScience*, 67(6). doi: 10.1093/biosci/bix014.

FAO (2000) *Global Forest Resources Assessment*. Rome.

Ferraro, P. J. and Simonrangkir, R. (2020) 'Conditional cash transfers to alleviate poverty also reduced deforestation in Indonesia', *Science Advances*, 6(24).

Foster, A. D. and Rosenzweig, M. R. (2003) 'Economic growth and the rise of forests', *The Quarterly Journal of Economics*.

Freyaldenhoven, S., Hansen, C. and Shapiro, J. M. (2019) 'Pre-event trends in the panel event-study design', *American Economic Review*, 109(9).

Hansen, M. C. *et al.* (2013) 'High-Resolution Global Maps of 21st-Century Forest Cover Change', *Science*, 342, pp. 850–854.

Kahn-Lang, A. and Lang, K. (2018) 'The promise and pitfalls of differences-in-differences: Reflections on "16 and Pregnant" and other applications', *NBER Woring Paper Series No. 24857*.

Kreibaum, M. (2016) 'Their Suffering , Our Burden? How Congolese Refugees Affect the

Ugandan Population', *World Development*. Elsevier Ltd, 78, pp. 262–287. doi: 10.1016/j.worlddev.2015.10.019.

Martin, A. (2005) 'Environmental conflict between refugee and host communities', *Journal of Peace Research*, 42(3).

Maystadt, J.-F. *et al.* (2020) 'Vegetation changes attributable to refugees in Africa coincide with agricultural deforestation', *Environmental Research Letters*, In press. Available at: https://doi.org/10.1088/1748-9326/ab6d7c.

Maystadt, J.-F. and Duranton, G. (2019) 'The development push of refugees : evidence from Tanzania', *Journal of Economic Geography*, 19, pp. 299–334. doi: 10.1093/jeg/lby020.

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.

Maystadt, J. *et al.* (2019) 'Impacts of Hosting Forced Migrants in Poor Countries', *Annual Review of Resource Economics*.

Miteva, D. A. *et al.* (2017) 'Spatial patterns of market participation and resource extraction: fuelwood collection in northern Uganda', *American Journal of Agricultural Economics*, 99(4).

Mulumba, D. (2011) 'The gendered politics of firewood in Kiryandongo refugee settlement in Uganda', *African Geographical Review*, 30(1).

Oxfam International (2015) *Nduta refugee camp reopens to welcome thousands of Burundian refugees*.

Rambachan, A. and Roth, J. (2020) 'An honest approach to parallel trends'.

Rivoal, M. and Haselip, J. A. (2017) 'The true cost of using traditional fuels in a humanitarian setting: Case study of the Nyarugusu refugee camp, Kigoma region, Tanzania', *UNEP DTU Partnership Working Paper Series*, 3.

Roth, J. (2020) 'Pre-test with caution: Event stude estimates after testing for parallel trends'.

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.

Ruiz, I. and Vargas-Silva, C. (2016) 'The labour market consequences of hosting refugees', *Journal of Economic Geography*, 16. doi: 10.1093/jeg/lbv019.

Scrieciu, S. S. (2007) 'Can economic causes of tropical deforestation be identified at a global level?', *Ecological Economics*, 52.

Sexton, J. O. *et al.* (2013) 'Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS Vegetation Continuous Fields with lidar-based estimates of error', *International Journal of Digital Earth*.

Tropek, R. *et al.* (2014) ‘Comment on “ High-resolution global maps of 21st-century forest cover change ”’, *Science*, 344(6187).

UNHCR (2019) *Figures at a Glance*, *About Us*. Available at: https://www.unhcr.org/figures-at-a-glance.html (Accessed: 16 August 2019).

UNHCR and FAO (2018) *Managing Forests in Displacement Settings*.

Verme, P. and Schuettler, K. (2019) ‘The Impact of Forced Displacement on Host Communities: A Review of the Empirical Literature in Economics’, *World Bank Policy Research Paper 8727*.

Verme, P. and Schuettler, K. (2021) ‘The impact of forced displacement on host communities: A review of the empirical literature in economics’, *Journal of Development Economics*. Elsevier Ltd, 150(January), p. 102606. doi: 10.1016/j.jdeveco.2020.102606.

Verwimp, P. and Maystadt, J.-F. (2015) ‘Forced Displacement and Refugees in Sub-Saharan Africa An Economic Inquiry’, *World Bank Policy Research Working Paper 7517*.

Whitaker, B. E. (2002) ‘Refugees in western Tanzania: The distribution of burdens and benefits among local hosts’, *Journal of Refugee Studies*, 15(4).