Essays on the Economic Impact of Forced Displacement

Olive U. Nsababera

Thesis submitted for the degree of Doctor of Philosophy

Department of Economics

University of Sussex

December 2020

Declaration

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

The work presented in the first chapter of this thesis ‘Refugee camps – A lasting legacy? Evidence on long-term health impact’ is published in Economics and Human Biology (Nsababera, 2020). I confirm that I conceived of the idea, performed the statistical analysis, and wrote the manuscript. Richard Dickens and Richard Disney supervised the research process. The final publication is available at https://doi.org/10.1016/j.ehb.2020.100926.

Signature:

Olive Nsababera

UNIVERSITY OF SUSSEX

OLIVE UMUHIRE NSABABERA

DOCTOR OF PHILOSOPHY IN ECONOMICS

ESSAYS ON THE ECONOMIC IMPACT OF FORCED DISPLACEMENT

SUMMARY

This thesis examines the economic impact of forced displacement on hosting regions – a challenge facing several countries but one for which empirical evidence is limited. It focuses on Tanzania, one of the countries that has hosted a large and prolonged presence of refugee camps in order to assess their long-term economic impact. It also examines persistence of the effects – a key contribution to the literature. The first chapter introduces the studies and describes the context.

The first study is an individual-level analysis examining the long-term health effect of childhood exposure to refugee camps. I exploit the geographic variation and the fact that different birth cohorts were exposed to different stages of the camp life cycle - from camp establishment to camp closure - to study the differential effects from exposure. I find that individuals exposed to the early stages of camps were negatively affected and that the effect persisted into adulthood. I investigate possible channels through which health was affected and conclude that the primary channel was poor sanitation in initial camp phases. I find no effect on the subsequent generation.

Moving from individual welfare, the second study focusses on data. Lack of subnational data is a challenge prevalent in the force displacement literature, and, indeed, in most developing countries. The study examines whether daytime satellite imagery can be used to predict local measures of economic activity and welfare. I first examine the predictive ability in a cross-sectional setting and find that features derived from remote sensing imagery can predict a large share of the variation in subnational agricultural occupation, wealth and consumption. Next, I evaluate whether the model can generalise across time. The results are mixed. Additionally, error rates of the predictions have implications for potential applications.

The third study turns to forced displacement as a potential, but hitherto understudied, driver of urban growth. It examines whether refugee camps, as centres of resource inflows such as humanitarian aid, infrastructure investment, and local trade, had an urbanising effect on surrounding localities. Built-up area from daytime satellite imagery for the period 1985-2015 is used as a proxy for settlement and non-agricultural economic activity. The main finding is that camps had a small positive urbanising effect. The results are consistent with evidence that camp presence was not associated with profound structural change.

Acknowledgments

My time as a PhD student has been a truly rewarding experience. For that I am deeply grateful to my supervisors Richard Dickens and Richard Disney for their invaluable guidance and support. They allowed me to fully explore my interests and in the unexpected twists and turns, they patiently guided me while never losing the eye for detail or the bigger picture. I am truly grateful to have been their student and to have benefitted from the thoughtful attention, constructive feedback and encouragement they put into training me to be a competent researcher. This thesis is only a small reflection of what I have learned from them. I hope it is of some reassurance to them that the many seeds they have planted will continue to bear fruit beyond this thesis. I wish to also thank Sanchari Roy who was my supervisor in the initial stages of this thesis.

My sincere gratitude also goes to Barry Reilly, my academic advisor, for maintaining a keen interest in my progress and for helpful comments. His dedication and concern for students also deepened my appreciation for teaching and enriched my time as a doctoral tutor.
Looking back, I have been very fortunate to have had nurturing teachers. That I was able to undertake this thesis at all, I owe in large part to the help of the many that have graced my path. I remain forever grateful to Perpetua Macharia who generously took it upon herself to help me and my siblings adjust to a foreign land. To Benson Mulama and many more whose inspiring influence and encouragement endures – “Look at what I have written! Who would have thought? You would have. Thank you!” To Mrs. Kimei who showed me a door that has since led to opportunities that seemed unimaginable – I can never thank you enough. I also remain grateful to Douglas Almond who during my masters’ studies generously made time to supervise my independent study on refugees in the Middle East. That experience was instrumental in my decision to pursue a PhD and my choice of topic.

I am thankful to the University of Sussex for generously funding my PhD. Many thanks also go to the faculty in the Department of Economics for helpful comments and feedback received during my studies. My fellow PhD students have been wonderful compatriots and I am thankful to them for fostering a supportive academic environment. I would like to especially thank my “seniors” - Subhani Keerthiratne, Eugenia Go, Nihar Shembavnekar, Hannah Sam, Panka Bencsik, Nick Jacob, Antonia Schwarz, Hector Rufrancos and Rashaad Shabab for sharing their accumulated wisdom. Eugenia Go patiently listened to me brainstorm and was a source of sound advice and encouragement throughout. Her magnanimity and commitment to her work continue to be an example for me.

Access to data was critical to this thesis and presented a wonderful window into other aspects of what makes a good researcher. I am humbled and inspired by the help I have received from several researchers and institutions. I owe many thanks to my supervisor Richard Dickens for making numerous introductions and inquiries on my behalf even without my asking. Andy McKay and Barry Reilly graciously stepped in to help facilitate access to data and provided valuable introductions. I thank Francois Maystadt, Zhou Yang-Yang, and the

United Nations High Commissioner for Refugee (UNHCR) office in Tanzania for fielding my questions and generously sharing data on refugee camps in Tanzania. I am also thankful to individuals from Tanzania who not only responded to my requests but did so with the warmest of wishes. Ani Silwal kindly provided data on *Ujamaa* coverage in Tanzania. I am especially thankful to James Muthoka who introduced me to remote sensing data and patiently answered my many questions. My gratitude also goes to Ran Goldblatt and Rick Lawrence for graciously providing valuable insights on remote sensing data. I am much obliged to Mattia Marconcini of the German Aerospace centre (DLR) who graciously provided the data that is critical to chapter 3 and 4 of this thesis. All errors are my own.

I am very thankful to my friends who have encouraged me during this thesis. I am particularly thankful to Rahma Ahmed for being an excellent sounding board and I thank her, Eugenia Go, and Helen Connolly for proof reading sections of this thesis.

I am grateful to have had a chance to study a topic that is very close to my heart. I know too well that this thesis cannot fully capture the many dimensions of forced displacement and I am heartened by the surge in academic interest on this topic. That I have been able to contribute in any way to this, I owe much to my family. I am very fortunate to have had the support of my siblings Celestin, Yves, Olivier and Yvette. They inspire (and challenge!) me to roll up my sleeves and to have the courage to forge new paths. Thank you for sharing so fully in my aspirations. To my father and mother- my unsung heroes. This thesis would have been impossible without you. For your loving sacrifices, your courage and the many lessons you have taught me – I dedicate this thesis to you.

Finally, to a trio of ancient refugees - Jesus, Mary and Joseph - who have been the most faithful of companions on this and every journey - thank you.

Table of Contents

Chapter 1

Introduction

Forced displacement is among the most pressing socio-economic challenges of the 21st Century. According to the United Nations High Commissioner for Refugees (UNHCR), the rate of displacement has doubled since 2010 and the most recent estimates suggest that there are about 79.5 million displaced people globally – the highest on record. Of these, 29.6 million are refugees and individuals who, although not legally considered refugees, are also forcibly displaced outside their countries. A majority of refugees (85%) are in developing countries and a significant number of them live in camps which have been the conventional policy response for hosting displaced populations (UNHCR, 2020).

For a long time, the issue received little attention in the Economics literature. Two reasons could explain this. First, forced displacement has predominantly been viewed as a humanitarian issue. It is only recently that there is increasing recognition that addressing forced displacement is indispensable to economic development (World Bank, 2017). Recent studies demonstrate that forced displacement has implications for the economic development of hosting regions. Forced displacement shocks affect the local labour market, household welfare, human capital and the environment in hosting regions (Alix-Garcia et al., 2013; Alix-Garcia & Saah, 2010; Baez, 2011; Maystadt et al., 2020; Maystadt & Duranton, 2019; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015, 2016). Second, the shortage of data poses considerable challenge to research in this area. Data gaps are rampant in developing countries (Beegle et al., 2016; Dang et al., 2019; Devarajan, 2013; Serajuddin et al., 2015).

The scarcity of data is more acute at sub-national levels where regular data on measures of economic activity and welfare are usually lacking or collected infrequently.

This thesis contributes to filling this gap in the literature using the context of a refugee influx from Rwanda and Burundi into Tanzania in 1993 and 1994. The experience of Tanzania offers an opportunity to causally identify the impact of forced displacement. The availability of recent studies on the impact of proximity to camps in Tanzania is a unique opportunity to compare the results and extend the frontiers of literature with new insights and findings on long-term effects. Most of the empirical evidence on the impact of refugee camp presence on hosting regions has tended to focus on short term effects. Yet, displacement is often protracted and camps that are set up as an emergency, temporary response to cater for the needs of the displaced while they await eventual return, often evolve into long-term settlements. A household panel survey – the Kagera Health Demographic Survey (KHDS) – which, coincidentally, started prior to the refugee influx, paved the way for much of the pioneering work in this area. However, the KHDS covers only one region of Tanzania, and is limited for understanding the impact of prolonged camp presence – an important issue given the tendency of refugee camps to be protracted. Broadening the sample to include individuals and areas near protracted camps is necessary for understanding impacts from the later evolution of the relationship of camps to local communities. The thesis will exploit the variation in duration of camps in the empirical strategy in a novel way that allows for a richer study of the differential effects from exposure and their persistence after camp closures. Besides augmenting the scope of the household survey and camp sample, data obtained from machine learning methods and satellite imagery also allows me to extend the geographical breadth and to do so at a spatially granular level which is key to capturing localized effects and the effects of protracted camp presence.

The first paper of the thesis examines the long-term health effect of childhood exposure to refugee camps. The focus on childhood health is important given the well-established understanding that this has important long-term implications for cognitive development and labour market outcomes. In hosting regions, the immediate aftermath of a forced displacement shock is typically characterized by pressure on resources, land degradation and outbreaks of communicable diseases owing to the sudden population increase. In subsequent years, however, refugee camps become centres of resource inflows such as humanitarian aid, infrastructural investments and local trade. There is evidence that camps are associated with an increase in household consumption and wealth, although the effects are heterogeneous across economic activities (Alix-Garcia & Saah, 2010; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015). To the best of my knowledge, only Baez (2011), has examined the effect on the health of the local population, and found that children's health in hosting areas was adversely affected. There is, however, no research to date on whether the health impact persists into adulthood. This is an important question given the possibility that children may recover from initial poor health. This is particularly true if some of the positive spillovers take a longer time to arise, than the short-run negative shock arising from the mass influx of refugees.

Taking height-for-age z-score (HAZ) as a proxy for health, the paper exploits the fact that different birth cohorts were exposed to different stages of the camps' lifecycle − from camp establishment to camp closure. I find that individuals exposed to the early stages of camps were negatively affected and provide new evidence for the first time in the literature, that the effect has persisted into adulthood. I investigate possible channels through which health was affected and conclude that the effect can be explained by poor sanitation in initial camp phases. A further contribution of this paper is to examine the effect on a subsequent generation exposed to the later camp stages and those born after camps. On one hand, if we take the evidence that refugee presence exerted a pressure on resources, those born after camps closed, when this pressure is alleviated, may have better outcomes than those exposed to camps. On the other hand, if refugee presence generated positive spillovers, those born after camps closed may be worse off. This question has yet to be considered in the literature. I find no difference in health between those born in the post-camp period and those born when the camps were in operation.

The second paper of the thesis focusses on data − a challenge prevalent in the forced displacement literature, and, indeed, in most developing countries. While data scarcity can be both a consequence of forced displacement and a characteristic of forced displacement contexts, the paper does not focus on data as an impact of forced displacement. Rather, it is motivated by the question of how we can address data gaps that limit research in these contexts. To this end, the study examines whether daytime satellite imagery can be used to predict local measures of economic activity and welfare. The paper contributes to a nascent line of enquiry that investigates whether machine learning techniques applied to daytime satellite data can be used for predicting subnational economic well-being in developing countries. This literature is still inconclusive on whether the predictive power of daytime data holds in different contexts and for which outcomes and whether prediction can be improved by combining with other routinely collected sources of data.

I combine data from a remote sensing satellite (the Landsat) and two waves of household survey data. The availability of an extensive and publicly available archive of Landsat images mean it could have an important potential for filling data gaps for economics applications. First, I examine the ability of features from the Landsat satellite to predict welfare and economic activity in a cross-sectional setting. I find that features derived from Landsat data, combined with weather and geophysical characteristics of the locations, can predict a large share of the variation in the sub-national share working in agriculture, wealth and consumption although error rates remain high. Next, I evaluate performance over time and find that the ability to generalise over time is mixed. I also find that models using built-up area as a predictor perform better than directly using imagery characteristics. The paper concludes that there is value to be gained from efforts that seek to use high dimensional satellite imagery to derive metrics that are predictive of economic outcomes but also intuitively easier to interpret.

The third paper of the thesis turns to forced displacement as a potential but understudied driver of urban growth. The emergence of new urban agglomerations in previously rural areas has motivated a need to examine other drivers of urbanization that depart from the traditional story of structural transformation excess – labour moving from agriculture to the other sectors. While evidence from some contexts suggests that an inflow of displaced populations could lead to emergence of ‘boomtowns’, lack of data has left this channel largely unexamined. Understanding how forced displacement affects urban growth processes is critical to effectively managing the pace of urban growth. To overcome the data limitations, I use built-up area from daytime satellite imagery for the period 1985-2015 as a proxy for settlement and non-agricultural economic activity. This allows me to examine whether there was expansion of existing urban localities and/or rural-urban transformation of rural localities. In so doing, the paper contributes to a deeper understanding of the effects of forced displacement beyond individual and household level outcomes to wider processes such as urban growth. I find that camp presence is associated with increased settlement and non-agricultural economic activity, albeit modestly and primarily in rural areas. I find no strong evidence that camp presence was associated with densification of localities that were already more urban. I suggest that the constrained institutional context in which refugee camps operate may explain why, despite engendering conditions that may be conducive to transformation of surrounding localities, camp presence is associated with a small increase in settlement and non-agricultural economic activity. This may also explain why results may differ in other contexts. In the following section, I describe the Tanzanian context relevant for the empirical analysis conducted in this thesis.

1.1 Context

In 1993, civil war erupted in Burundi following the assassination of the president. In October of that year, about 250,000 Burundians fled across the border into Tanzania. A few months later in April 1994, the plane carrying the president of Rwanda and the new president of Burundi was shot down. Within 24 hours, 250,000 refugees from Rwanda fled into Tanzania and more continued to pour in over the subsequent months of 1994. Based on UNHCR sources, Maystadt and Verwimp (2014) notes that the scale and pace of the influx was unprecedented. In total, Tanzania received more than 800,000 refugees from Burundi and Rwanda in the short period. It is estimated that the refugees represented more than a third of the local population in the two recipient regions of Kagera and Kigoma (Adisa, 1996; Maystadt & Verwimp, 2014; Whitaker, 1999). Prior to the refugee influx, Kagera and Kigoma being among the remotest regions of Tanzania, were also among the poorest regions (Green, 1995).

The sudden nature of the events and the scale of the inflow caught the UNHCR and the government of Tanzania off guard. The refugees received little humanitarian assistance at the onset of the influx in October 1993 and in some instances, members of the local community are reported to have helped them with food supplies until camps were operational (Veney, 2007). They settled near the border, often in the open, because the limited means of transportation and the terrain of the region limited their mobility. The choice of camp sites was taken by the Tanzanian Ministry of Home Affairs and the UNHCR. The reported criteria in the designation of sites were that refugees should be able to reach them and that costs of future repatriation when the situation in the sending countries became more peaceful should be minimised. The choice of sites was limited because by the time the UNHCR and the government of Tanzania moved into action in 1994, it was deemed too costly to move them farther away from the border (Lupala, 2015; Maystadt & Verwimp, 2014).

Because of the large number of refugees entering during a short period of time, relief agencies were constrained. As a result, the initial camps that were operational in 1994 grew rapidly and had larger populations than subsequent camps that were opened between 1995 and 1999. As an example, one of the early camps, Benaco, became the world’s largest refugee camp at the time. Its rapid development is also attributed to the fortuitous presence of a construction company with the requisite machinery. The government and the UNHCR contracted the company for camp construction (Adisa, 1996; Maystadt & Duranton, 2019; Tanzanian Affairs, 1994). Map 1 shows the location of the refugee camps.

The sudden population shock caused pressure on natural resources. The presence of refugees led to an increase in demand for firewood and water and resulted in deforestation and land degradation of some regions close to the camps (Green, 1995; Whitaker, 1999). Relief agencies reported outbreaks of diseases such as malaria, cholera and dysentery in local areas (Eriksson et al., 1996). Whitaker (1999) further documents events during this period; the author notes that the refugee influx was also accompanied by a proliferation of humanitarian aid agencies and expatriate workers. Relief agencies hired local labour and many local employees from government hospitals and schools are reported to have left their positions in favour of the higher salaries from aid organizations. Due to the demand for housing by expatriate workers, housing prices increased. Foreign workers, however, also created a demand for goods such as chocolate and cheese, and enterprising locals took advantage of the new market opportunities (Whitaker, 1999).

Although the refugees were kept in camps, evidence shows that refugees interacted with the local population. They are reported to have worked as labourers in neighbouring villages weeding, harvesting, clearing land, tending livestock or fetching water and firewood. Commercial centres also developed around the camps. There were daily markets and farmers who had previously traded across the border before the camps, now sold their products in nearby camps. According to the World Food Program that distributed aid, refugees traded about 75 % of their aid receipts (Whitaker, 1999).

During the duration of the camps, vast sums of money were injected into Kagera and Kigoma. By 1999, it was reported that more than $15 million had been spent on improving roads, airstrips and communication infrastructure in Kagera region alone. Locals around the camps also gained access to health facilities in the refugee camps. Before the arrival of the refugees, Tanzania had just introduced a cost-sharing scheme requiring locals to start paying part of their health care costs, but at refugee health facilities, locals could access health services free of charge. In the event of referrals, free transportation was provided for both refugees and the local population to district hospitals (Lupala, 2015; Whitaker, 1999). Locals also benefitted from UNHCR-installed water outlets in the camps and surrounding villages (Wolfcarius, 2008).

In 1996, when the security situation was deemed to have improved, most Rwandan refugees were repatriated and camps that had hosted them closed down. The return of Burundian refugees was more gradual because of the continued poor security situation in Burundi. In 2002, the UNHCR and the Tanzanian government officially launched a repatriation programme to facilitate the return of Burundian refugees. Repatriation was initially limited to relatively safe areas in northern Burundi. In subsequent years, camp closures continued as they emptied out upon the return of the refugees. The closure of the camps was governed by a tripartite agreement between the UNHCR and the governments of Tanzania and Burundi that all camps with a refugee population of under 10,000 would be closed. This was to consolidate camps to reduce operation costs (Ongpin, 2008; Veney, 2007). Camp closing involved screening to assess which refugees were deemed in need of continued international protection and which would be repatriated (UNHCR, 2011). In 2006, about 350,000 refugees still remained. Mass repatriation efforts reduced the figures even further over the following three years. By 2009, only about 50,000 Rwandan and Burundi refugees remained – the lowest figure for the first time in 15 years (See Fig. 1). Several UN sources viewed 2009 as the first "camp free" year in the Kagera region. In Kigoma, by 2009 camps had all but closed down with the exception of three that housed the remaining refugees (Zhou, 2014). Overall, nine of the camps closed in 1996, one in 2005, eight in 2007 and 2008, one in 2009 and two in 2012. Fig. 1 shows the number of Rwandan and Burundian refugees in Tanzania up to 2014.

Figure 1: Number of Refugees from Rwanda and Burundi in Tanzania, 1961-2014

Source:(UNHCR, 2019b)

The presence of the refugees had so altered local conditions that the departure of refugees was predicted to result in a vacuum. Citing Whitaker (1999), Han (2009: p.16) noted "if, over the course of refugee presence, local farmers learned to change production patterns to cater to new local market demands, then such farmers could face food insecurity and loss of income after camp disbandment . . . "

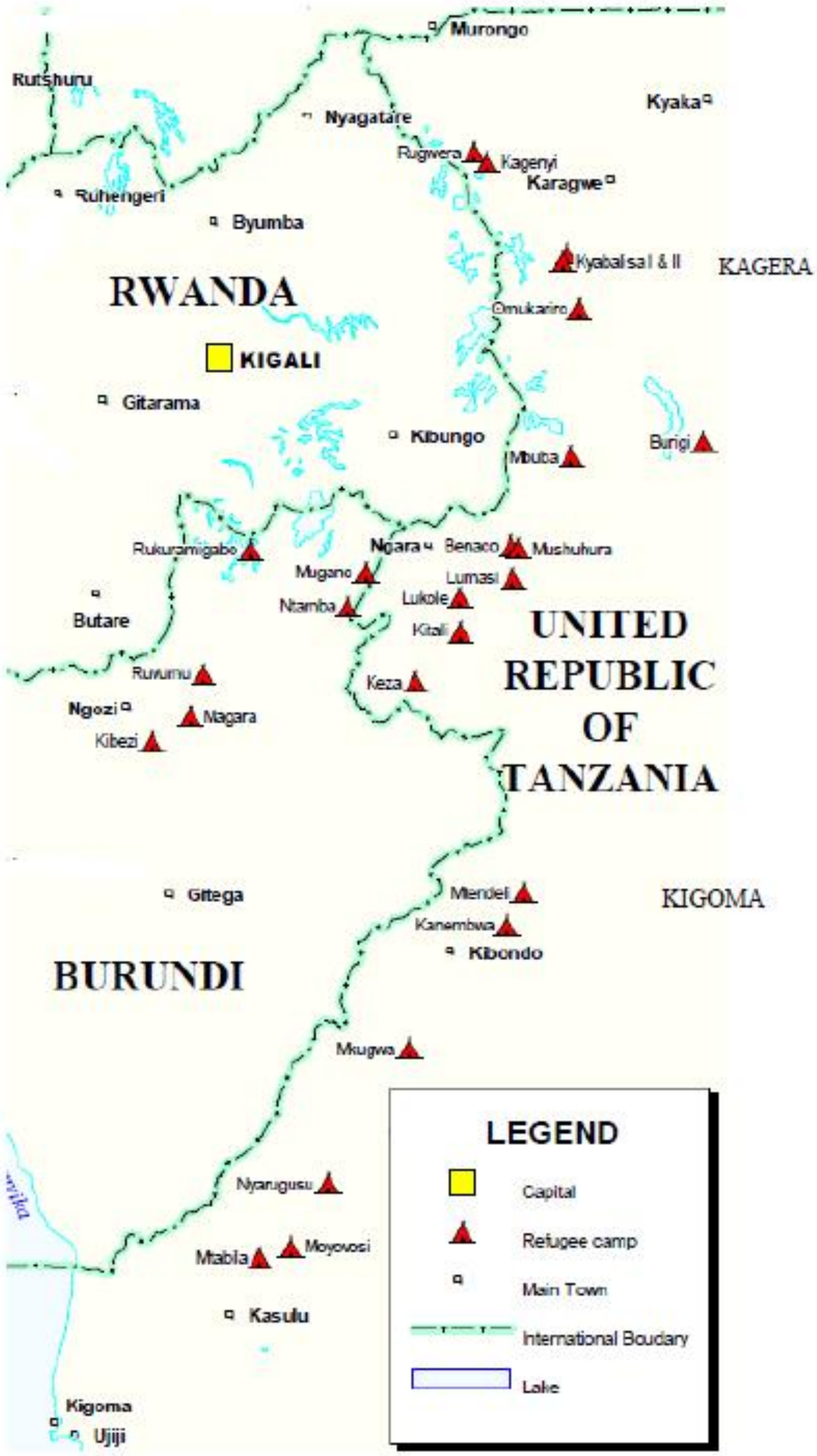
After camps closed, premises and facilities were handed over to local district authorities. The UNHCR together with the local government established a programme for the rehabilitation and reconstruction of former campsites referred to as the Joint Programme J.P. 6.1 Transition from Humanitarian Assistance to Sustainable Development in North-western Tanzania. It was intended to address the gap left by the withdrawal of humanitarian agencies. The programme which ran from 2008 to 2011 with a budget of over US $10 million had three areas of intervention: Wealth Creation, Social Services and Governance, and Sustainable Management of Natural Resources. A component of the programme was the training of health personnel. Some former campsites were transformed into schools, others into health centres and one became a military training centre (Han, 2009; Lupala, 2015; UNDP, 2009).

During this period, the prevailing stance was that the camps were a temporary solution and refugees were expected to eventually return to their countries. Initially, there were few movement restrictions. Consequently, refugees are reported to have provided labour on Tanzanian farms and engaged in trade in local markets and in the camps. In 1998, the government tightened control on movement. Refugees had no right to work or to own land. Despite restrictions, it is documented that locals still managed to hire refugee labor. Tanzanians continued to use camp facilities such as water and health centres and operated small businesses that were frequented by refugees. Refugees exchanged or sold aid items such as foodstuffs, clothing, construction material and cooking implements for items that were not included in their aid baskets (Veney, 2007; Whitaker, 1999). The restrictive policies were meant to discourage refugees from viewing Tanzania as a permanent place to settle. This was a break from Tanzania's socialist past under which the country had maintained an "open-door" policy (Veney, 2007).

More broadly, Tanzania's refugee hosting experience and policies mirror some of the currents in the region. Between 1960 and 1990, refugee policies in Africa were fairly open and those fleeing in search of safety had few restrictions on economic activity and they were encouraged to be self-sufficient. The marked shift from this stance in the 1990s has been attributed to several reasons. First, there were concerns about possible spill over of conflict into internal borders. Second, the economic decline underway at the time in many African countries fuelled concerns over the fiscal effects of hosting refugees. Additionally, the magnitude of displacement in the 1990s reached an unprecedented level, owing in large part to the dramatic influx from Rwanda and Burundi (Rutinwa, 2017; Veney, 2007). Tanzania's neighbouring countries, Kenya and Uganda, also witnessed an influx of refugee populations due to conflicts in Burundi, Rwanda, the Democratic Republic of Congo, South Sudan, Eritrea and Somalia. Between 1993 and 2002, Tanzania and Kenya were among the top-ten refugee hosting states in Africa with Tanzania being the largest refugee hosting country on the continent between 1997 and 2002. In 2002, Tanzania's refugee population was the third largest globally (Veney, 2007). Kenya similarly adopted an encampment approach and established camps that are still operational due to ongoing insecurity in countries of origin. In recent years, the country has threatened to close these camps and to repatriate the inhabitants ("Kenya Says Go Home," 2016; Sanghi et al., 2016; UNHCR, 2021). This is predicted to adversely affect both refugees and hosting regions given the symbiotic relationship between the camps and local communities (Sanghi et al., 2016). By contrast, in Uganda refugees are given the right to work and freedom of movement. Uganda's approach has been juxtaposed with the encampment model (Betts et al., 2019). The recognition that forced displacement situations are often not temporary but rather last for many years has led to debates on the appropriate policy responses. The scale of the influx in Tanzania, prolonged presence of camps and subsequent closures present an opportunity for a deeper understanding of how the documented interactions may impact economic outcomes in hosting regions.

This thesis sets out to examine the long-term impact of the camps on hosting regions providing new empirical evidence on effects along the different stages of camp operation from initial opening until after camp closures. The rest of the thesis is structured as follows: The second chapter investigates the legacy of refugee camps on the health of children in hosting regions. The third chapter turns to investigating potential methods of addressing data gaps that are prevalent in such contexts. The fourth chapter examines whether the camps, as centres of resource, inflows into hosting regions, provided an impetus to urban growth. The fifth chapter is the conclusion.

Map 1: Refugee camps in Tanzania



Source: UNHCR, 1999

Chapter 2

Refugee camps – A lasting legacy? Evidence on long-term health impact

2.1 Introduction

Conflict is often accompanied by displacement of individuals within and across national borders. In the host regions, the immediate aftermath of such forced migration shocks is typically characterised by pressure on existing resources. Of the 70.8 million forcibly displaced individuals globally, a majority live in refugee camps in developing countries (UNHCR, 2019a). These camps are intended as an emergency response, allowing humanitarian organisations to deliver emergency shelter, food, water and medical care. Owing to the increasingly protracted nature of conflict, however, these camps often evolve into long-term settlements, lasting for several years and becoming loci of continued resource inflows such as humanitarian aid, infrastructural investments and evolving local trading.

Previous micro-level analysis has found that proximity to camps has an effect on local labour market outcomes, household consumption, human capital and the wealth of host communities (Alix-Garcia & Saah, 2010; Baez, 2011; Becker & Ferrara, 2019; Maystadt & Duranton, 2019; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015). Of these studies, however, only a few have looked at the long-term impact (Maystadt & Duranton, 2019; Ruiz & Vargas-Silva, 2015) . Studies tend to focus on cohorts that were alive at the beginning of the refugee influx and are also based on a sample of camps that were in operation for a relatively shorter period.

Other long-term impacts on host communities, including the impact on cohorts of individuals born in the period the camps were in operation and thus exposed to the later evolution of the relationship of camps to local communities, and indeed those born after camps closed, are still under-researched. This paper contributes to this gap in the literature using data arising from the refugee influx from Rwanda and Burundi into Tanzania in the period 1993– 2012.

The paper uses height-for-age z-score (HAZ) as a proxy for health, to explore one impact of the fact that different birth cohorts were exposed to different stages of refugee camps. The temporal variation through birth cohorts is coupled with variation in distance to refugee camps to create a difference-in-difference estimation approach.

This paper is closely related to the work of Baez (2011) who studies the impact of exposure to refugee camps on human capital in Tanzania. However, the perspective and methodology of this paper differ in two important respects. First, whereas Baez (2011) solely examined the effect on the health of individuals aged 10–15 who were exposed to the refugee influx in their childhood, this paper examines whether the effect of childhood exposure, if any, lasts into adulthood i.e. age 20 and above. Qualitative evidence shows that, in the immediate aftermath of the refugee influx, the increase in population was associated with an increase in deforestation, communicable diseases and pressure on resources. In subsequent years, however, there were some positive spillovers. Locals are reported to have benefitted from roads and healthcare services provided to camp populations (Whitaker, 1999). Other studies also found that locals around refugee camps experienced an increase in consumption levels and in assets, although the effects are heterogeneous across economic activities (Alix-Garcia

& Saah, 2010; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015). In light of these possible positive spillovers, and because poor health in childhood may still be recouped in the growth window, the first objective is to establish whether the effect of camp exposure is discernible in later life. Secondly, while previous research has focused on the initial phases of refugee arrival, this paper considers another juncture that is crucial for host regions – the departure of refugees and closure of camps. By comparing the generation of local children that were born after the refugee influx to those born after camps closed, I examine the effect of having been exposed to the camps relative to being born in a post-camp era. These later cohorts have not previously been studied. The closure of the camps could have led to a loss of positive spillovers which may have negatively affected those born after the camps closed. On the other hand, the departure of refugees and closure of camps could have alleviated pressure on local resources to the benefit of those born after camps closed. This paper seeks to examine what the effects were.

Another difference of the present paper to other studies is the use of a different dataset (the World Bank Living Standards Measurement Survey (LSMS) rather than the Kagera Health Demographic Survey (KHDS)), which, combined with new data on camps in the Kigoma region, improves upon the geographic scope of previous studies.

The paper finds a negative, albeit localised, effect of exposure to camps discernible through to adulthood, for individuals that were children at the time of the refugee influx. The effect is a reduction in HAZ of 0.29 between those who lived within 50 km of a camp and those that were more than 50 km away. The result is comparable to a 4.3% difference in adult hourly earnings. However, those that were exposed for a longer duration were less affected. These results are robust to different measures of exposure and provide evidence that the negative effect of camps may have dissipated over time. Among the later generation who are still children (at the end of the period of study), the results indicate that there is no observable difference in the HAZ between those born during camp operations and in the post- camp period.

The remainder of the paper is structured as follows: section 2.2 provides an overview of the pertinent literature, section 2.3 describes the data used and section 2.4 outlines the empirical methodology. The results are discussed in section 2.5. Section 2.6 provides some robustness checks while section 2.7 concludes.

2.2 Literature review

Health shocks may arise from an increase in, or lack of, employment opportunities. Forced migration constitutes a labour supply shock in recipient regions. Braun and Mahmoud (2014) argue that the magnitude of the effect depends on the degree of substitutability between migrants and locals. Using data from the 1950s after the expulsion of millions of Germans from Eastern Europe to West Germany after World War II, they find that the forced migration shock reduced the employment rate of the local workforce. In this case the forced migrants and natives are very close substitutes: the forced migrants spoke the same language (German) and had been educated in German schools. Similarly, in this context, we can note that Rwandans and Burundians are not linguistically distant from the locals in the Kagera and Kigoma region and could, therefore, be considered close substitutes.

Calderón and Ibáñez (2009) find that the labour market effects are largest for low-skilled workers. They examine the impact of the Colombian conflict on regions within Colombia that hosted the internally displaced population. They distinguish between high and low-skilled workers in recipient regions and find a general decline in wages, but that it is greatest among the low-skilled.

In Tanzania, there is evidence that the forced migration shock had an impact on host regions' labour market and goods market. Ruiz and Vargas-Silva (Ruiz & Vargas-Silva, 2015, 2016) examine the impact of the forced migration shock on the type of employment held by the locals. They find that exposure to the refugee shock resulted in locals having a higher likelihood of working in household farms or tending household livestock, and a lower likelihood of working outside the household as employees. They find that the low likelihood of being employed outside the household was particularly strong among Tanzanians that had been doing casual work before the shock where they faced competition from refugees. Their sample is composed only of individuals that were alive at the beginning of the refugee influx.

A second potential impact on health arises via consumption goods and wealth. Alix-Garcia and Saah (2010) examine the impact of proximity to refugee camps and access to aid on the prices of Tanzanian agricultural goods in adjacent markets between 1992 and 1998. The authors find that there was an increase in the prices of those agricultural goods that were staples of the refugee diet and that were not provided as part of the aid ration. Whitaker (1999) provides extensive and useful qualitative insights into these price effects that resulted from the composition of aid rations to the refugees. Food aid was typically in the form of maize, cooking oil and beans. The diet of Rwandans and Burundians is, however, primarily cassava and green bananas while Tanzanians prefer maize. Refugees therefore sold their aid goods to locals in exchange for cassava and plantain. As a result, the prices of bananas and cassava increased sharply and there were reports of locals endangering the food security of their households by selling large amounts of their food stocks in order to take advantage of the high prices of these goods. In contrast, the price of maize declined as the local markets were flooded with maize from the refugees and local farmers were unable to sell their own maize produce. These authors also assess the impact of the refugee presence on short-term household wealth. They find, overall, an increased incidence of wealth indicators such as radios and bicycles in rural households closer to the refugee camps which they attribute to the price impact arising from the change of diet. The wealth effects, therefore, differ across households. Rural residents living near refugee camps benefitted from selling their stock of agricultural products. On the other hand, the authors explain that because urban households are more likely to be buyers of agricultural goods for consumption, they are affected by the high prices and, therefore, experience negative wealth effects.

Similar findings arise from Maystadt and Verwimp (2014), who study the impact of the forced migration shock on the welfare of households in the Kagera region of Tanzania between 1994 and 2004. They find an overall positive albeit heterogeneous impact on household consumption. Agricultural workers were worse off, and the authors suggest that this is due to increased competition from refugees for those jobs. They also find that, despite the reported surge in local entrepreneurship, those self-employed in non-agricultural activities experienced a relative welfare drop, perhaps because of increased competition from other local entrepreneurs who came from other Tanzanian regions. In a subsequent paper, Maystadt and Duranton (2019), use data from 2004 and 2010 and still find that overall, the local population had higher consumption levels. They provide evidence that this effect is driven by infrastructural investments in the regions, specifically roads built to serve the refugee camps.

One study directly focuses on the health of the local population. Baez (2011) finds a negative impact of the refugee influx on the human capital of local Tanzanian children. First, the author examines the impact of hosting refugees on the health of 0–5-year-old children in 1996, less than two years after the arrival of the refugees. To do so, he compares the children in more affected areas with those in less affected areas in 1992 and 1996. He finds a 15–20 percentage point increase in the incidence of infectious diseases, an increase of roughly 7 percentage points in mortality for children under five and a decline in HAZ of 0.3 standard deviations. Second, the author employs a double difference comparing individuals in highly affected areas with those in lesser-affected areas at the onset in 1994, (when they are 0–5 years old) and in 2004 (when they are 10–15 years old). He still finds a negative effect on health.

These results warrant further discussion. Individuals are believed to keep growing until age 20 (see, for example, Moradi (2010). Empirical evidence shows that children that are lagging behind in growth can catch up to their peers and that this "catch up growth" happens in puberty. Indeed, evidence from Tanzania corroborates this phenomenon (Hirvonen (2014). At the time that Baez (2011) observes individuals in 2004, (when they are 10–15 years of age) they have not attained full adult stature. Therefore, whether the effect of exposure in childhood, if any, was permanent and lasting into adulthood, is an important question and a gap in the literature. This is particularly true if some of the positive spillover effects discussed above take a longer time to arise than the short-run negative shock arising from the mass influx of refugees. Hence, using more recent data from 2012, I link adult health outcome (proxied by HAZ) with data on the year of opening and closing of the nearest camp during an individual's growth window (defined as age 0–19). Exposure to refugee presence was determined both by proximity to the camp and by the age of the individual during the operational years of the nearest camp. This approach will capture the long-term impacts on the health of individuals that were children at the beginning of the influx and are now adults as well as the health of later cohorts with shorter exposure. A similar strategy has been used to study the impact of school construction during childhood on education and earnings in adulthood (Duflo, 2001). Such cohort analyses rely on two types of variation – variation in the cohorts that are exposed to a treatment, and geographic variation in the treatment. Implementing such a cohort dimension by exploiting years of camp operation during an individual's childhood is the novelty of this paper. Beyond shedding light on whether the effect of childhood exposure is discernible in adulthood, this paper can go a step further and examine whether the impacts are heterogeneous by the duration of the exposure during the growth window. Having a longer exposure could have afforded more opportunity to benefit from any positive effects but it could also imply more exposure to the negative effects.

A further and principal contribution of this paper is that it attempts to examine the effect on individuals that have previously not been the focus of any study in this literature - the groups born during and after camp operations. The aforementioned studies all examine impacts on individuals that were alive at the beginning of the refugee influx. The effect on the generation of individuals that are born during the refugee era or after camps close is a question that has yet to be considered in this literature. Given that the migration shock was associated with effects on the labour market, the goods market, household consumption and household wealth, it is reasonable to hypothesise that the children born during the refugee era would fare differently from those born after camps closed. On one hand, if we take the evidence that refugee presence exerted a pressure on resources, those born after camps closed when this pressure is alleviated may have better outcomes than those exposed to camps. On the other hand, if refugee presence generated positive spillovers, those born after camps closed may be worse off. Two studies outside the economics literature provide support for the latter hypothesis. Lupala (2015) conducted a case study on the impact of closing the Mtabila refugee camp in Tanzania on the livelihoods of surrounding residents. The author conducted qualitative interviews with 198 households in three villages near Mtabila refugee camp after its closure. Villagers reported that the closure of the camp had led to loss of services such as health centres, schools, water supply and routine road maintenance. They also reported that after the closure of the camp there had been a decline of agricultural production, which they note had been dependent on unskilled refugee labour. Similarly, Han (2009) conducted qualitative interviews in three villages near two refugee camps in Tanzania (Kanembwa and Karago) after they closed. Among her findings, she notes that while initially the demand for construction materials by NGOs for building camp facilities led to increases in prices of construction materials in the towns that hosted the refugees, the high prices persisted even after the camps closed (Han, 2009). Although Lupala (2015) and Han (2009) are descriptive qualitative studies and do not aim to draw causal links, they nevertheless hint at possible impacts of closing refugee camps. This paper intends to contribute to this gap in literature.

This paper also differs from the aforementioned existing studies in its use of the LSMS. Previous studies on Tanzania have primarily used the KHDS, which only contains data on one region of Tanzania. The LSMS, on the other hand, additionally allows for the study of the impacts on the Kigoma region, which was also affected by the forced migration shock. The inclusion of camps in the Kigoma region adds more temporal variation which is exploited in the empirical strategy. Camp protraction was more pronounced in Kigoma.

Globally, refugee camps tend to be protracted so broadening the sample to include children that were near protracted camps is useful for understanding the impact of protracted exposure. Finally, the LSMS has height data on respondents of all ages and gender, which allows for richer study along these dimensions.

As a measure of exposure to the refugee presence, the literature uses geographic variation in refugee presence generated by the shock. Baez (2011), classifies the western districts that border Rwanda and Burundi as treatment districts, whereas the eastern part of Kagera are used as controls. In the same paper, the author also uses distance of villages from the border of Rwanda. The use of the district level variation is potentially problematic as the Eastern district of Kagera also hosted refugees. In view of this, subsequent work (Maystadt & Duranton, 2019; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015) proposed an intensity index based on distance to all camps weighted by an estimate of the population of the camps as a less noisy measure of refugee presence. In this paper, I will provide a further refinement presenting analysis based on distance to the nearest camp. In addition, I will also present analysis using the intensity index with a modification that allows the index to not only be village- specific but also cohort-specific.

2.3 Data description

2.3.1 Birth cohorts

This analysis uses a nationally representative household survey – the 2012 Tanzania Living Standards Measurement Survey (LSMS) – to look at health outcomes across defined birth cohorts. I limit the sample to the two administrative regions of Kagera and Kigoma where the camps were located and the three surrounding regions of Mara, Mwanza and Shinyanga.[1] I thereby retain 3,934 individuals.

The analysis uses only the 2012 data. The 2014 wave of the LSMS, data collection for which was completed in 2015, is the most recent but it is not preferred – there was renewed outbreak of violence in Burundi in 2015 that caused a new movement of refugees into Tanzania. Although the magnitude was small compared to that of 1993, at least two camps that had been closed were reopened. Although not a concern for individuals that were already height mature by the times camps closed, it may be a concern for younger individuals that were interviewed in 2015. Using the 2012 round avoids this concern altogether. Because the LSMS is not available pre-1993 (i.e., before the camps opened), I use the 2012 wave and construct birth cohorts for temporal variation. I restrict the sample to individuals that are between ages 0–49 at the time of the survey in 2012. Age 49 is chosen as the upper limit because the normal process of ageing typically begins at this point and height shrinkage or loss of stature begins (Cline et al., 1989; Moradi, 2010). Height growth typically occurs between age 0–19 and most individuals attain final adult height by age 20 (See Moradi, 2010). I therefore exploit the fact that different age groups had differing exposure to the refugee influx. I define two broad groups based on survey year, year of birth and camp operation dates.

2.3.2 Adults at the time of the survey

At the time of the survey in 2012, all these individuals are adults i.e., they are above 19 years of age. However, individuals who were age 0–19 at camp opening were susceptible to the

[1] Given the potential channels through which refugee presence may affect host communities discussed in the literature review and context sections, it is unlikely that individuals beyond those regions would have interacted with the camps. In the methodology section, I discuss how I farther limit the sample by distance.

effects of the refugee influx as they were still children. On the other hand, individuals who were age 20–30 years at camp opening were already height mature at the time of the influx – they may have been affected in other ways such as in the labour market, but not height-wise. As will be discussed further in the empirical strategy, comparison of adults in the survey who were children at camp opening and those who were mature at camp opening will be the basis for examining the long-term impact of exposure to camps.

2.3.3 Children at the time of the survey

These individuals are still children at the time of the survey. They were either (i) exposed to a camp at some point in their childhood or (ii) they were born after the camps closed. Comparison of children born before camps closed and those born after camps closed will form the basis for examining how, in this generation of individuals that are still children, individuals born during the camp era fare relative to those born after camps closed.

For each village, I calculate how long each cohort would have been exposed to each camp in childhood. The result is a by village by camp cohort-specific duration measure for which a given camp (c) was operational during an individual of village v of cohort k’s childhood (age 0–19). Table 2.1 illustrates how various cohorts were exposed to the refugee camp presence, by way of illustration, using one camp, Lukole. In the analysis however, this is done for each camp.

Table 2.1: Cohort exposure (time dimension) to camp presence

Subgroup	Description	Sample type	Age at survey	Age in 1993	Year of birth	length of exposure to camp Example: Lukole A
Alive at the beginning of the influx						
1	Mature at the beginning of the influx (1993)	Mature at survey. Long-run effect – placebo	39-49	20-30	1963 - 1973	0
2	Child at the beginning of influx (1993)	Mature at survey. Long-run effect	32-38	13-19	1974	1
					1975	2
					1976	3
					1977	4
					1978	5
					1979	6
					1980	7
			26-31	7-12	1981	8
					1982	9
					1983	10
					1984	11
					1985	12
					1986	13
			20-25	0-6	1987	14
					1988	15
					1989	15
					1990	15
					1991	15
					1992	15
Born during the refugee era or after camps closed						
3	Not alive at the beginning of the influx (born during the refugee era or after camps closed)	Child at survey. Young generation.	13-19	N/A	1993	16
					1994	15
					1995	14
					1996	13
					1997	12
					1998	11
					1999	10
			7-12	N/A	2000	9
					2001	8
					2002	7
					2003	6
					2004	5
					2005	4
			0-6	N/A	2006	3
					2007	2
					2008	2
					2009	1
					2010	0
					2011	0
					2012	0

Note: Based on the age at survey (2012) individuals are first split into those that are adults and those that are still children. The adult sample (subgroup 1 and 2) is used to examine the long-run effect in the old generation, while the child sample (subgroup 3) is used to examine the effect within the young generation. Within each sample, the year of birth, year of opening and closing of each camp are further used to determine duration of exposure of a particular cohort to a particular camp. Duration is defined as the number of years that a cohort was exposed during its childhood window of age 0-19. The example shows the number of years of exposure to Lukole A camp. (E.g. Lukole A was opened in 1993 and closed in 2008. An individual born in 1987 would have been 6 years old when it opened, therefore, upto and including age 19, they would have had 14 years of exposure.) The cohort exposure dimension is then combined with variation in distance from a camp to an individuals' village. Age bands (0-6; 7-12; 13-19) are chosen to align with growth spurts.

2.3.4 Location of clusters/villages[2]

The geographic coordinates of clusters are obtained from the LSMS. The sample comprises 194 clusters. For confidentiality purposes, the LSMS does not provide the precise geographic coordinates of a household. Rather, the LSMS assigns each household in a cluster the average of the GPS coordinates of households in that cluster, randomly offset within a range of 0-5km. Because the empirical strategy relies on exposure by distance from camps (see section on methodology), the offsetting introduces measurement error in distances constructed using cluster locations. However, as the offset is random, the measurement error should be random therefore mitigating concern regarding bias in estimated effects. One approach that is suggested for further mitigating this concern is to use distance bands (Perez-Heydrich et al., 2013) and this paper will present results that use this approach.

Since the location of individuals is based on residence at the time of survey, it may be different from their location before the refugee influx. This is less of a concern in the younger generation as most of the sample (84%) was born in the 2012 location, compared to 41% in the adult sample. In the absence of detailed migration history, analysis on these sub-samples of non-migrants will be conducted to address concerns that individuals may have been born elsewhere and exposed to a different environment from that of their residence at the time of survey.

2.3.5 Camps

Distance to camps will form the basis for estimating exposure to refugee presence. I discuss this distance measure in the next subsection. Information on these camps is compiled from

---

[2] Clusters are typically defined by village boundaries (LSMS, 2012).

various sources. The geographic coordinates of the camps in Kagera are obtained from Maystadt and Verwimp (2014). The geographic coordinates of the camps in Kigoma, and the opening and closing dates of camps in both Kagera and Kigoma are obtained from Zhou (2014) and from the UNHCR field office in Tanzania. For the population of the camps, I use estimates from Maystadt and Verwimp (2014) and various UNHCR reports. I use the highest population estimate available as a proxy for the size of the camp. The total camp sample comprises 22 camps, 11 in Kagera and 11 in Kigoma. The population across camps ranges from 2,155 to 350,000 over these years.

2.3.6 Measures of exposure to camps

In addition to exposure to refugee presence by cohorts described above, I construct measures of exposure to refugee presence as follows.

2.3.6.1 Nearest camp analysis

First, I consider only the nearest camp. After excluding observations that do not have a HAZ score (outcome variable), of the individuals that have a HAZ score and are in my sample of interest 0–49-year olds, 31 individuals were more than 430 km away from a camp. After excluding these outliers, I end up with a final sample of 3,934 individuals.

Table 2.2: Summary statistics of distance to the nearest camp

	Mean	Std. Dev.	Min	Max	p10	p25	p50	p75
Distance to nearest camp (km)	173.9	112.6	2	404.2	21.1	78.5	168.0	278.5
Observations	3934							

Instead of a continuous measure of distance to the nearest camp, I define a discrete treatment variable where I classify individuals having a camp within 20 km as the treatment group and those farther than 20 km as the control group. I then vary this threshold to 50 km, 80 km, 100 km and 170 km. The choice of the thresholds is motivated by the reality of accessibility in North Western Tanzania as well as by the distribution of the distance to nearest camp variable (see Table 2.2). For instance, given that one of the ways locals interacted with camps was access to services, there is a limit to what is a feasible distance to travel. The 20 km also corresponds to the 10th percentile of the distance to the nearest camp, while 80 km and 170 km correspond to the 25th and 50th percentile respectively.

2.3.6.2 Intensity of all camps

Second, I consider all camps and create an intensity exposure index. Specifically, following the approach of Maystadt and Verwimp (2014) and Ruiz and Vargas-Silva (Ruiz & Vargas-Silva, 2015, 2016), I create a *camp exposure intensity index* by weighting the inverse distance from each village to each refugee camp by the population of each camp ($Population_c$) as a proportion of the population of all camps. The inverse of the distance is taken to reflect the inverse relationship between intensity of exposure and distance – villages closer to the camp experienced greater intensity than those farther away. Weighting by population ensures proportionality so that the possible effect of smaller camps is not exaggerated and that of larger camps underestimated.

I deviate from Maystadt and Verwimp (2014) and Ruiz and Vargas-Silva (Ruiz & Vargas-Silva, 2015, 2016) in that I weigh each camp by the duration for which a camp was open during an individual's childhood (which varies by cohort) rather than weighting by the total number of years the camp was operational (which is fixed). As discussed above, cohort exposure Table 2.1 provides an illustration of how duration of exposure to each camp for each cohort is determined based on the year of birth, year of opening and year of closing of the camp. As a result, whereas the intensity index in Maystadt and Verwimp (2014) and Ruiz and Vargas-Silva (Ruiz & Vargas-Silva, 2015, 2016) is fixed by village, I introduce time variation.

The by village by cohort intensity exposure variable (I_{vk}) is thus:

$$I_{vk} = \left(\sum_{c=1}^{22}\left[\left(\frac{1}{distance_{v,c}}\right)\left(Duration_{v,k,c}\right)\left(\frac{Population_c}{\sum_{c=1}^{22} Population_c}\right)\right]\right)$$

Each of the above measures will be incorporated into the estimation equation discussed in the methodology section below.

2.3.7 Outcome variable (Height-for-age-z score (HAZ))

The outcome variable of interest is height for age z-score – HAZ. HAZ is defined as the difference, expressed in standard deviation units, between an individual's height and the median height of a healthy and well-nourished population of the same age and gender ("reference population"). I use the WHO 2007 standard as the reference population as it is the most recent and also contains a more diverse pool of ethnic and cultural backgrounds than previous standards (De Onis et al., 2007). The average HAZ in Tanzania is -1.5 (DHS, 2015).

HAZ is a reliable indicator of long-run nutritional status(Akresh et al., 2012; Thomas et al., 1996; WHO, 2005). Because it reflects household income and consumption, it is considered a good non-monetized measure of welfare and has been used to estimate effects of childhood exposure to crop failure (Akresh et al., 2011), war (Akresh et al., 2012) and agricultural price shocks (Cogneau & Jedwab, 2012). The effects of economic shocks experienced in childhood are reflected in HAZ (Akresh et al., 2011; Cogneau & Jedwab, 2012; Micklewright & Ismail, 2001). Furthermore, height is associated with labour market outcomes, educational attainment and effects over the lifecycle (Ahsan & Maharaj, 2018; Alderman et al., 2006; Schultz, 2002; Thomas & Strauss, 1997). Previous cohort studies have found that low HAZ in childhood is associated with lower earnings in adulthood (Galasso et al., 2016; Hoddinott et al., 2011; Victora et al., 2008). Having a HAZ of less than – 2 ("stunting"), is of particular concern as it impairs cognitive development and has grave implications for later socio-economic outcomes (WHO, 1995).

HAZ is preferred over BMI because the latter fluctuates and may reflect both long and short-run health dimensions (Strauss & Thomas, 1998). Additionally, the identification strategy exploits the height growth window (see sections 2.3.1 and 2.4). This allows us to overcome the lack of data pre-camps. By contrast, lack of pre-camp data on other outcomes such as on the labor market preclude their use.

To obtain the HAZ, I use individuals' height measurement and age from the LSMS. Rather than the self-reported age in years, for better precision, I calculate an individual's age in months as the difference between the month of birth and the month when the height measurement was taken.

Table 2.3 provides the mean of the HAZ between individuals within 100 km of a camp and those farther than 100 km, across age cohorts. Individuals within 100 km of a camp on average have a statistically significant lower HAZ than those in villages far away. A key point is that the effect is also statistically significant for the oldest cohort i.e., individuals age 39–49 at the time of the survey. Since these individuals were mature before 1993, the fact that we observe a statistically significant difference indicates that villages closer to areas that later became camps were worse off than those farther away, even before the refugee influx.

Table 2.3: Differences in means of HAZ between those within 100km of a camp and those farther than 100km but less than 430km, by age groups

	(1) Full mean	(2) Within 100km mean	(3) More than 100km mean	(4) diff
Age 39-49	-0.84	-1.08	-0.75	0.32**
Age 32-38	-0.80	-1.08	-0.69	0.40**
Age 26-31	-1.04	-1.52	-0.84	0.68***
Age 20-25	-0.93	-1.46	-0.73	0.74***
Age 13-19	-1.41	-1.72	-1.29	0.43***
Age 7-12	-1.49	-1.88	-1.33	0.55***
Age 0-6	-1.40	-1.81	-1.23	0.58***
Observations	3934	1134	2800	3934

2.4 Empirical methodology

I exploit temporal variation (through birth cohorts) and spatial variation (through distance thresholds) and use a difference-in-difference with a birth cohort fixed effects model.

Identification of the effect of camp presence and the effect of closing camps on HAZ could be challenged by endogeneity in camp locations. The crucial identification assumption relied upon is common trends – there could be differences in levels of HAZ between locations as long as trends would have been the same in the absence of camps. Regarding location of the camps, the existing empirical studies on Tanzania (see literature review), rely on the fact that the refugees were pushed into Tanzania by the internal conflict within their own countries, which is unlikely to have been affected by socio-economic outcomes in Tanzania (Alix-Garcia & Saah, 2010; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015, 2016). The refugees settled where they could and by the time the UNHCR responded, it was deemed unfeasible, cost-wise, to relocate them far from the borders (Maystadt & Verwimp, 2014). With regard to camps closing, this was largely exogenous to the socioeconomic conditions in Tanzania. Camps closed because refugees left. According to the UNHCR, the possibility of return was dependent upon the security situation in their countries. In the robustness section, I conduct a test of pre-trends to explore the validity of the identification strategy.

2.4.1 Impact of exposure to camp presence on adults who were children at the opening of camps (long-term effect)

Comparison of adults that were children at the opening of camps to those that were height mature, forms the basis of examining the long-term impact of exposure to camp presence on those who were children at the opening of camps. These individuals (subgroups 1 and 2 in Table 2.1) are observed as adults in 2012. I use the following equation:

$$HAZ_{ivk} = \beta_0 + \beta_1 (TREAT_d) + \beta_2 \left(TREAT_d * Child_{open}\right) + \beta_3\, male_i + \delta_k + \theta_j + e_{ivk} \quad (2.1)$$

The dependent variable, HAZ_{ivk}, is the height-for-age z-score of an individual i in village v born in year k. $Child\_open$ is a dummy variable that indicates whether or not an individual was a child at camp opening. $TREAT_d$ is a dummy indicating whether the individual's village is within d km of the nearest camp. I estimate the equation using different distance thresholds (20 km, 50 km, 80 km, 100 km and 170 km respectively). Construction of the distance thresholds are described in section 2.3.6. θ_j and δ_k are region fixed effects and year of birth fixed effects respectively. e_{ivk} is a random idiosyncratic error term. Identification of the impact of the refugee shock comes from comparing individuals who were children at the opening of the camps and those who were not, together with the variation in the exposure of villages. The difference-in-difference estimator of interest is β_2 .

As an alternative measure of exposure, I use the intensity index (I_{vk}) discussed in section 2.3.6, to estimate the following:

$$HAZ_{ivk} = \beta_0 + \beta_1 (I_{vk}) + \beta_2\, male_i + \delta_k + \theta_j + e_{ivk} \quad (2.2)$$

2.4.2 Impact on younger generation

Identification of the effect of having been exposed to a camp for the younger generation comes from comparing children born before camps closed to those born after camps closed and by variation in the exposure of villages. These individuals are subgroup 3 in Table 2.1.

I estimate a variant of the long-term exposure Eq. 2.1 in which I use a discrete binary variable C_i, which takes the value of 1 if the nearest camp was operational at some point in the individual's childhood (regardless of duration) and 0 otherwise. In subsequent analyses I refine the variable further by using different categories of duration of exposure during childhood.

To examine the aggregate exposure to all camps in the younger generation using the intensity index (I_{vk}), I estimate Eq. 2.2 for the younger generation.

2.5 Results and discussion

In the following sections, I present the results for adults who were children at the opening of the camps (long-term effect), followed by results on the younger generation.

2.5.1 Impact of exposure to camp presence on adults who were children at the opening of camps (long-term effect)

Table 2.4 shows the results of the baseline specification Eq. (2.1) when treatment is defined using the discrete distance thresholds. Each column represents a different distance cut-off. In the absence of strong priors about exposure cut-offs, I present results from different cut-offs. The Table shows a negative effect of camp exposure. The effect is larger when treatment cut-off is defined narrowly (a decline in the HAZ by 0.4 standard deviations). When the distance threshold is increased to 170 km there is no statistically significant effect, suggesting that the effect of exposure was localised. In the robustness checks, I compare the effect among different distance bands to further examine the decreasing gradient of the effect and confirm the effect was stronger for areas closest to camps. Additionally, as discussed in section 2.3.4 one may be concerned that individuals born elsewhere may have been exposed to a different environment from that of their residence at the time of survey. After separating the sub-samples of adults that reside in their place of birth from those that do not, I find the estimated magnitudes for the non-migrant sample are larger than in the full sample. The results are presented in Table A9 in the Appendix.

Turning to the effect identified in the literature, (Baez, 2011), finds a worsening of HAZ by 0.3 standard deviations in childhood. In adulthood, I find a decline ranging between 0.2-0.4 standard deviations (1.38–2.76 cm) depending on how strictly the treatment threshold is defined. Notwithstanding the differences in methodology, dataset and cohorts between this paper and (Baez, 2011), the findings in this section confirm a negative effect of camp exposure on the health of this older generation and provide first evidence that the negative effect lasted into adulthood.

Viewing the results in the context of the HAZ of Tanzania may better contextualise this finding. A decrease in HAZ of 0.2-0.4 from the average of -1.5, is a non-trivial effect. This implies a reduction of the HAZ to a range of -1.7 to -1.9 which would tip the individual closer to stunting (HAZ of less than -2).

The magnitude of the effect can also be expressed in terms of earnings using existing estimates of the association between HAZ and labour market outcomes in developing countries. One such estimate, though not from Tanzania, suggests that a 1 standard deviation increase in HAZ at 3 years raises hourly earnings by 14.8% (Hoddinott et al., 2011). Using this as a simple guide would translate the magnitude of the effect to a 2.9% – 5.9% reduction in adult hourly earnings for an individual that was 3 years of age.

The negative coefficient on the male variable in Table 2.4 and indeed in all regression Tables reflects that males tend to have a lower height-for-age z-score than females. This is confirmed by other data sets on Tanzania. For example, the 2015 Demographic Health Survey for Tanzania shows males have an average HAZ of -1.5 compared to -1.4 for females.

Table 2.4: Long-term impact of exposure to camps on adults who were children in 1993 (using distance thresholds)

	HAZ				
	(1)	(2)	(3)	(4)	(5)
treat20*childopen	-0.400**				
	(0.161)				
treat50*childopen		-0.285**			
		(0.126)			
treat80*childopen			-0.195*		
			(0.114)		
treat100*childopen				-0.236**	
				(0.105)	
treat170*childopen					-0.061
					(0.104)
Male	-0.320***	-0.319***	-0.315***	-0.316***	-0.324***
	(0.048)	(0.047)	(0.048)	(0.047)	(0.047)
Observations	1,313	1,313	1,313	1,313	1,313
R-squared	0.145	0.145	0.146	0.147	0.153
Birth year fixed effects	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes

Note: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. “treat d” is a dummy variable that takes 1 for individuals within d km of a camp and 0 otherwise. All models include a constant term and the base treatment term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

2.5.1.1 Does the effect vary by duration of exposure?

In this exercise, I explicitly examine whether the effect on HAZ score varies by length of camp exposure. It should be noted that year-of-birth fixed effects do not fully control for the effect of duration because individuals born in the same year may have different durations based on which camp is nearest.

On average, individuals in the adult sample were exposed to a camp in their childhood for 3 years. I classify individuals into three categories of exposure based on the duration – 0 years of exposure (no exposure), 1–2 years ($duration_{short}$) and 3 or more than 3 years ($duration_{long}$). The 0 years of exposure serves as the reference group. This can be thought of as an extension of Eq. 2.1 above where I now use different levels instead of the binary $Child\_open$ variable. Specifically, I interact duration with the discrete distance thresholds to examine if treatment varies by duration.

$$HAZ_{ivk} = \beta_0 + \beta_1 (TREAT_d) + \beta_2 (TREAT_d * duration_{short}) + \beta_3 (TREAT_d * duration_{long}) + \beta_4 male_i + \delta_k + \theta_j + e_{ivk} \quad (2.1.1)$$

The results in Table 2.5 reveal interesting insights into the effect of duration of exposure. First, when treatment is restricted to areas in close proximity to the camps (where treatment is defined as being within 20 km or 50 km of a camp), the effect of being exposed for a shorter duration is greater than that of being exposed for a longer duration relative to the reference group (i.e., exposure of 0 years).

The opposite is true when the distance is increased to include areas farther from the camps (where treatment is defined as being within 80 km or 100 km away from the nearest camp). In this case, the effect of being exposed for a shorter duration relative to 0 years is not statistically significant. Rather, it is the effect of being exposed for a longer duration relative to 0 years that is statistically significant.

Why is it that for areas closest to the camps short exposure in childhood is more detrimental than longer exposure? A likely explanation is that the initial negative effect was due to outbreaks of disease. Disease outbreaks would have been most severe closest to camps. This explanation is consistent with reports of the nature of the refugee crisis at the time – both the government of Tanzania and the UNHCR were not prepared for the scale and swiftness of the crisis. Outbreaks of disease, pressure on resources, deforestation and land degradation were reported. In section 2.5.3 I provide suggestive evidence that diseases were more prevalent in initial camp phases.

For the long exposure group, continued camp presence may have afforded opportunities that mitigated the initial adverse effect. Disease outbreaks tended to be severe in the immediate aftermath of camp establishment but eventually peter out in the long run. The long exposure group could have recovered from some of the initial negative effects. As discussed in section 2.2, poor health in childhood may still be recouped in the growth window. That health of children nearest to camps and exposed long-term could improve is consistent with findings from the existing literature of increased household welfare in proximity to camps, attributed to infrastructure investment (Maystadt & Duranton, 2019). The inflow of humanitarian aid and infrastructure investment may also have had a direct effect on health for example by facilitating greater access to health services and thereby helping to mitigate the negative effect among the long-term exposed. In section 2.5.3, I test this possibility of a direct health effect through utilisation of health services.

When the threshold is widened to areas farther away from the camps, exposure for a short time had no statistically significant effect, consistent with the explanation above that the negative effect was due to outbreak of disease. This was more severe nearer camps and less so farther away. It may seem counterintuitive that longer exposure is detrimental when including children farther from the camps. But this can be reconciled with existing evidence that household welfare improvements were stronger closer to the camps. Studies found that wealth increases are strongest for households near camps, whereas distant households showed wealth decline (Alix-Garcia & Saah, 2010). It is, therefore, unsurprising that nutritional conditions of children from more distant households exposed long-term, experience welfare declines or are negatively impacted.

Table 2.5: Long-term effect of camp exposure by distance thresholds from nearest camp and duration of exposure

	HAZ				
	(1)	(2)	(3)	(4)	(5)
treat20	0.220				
	(0.193)				
treat20*duration short	-0.618**				
	(0.282)				
treat20*duration long	-0.323*				
	(0.176)				
treat50		0.120			
		(0.159)			
treat50*duration short		-0.347*			
		(0.186)			
treat50*duration long		-0.243			
		(0.149)			
treat80					
			-0.107		
treat80*duration short			(0.118)		
			-0.0372		
treat80*duration long			(0.123)		
			-0.337**		
treat100				-0.130	
				(0.121)	
treat100*duration short				-0.0688	
				(0.112)	
treat100*duration long				-0.393***	
				(0.129)	
treat170					-0.279**
					(0.116)
treat170*duration short					-0.00945
					(0.112)
treat170*duration long					-0.121
					(0.123)
male	-0.321***	-0.320***	-0.313***	-0.313***	-0.323***
	(0.0477)	(0.0466)	(0.0476)	(0.0474)	(0.0476)
Observations	1,313	1,313	1,313	1,313	1,313
R-squared	0.146	0.145	0.148	0.150	0.154
Birth year fixed effects	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	1,313	1,313	1,313	1,313	1,313

Notes: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. "treat d" is a dummy variable that takes 1 for individuals within d km of a camp and 0 otherwise. All models include a constant term. The reference group is those that have 0 duration of exposure. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05,* p<0.1

2.5.1.2 Does the effect vary by age at exposure?

While the results in Tables 2.4 and 2.5 suggest that there is a negative effect on HAZ as a result of exposure to the refugee shock for those adults that were closest to camps, an important question is whether certain ages were more affected. To investigate this, I consider whether the effect of the refugee shock is heterogeneous across age groups. The differentiation of exposure by age group makes it possible to examine whether being exposed to the refugee shock had a differential effect on a certain critical age group compared to another. I group those who were children at the beginning of the influx into age categories 0-3, 4-8, 9-12, and 13-19, based on their age at the opening of the nearest camp. This age grouping aligns with the timing of growth spurts and is constructed with reference to the relationship between age and growth (Akresh et al., 2017; Case & Paxson, 2008). I estimate the following regressions that build on Eq. 2.1. The oldest age group, age category 13-19 serves as the reference group.

$$HAZ_{ivk} = \beta_0 + \beta_1 (TREAT_d) + \beta_2 \left(TREAT_d * ageatexposure_{0\_3}\right) + \beta_3\left(TREAT_d * ageatexposure_{4\_8}\right) + \beta_4 \left(TREAT_d * ageatexposure_{9\_12}\right) + \beta_5\, male_i + \theta_j + e_{ivk} \quad (2.1.2)$$

The results in Table 2.6 show that the effect is larger for age groups 0-3 and 9-12, consistent with the established view that these are periods of greater sensitivity and strongest growth. Relative to those who were age 13-19 at the time of the refugee shock, age group 0-3 (exposed in early childhood) was most negatively affected.

These findings are not inconsistent with the results from the previous subsection which shows short duration exposure to be more detrimental compared to long duration exposure. This is because age at exposure does not necessarily correlate with duration of exposure because they vary by camp. To illustrate this point, consider an individual who was in the youngest age group and whose nearest camp was Lukole A (closed in 2008). Such an individual would have been exposed for a long duration. On the other hand, an individual in the youngest age group whose nearest camp was Omukariro (closed in 1996) would have been exposed for a short duration.

Table 2.6: Long-term effect of camp exposure across age groups

	HAZ				
	(1)	(2)	(3)	(4)	(5)
treat20	0.241				
	(0.185)				
Age 9-12 at opening*treat20	-0.647***				
	(0.181)				
Age 4-8 at opening*treat20	-0.386				
	(0.238)				
Age 0-3 at opening*treat20	-0.756**				
	(0.332)				
treat50		0.170			
		(0.165)			
Age 9-12 at opening*treat50		-0.506***			
		(0.192)			
Age 4-8 at opening*treat50		-0.327**			
		(0.145)			
Age 0-3 at opening*treat50		-0.536***			
		(0.195)			
treat80			0.0306		
			(0.134)		
Age 9-12 at opening*treat80			-0.386**		
			(0.169)		
Age 4-8 at opening*treat80			-0.344***		
			(0.130)		
Age 0-3 at opening*treat80			-0.403**		
			(0.172)		
treat100				-0.0703	
				(0.138)	
Age 9-12 at opening*treat100				-0.383**	
				(0.150)	
Age 4-8 at opening*treat100				-0.371***	
				(0.118)	
Age 0-3 at opening*treat100				-0.404**	
				(0.156)	
Male	-0.328***	-0.321***	-0.320***	-0.321***	-0.346***
	(0.0619)	(0.0618)	(0.0622)	(0.0616)	(0.0621)
Observations	910	910	910	910	910
R-squared	0.155	0.156	0.157	0.162	0.168
Region fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. “treat d” is a dummy variable that takes 1 for individuals within d km of a camp and 0 otherwise. All models include a constant term. The reference group is those that were age 13-19 at exposure. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

2.5.1.3 Intensity analysis

The results presented thus far only consider the effect associated with the nearest camp. To investigate the aggregate effect of all camps, I present the results of Eq. 2.2 that uses the by village by cohort intensity index described in section 2.3.6. The descriptive statistics of the intensity index are provided in the Appendix.

Table 2.7 shows the results where I estimate the aggregate effect of all camps only on those that were exposed. Column 1 presents the results of the effect of proximity and duration whereas column 2 presents the results that include the population of the camps as weights. I find that intensity has a negative and statistically significant effect on HAZ. The magnitude increases when I use the intensity index that is adjusted for camp population, implying that it is not just distance to camps and duration that have an effect, but also the size of camps. This would be consistent with the reports that refugee labourers competed with locals for low wage jobs and other resources; the bigger camps would have had more competition for resources. Additionally, it would also be consistent with reports of disease outbreaks at the beginning of the influx; bigger camps would have been more susceptible to epidemics.

To interpret the result more easily, we can consider the aggregate effect of exposure for a child that had average exposure to all camps relative to one that had the least exposure i.e., lower intensity index. The coefficient in column 1 suggests that exposure to camps for an individual of average exposure relative to the least exposed is associated with a decrease in HAZ of 0.09.[3]

[3] The calculation is obtained by multiplying the coefficient on *Intensity* by the difference between the average and minimum of the intensity index I_{vk} . See Table A4 in the Appendix.

Table 2.7: Intensity analysis – adult sample

	HAZ	
	(1)	(2)
Intensity	-0.147***	
	(0.0523)	
Population adjusted intensity		-5.549**
		(2.141)
Male	-0.319***	-0.319***
	(0.0476)	(0.0477)
Observations	1,313	1,313
R-squared	0.148	0.150
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Note: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. All models include a constant term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

2.5.2 Impact of exposure to camp presence on the younger generation

Table 2.8 shows that there is no statistically significant effect in the younger generation in 2012 when comparing those born before and after camp closures across all distance thresholds. There are also no differential effects by duration of exposure (Table 2.9) when I conduct a similar exercise as that described in section 2.5.1 using Eq. 2.1.1 on the younger sample. These results hold even when using the intensity index - those who were exposed to the refugee presence do not fare any differently from those born in the post-camp era (Table 2.10). I also do not find any statistically significant effect among the sub-sample of non-migrants (Table A10 in the Appendix). This result is consistent with the explanation that the initial negative effect was likely due to outbreak of disease. For the younger generation, these issues had been managed; hence, there would be little difference when the camp closes.

In 6.1.1, I suggested that relief efforts associated infrastructure investments in the region and the resultant local trading opportunities arising from continued camp presence, may have helped reduce the initial adverse effect of the camps. If so, one can expect negative consequences to be associated with camp closure as relief efforts cease and local economic opportunities that were hitherto generated by the camps shrink. However, I find that this is not the case - the cohorts born after the camps are not worse off than those cohorts born when the camps were in operation. A possible explanation is that the camp rehabilitation and reconstruction programme (described in section 1.1), that was established in 2008 to fill the gap left by the withdrawal of humanitarian agencies, mitigated the adverse effects from the loss of benefits that would otherwise have been left by the closure of the camps. The efficacy of this programme clearly warrants further empirical investigation. At the time of writing, it has not been possible to obtain the data required to do so and it is deferred to future research.

Table 2.8: Impact of exposure to camps on younger generation (using distance thresholds)

	HAZ				
	(1)	(2)	(3)	(4)	(5)
treat20* C_i	-0.232				
	(0.252)				
treat50* C_i		0.0493			
		(0.226)			
treat80* C_i			-0.225		
			(0.181)		
treat100*C_i				-0.0745	
				(0.171)	
treat170*C_i					-0.103
					(0.127)
Male	-0.305***	-0.303***	-0.308***	-0.305***	-0.302***
	(0.0446)	(0.0450)	(0.0456)	(0.0459)	(0.0447)
Observations	2,621	2,621	2,621	2,621	2,621
R-squared	0.152	0.155	0.159	0.158	0.152
Birth year fixed effects	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes

Note: The individuals in this sample are those in subgroup 3 in the exposure Table 2.1. “treat d” is a dummy variable that takes 1 for individuals within d km of a camp and 0 otherwise. All models include a constant term and the base treatment term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table 2.9: Effect of camp exposure on younger generation by distance thresholds from nearest camp and duration of exposure

	HAZ				
	(1)	(2)	(3)	(4)	(5)
treat20	0.0983				
	(0.189)				
treat20*duration short	-0.352				
	(0.255)				
treat20*duration long	-0.185				
	(0.277)				
treat50		-0.264*			
		(0.145)			
treat50*duration short		0.0128			
		(0.218)			
treat50*duration long		0.0707			
		(0.261)			
treat80			-0.252**		
			(0.126)		
treat80*duration short			-0.0701		
			(0.166)		
treat80*duration long			-0.346		
			(0.224)		
treat100				-0.366***	
				(0.127)	
treat100*duration short				0.0288	
				(0.158)	
treat100*duration long				-0.174	
				(0.222)	
treat170					0.0386
					(0.110)
treat170*duration short					-0.0395
					(0.118)
treat170*duration long					-0.159
					(0.166)
Male	-0.306***	-0.304***	-0.305***	-0.304***	-0.300***
	(0.0446)	(0.0450)	(0.0454)	(0.0460)	(0.0448)
Observations	2,621	2,621	2,621	2,621	2,621
R-squared	0.153	0.155	0.160	0.159	0.153
Birth year fixed effects	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes

Note: The individuals in this sample are those in subgroup 3 in exposure Table 2.1. “treat d” is a dummy variable that takes 1 for individuals within d km of a camp and 0 otherwise. All models include a constant term and the base treatment term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table 2.10: Intensity analysis– younger generation

	HAZ	
	(1)	(2)
Intensity	-0.102	
	(0.0916)	
Population adjusted intensity		-4.863
		(3.176)
Male	-0.301***	-0.302***
	(0.0448)	(0.0450)
Observations	2,621	2,621
R-squared	0.153	0.154
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Note: The individuals in this sample are those in subgroup 3 in exposure Table 2.1. All models include a constant term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

2.5.3 Investigating potential channels

The preceding discussion in sections 2.5.1 and 2.5.2 suggested that the results are consistent with an explanation that the negative effect was due to disease spread that was subsequently brought under control. I have also suggested that infrastructure and aid inflow may have helped in mitigating the initial negative effect through improved household welfare and by facilitating access to health services. Data limitations constrain the ability to pin down the precise mechanisms through which effects of camp exposure on long term health can be mediated.

There are, nevertheless, indirect ways of testing these hypotheses. First, I test the hypothesis that prevalence of disease in local areas was more severe in earlier camp phases than in later years of camp operations. Secondly, I test whether utilisation of maternal and child health services was higher in later camp phases. The assumption in this second test is that infrastructure in the later camp phases could have facilitated access to health services or a change in health seeking behaviour.

I examine diarrhoea among local children, since there were reports of its outbreak in the aftermath of camp opening. Specifically, for children ages 0–5, I examine whether they have had diarrhoea in the two weeks preceding the survey. For utilisation of health services, I examine the place of delivery and whether those above 12 months of age have been fully vaccinated.[4] Delivery in health facilities is considered an important aspect of reducing health risks to children since proper medical attention and hygienic conditions during delivery can reduce the risk of subsequent infections and poor health. Similarly, childhood immunisations can reduce morbidity (WHO, 2005). This information is made available by pooling the 1999 and 2010 Demographic Health Surveys for Tanzania.[5] As all camps had opened by the time of the survey years, these data only allow me to examine whether disease prevalence and health service utilisation vary by the phase of the camp.

I classify camps into three categories of phases – those within 3 years of opening (campphase1), more than 3 years in operation (campphase2) and closed (campphase3). For each individual, I determine the phase of the nearest camp. Individuals near a camp which is in the first phase serve as the reference group. I then examine whether the likelihood of having diarrhoea or receiving all recommended childhood vaccinations varies by distance from the camp and by the camp phase controlling for individual characteristics, household

[4] Children are considered fully vaccinated if they have received all recommended childhood vaccinations i.e., BCG, measles, and three doses of DPT and polio vaccines. It is recommended that children receive full immunisation by 12 months of age but in Tanzania children are followed up until five years of age (NBS, 2000).
[5] The earliest available geocoded DHS survey for Tanzania is in 1999.

characteristics and weather conditions. I also examine whether, for women, being near a camp in the later phases is associated with higher likelihood of delivering in a health facility.

I find that the likelihood of having diarrhoea among children near camps in later phases (phases 2 and 3) is lower relative to children near camps in phase 1 (recent camps). Being near a camp in phase 2 is associated with a reduction in probability of having diarrhoea by 11.7 percentage points relative to being near a camp in phase 1. (See Table A5 in the Appendix). The result provides some supporting evidence that outbreak of disease in earlier camp stages prior to its being curbed was the likely channel driving the negative effect.

I do not find evidence that the likelihood of delivery in a health facility, or of receiving all recommended childhood vaccinations, varies by camp phase. (Table A6 in the Appendix). As such, we can tentatively conjecture that the mitigating effect of infrastructure and aid inflow alluded to earlier likely had no direct effect through health service utilisation but rather operated indirectly through household welfare. Place of delivery and immunisation coverage, however, albeit informative, are not exhaustive indicators of health service use and it is not possible to rule out that other aspects of health service use may have been affected.

2.5.4 Results by gender

I examine the effect on boys and girls separately to allow for the possibility that effects can vary by gender. In the adult sample, the results show negative impact for both boys and girls (Table A7 in the Appendix). This is consistent with disease spread – overall, both would be affected from outbreaks spread through contamination of shared resources such as water. Whereas among girls, however, both those closest and at intermediate distances were similarly affected, among boys the effect is confined to those who were closest to camps. While there is unlikely to be bias in disease spread, there may be some bias in intra-household resource allocation which may explain why for girls but not for boys, those at intermediate proximity are as similarly affected as those in closest proximity. This is plausible, for example, in the presence of household welfare declines in the more distant households alluded to in section 2.5.1.1. The existence of a gender bias during resource scarcity has been found in similar contexts.[6]

In the younger generation, consistent with the previous results for this sample, there is no statistically significant effect for either males or females (Table A8 in the Appendix).

2.6 Robustness checks

2.6.1 Testing for pre-trends

To explore the validity of the identification assumption, I conduct a test for pre-trends. Specifically, I test for whether there were differential trends by distance to the nearest camp ($distance_{vc}$) for individuals who were mature in 1993 (subgroup 1 in Table 2.1). I assume the refugee shock took effect at an earlier date (placebo treatment), by splitting the subgroup halfway into those that were age 20-25 years and those that were age 26-30 years in 1993. I use a "placebo-after" dummy that denotes whether or not an individual was born after the placebo treatment and estimate the following equation.

$$HAZ_{ivk} = \beta_0 + \beta_1 (distance_{vc}) + \beta_2 (Placeboafter * distance_{vc}) + \beta_3\, male_i + \delta_k + \theta_j + e_{ivk} \qquad (A.1)$$

Since this placebo treatment precedes the actual refugee shock and the individuals considered here were already mature at the time of the influx, the difference-in-difference estimator β_2 should be statistically insignificant. Table A11 in the Appendix shows that this is the case.

[6] For example, Akresh et al. (2011) find in Rwanda, only girls' health was negatively affected by crop failure.

2.6.2 *Baseline socioeconomic conditions*

To further allay concern about correlation with local socioeconomic conditions pre-camp, I use coverage of a national policy (*Ujamaa)* as a proxy for local socioeconomic conditions in 1993. *Ujamaa* was a socialist policy implemented in the 1970s in Tanzania. Infrastructure projects and provision of social services such as education, health and water supply were a key pillar of the policy (Coulson, 2013). The intensity of coverage can, therefore, serve as a proxy for the level of local economic development in 1993 (pre-camp). The measure has been shown to have a persistent impact on socioeconomic outcomes (Osafo-Kwaako, 2012; Silwal, 2016). To measure coverage of the policy at the local level, previous work has used the percentage of a district's population in the 1978 census that lived in planned villages (Silwal, 2016). After including this variable, the result still holds (Table A12 in the Appendix). The coefficient estimates and statistical significance levels are similar to the main specification.

2.6.3 *Balancing test*

The results may also be undermined if there were compositional changes in the treatment and control areas over time. To test the validity of this assumption, I use the sample of individuals born between 1993 and 2012 and test whether a series of covariates changed by cohort in treated versus untreated regions. I estimate covariate balance regressions of this form:

$$Covariate_{ivk} = \beta_0 + \beta_1 (TREAT_d) + \beta_2 (TREAT_d * C_i) + \delta_k + \theta_j + e_{ivk} \qquad (A.2)$$

I test this for household size, the gender of the individual, whether the household is rural, education of the household head, age of the household head, gender of the household head and whether the household head's occupation is agriculture. Under the hypothesis that there were no compositional changes, we would expect β_2 to be 0.

I find the difference in difference is not significant for all the covariates across all treatment thresholds except for the covariate for gender of the respondent and the likelihood of having a male head. However, even the likelihood of having a male head is significant only at the 10% (see Table A13 in the Appendix). I take this as evidence that there were no compositional changes on those observables over time between treatment and control areas. I include gender as a control in all regressions and also do the analysis separately by gender.

2.6.4 *Alternative measures of exposure*

2.6.4.1 *Continuous distance*

To exploit the full variation in distance to camps, I re-estimate the baseline specification 2.1 using the continuous distance from an individual's village to the nearest camp as an alternative to the discrete distance thresholds. I also use the log of the distance as another alternative measure.

For the older generation, the effect of having been exposed to camps in childhood is negative and statistically significant (see Table A14 in the Appendix). Note the coefficients when using continuous distance and log distance are positive because the relationship between HAZ and distance is inverse; the negative effect is reduced as one moves away from the camp. To interpret the result more easily, we can consider the effect of moving 10km *away* from the nearest camp. The result suggests that this is associated with a small improvement in HAZ of 0.008.[7] The effect, however, is only statistically significant at the 10%.

Although when using continuous distance, the effect of having been exposed to camps in the younger generation, is negative, it is smaller in magnitude to that in the older generation (see Table A15 in the Appendix).

2.6.4.2 *Distance bands*

To examine the localisation of the effect suggested by the results, I classify distance to the nearest camp into three categories namely, 0-20 kilometres, 20-100 kilometres and greater than 100 kilometres. Under this classification, 18 villages lie within 20km of a camp, 37 villages lie between 20 - 100km of a camp, and 139 villages are more than 100km but less than 430km from a camp. I use this classification to assess the effect associated with being near, or at an intermediate distance to a camp, relative to being far from a camp.

I use the following equation, which is a variant of baseline equation 2.1[8]:

$$HAZ_{ivk} = \beta_0 + \beta_1 (distance0\_20km * Child\_open) + \beta_2 (distance20\_100km * Child\_open) + \beta_3 male_i + \delta_k + \theta_j + e_{ivk} \qquad (A.3)$$

The results also suggest that the effect was localised (See Table A16 in the Appendix). The difference between those who were nearest and farthest is a reduction in HAZ of 0.36. The difference between those that are at the intermediate distance (20-100 km) and farthest (more than 100km) is lower at 0.2. I also tested the p-values for equality between the distance

---

[7] The calculation is obtained by multiplying the coefficient on the interaction term of $distance_v * Child\_open$ by 10km.

[8] Piece-wise linear splines could be another potential way to model the relationship between HAZ and distance in this context.

categories. Specifically, the test for whether there is a differential effect between those in 0-20km and 20-100km is significant. The results indicate a decreasing gradient – exposure was associated with negative impact on the HAZ and the effect is stronger for areas closest to camps relative to those farthest away.

I also estimate the analogous regression for the younger generation where the indicator variable denotes whether the nearest camp was operational at some point in childhood, or whether the individual was born after the camp closed. In contrast to the result from the older generation, in the younger generation I find no statistically significant effect of camps on those who were closest (0-20km), relative to those who were farthest (more than 100km), or on those at an intermediate distance relative to those farthest (see Table A17 in the Appendix). This finding corroborates the previous results of a lack of statistically significant effect in the younger generation using either the distance thresholds or intensity index.

2.7 Conclusion

This paper investigated the long-term impact of refugee camps on the health of local residents. First, it considered individuals who were children at the beginning of the refugee influx and examined the effect on their health as adults. The findings show that for this generation, camp exposure has had a negative impact on their adult health. While the existing literature shows a negative effect in childhood, this paper provides evidence that this effect lasted into adulthood.

I find that camp exposure led to a decline of between 0.2 and 0.4 standard deviations in HAZ (a reduction of 1.38–2.76 cm). Using existing estimates of the association between height and labour market outcomes, a simple calculation would translate the effect to a 2.9%–5.9% reduction in adult hourly earnings.

When it comes to the younger generation, however, those who were exposed to the refugee camps do not fare any differently from those born in the post-camp era. This paper provided some evidence that the initial negative effect on child health was through disease outbreaks in early camp phases. Over the years and by the time camps closed, these issues had been brought under control, hence those exposed to the camps in the younger generation are not differently affected from those born in the post-camp period.

The results are robust to different specifications, measures of exposure and a control for local pre-camp socioeconomic conditions. Nonetheless, the identification strategy is not able to completely address the possibility of unobserved time-variant characteristics that may be related to the camps. The results presented are also conditional on survival, since only living individuals are observed and results are, therefore, likely to be a lower bound estimate of the actual effect (Harttgen et al., 2019). Similarly, because self-reported age is used to calculate HAZ, the results obtained are likely to be an underestimate of the true effect (Akresh et al., 2011).

The initial stages of refugee inflows are typically characterized by the same conditions that were experienced in Tanzania. The health effect on the local population at the outset of camp establishment is thus likely to be generalizable to other refugee camp contexts in developing countries. However, since the subsequent evolution of the camp to the hosting community can vary between contexts, the extent to which spill overs from camp presence mitigate the initial effect could differ. Additionally, due to data constraints it is not possible to analyse the potential role played by camp rehabilitation post-closure.

Nevertheless, the findings of this paper provide insights for policy in the current climate of unprecedented refugee flows and in view of the climate-induced forced displacement that is predicted to occur in the future. A major challenge currently confronting several countries, and one that is likely to remain relevant, is how to sustainably support host regions. This paper's findings demonstrate that the onset of forced displacement flows is a critical window – conditions at the onset could have persistent effects. Early implementation of public health measures is crucial. Due to data limitations, this paper focused on the health of those in the vicinity of refugee camps; an equally important and under-researched avenue for future research is the health of refugees.

2.8 Appendix

Table A1: Summary statistics of Intensity Index (Entire sample)

	mean	sd.	min	max
Intensity	0.53	0.77	0.00	5.65
Population adjusted intensity	0.02	0.02	0.00	0.20
Observations	3934			

Table A2: Summary statistics of Intensity Index (Adult sample)

	mean	sd.	min	max
Intensity	0.61	0.85	0.00	5.65
Population adjusted intensity	0.02	0.02	0.00	0.20
Observations	1313			

Table A3: Summary statistics of Intensity Index (Younger generation sample)

	mean	sd.	min	max
Intensity	0.48	0.72	0.01	5.59
Population adjusted intensity	0.01	0.02	0.00	0.15
Observations	2621			

Table A4: Least exposed cohort versus average exposed cohort (Intensity index)

		coefficient	Mean of intensity Index	Minimum of intensity index	Difference in intensity	Associated change in HAZ
1 unit increase in intensity (Adult sample)	Table 2.7	-0.147	0.61	0.00	0.61	-0.09
1 unit increase in intensity (Young sample)	Table 2.10	-0.102	0.48	0.01	0.47	-0.05

Table A5: Diarrhoea prevalence by camp phases (DHS surveys)

	Diarrhoea
treat100	0.261***
	(0.0699)
treat100*campphase2	-0.117**
	(0.0513)
treat100*campphase3	-0.153**
	(0.0590)
Observations	2,361
R-squared	0.086
Region fixed effects	Yes

Note: Data is from DHS 1999 and 2010. Sample comprises 0-5-year olds. "treat 100" is a dummy variable that takes 1 for individuals within 100 km of a camp and 0 otherwise. "campphased" is a dummy that takes 1 if the nearest camp is in phase d and 0 otherwise. The reference group is campphase1 (recent camp). Specification includes a constant term, year of birth fixed effects, controls for child's gender, household head's gender, mother's age and education, household size, average annual temperature and rainfall in the cluster and survey round fixed effects. Robust standard errors in parentheses, clustered at cluster level. *** p<0.01, ** p<0.05, * p<0.1

Table A6: Health services utilisation by camp phase (DHS surveys)

	(1) Delivered in a health facility	(2) Fully vaccinated
treat100	0.163	0.189**
	(0.251)	(0.0929)
treat100*campphase2	-0.0294	0.0386
	(0.212)	(0.0685)
treat100*campphase3	-0.212	-0.0826
	(0.213)	(0.0813)
Observations	2,562	2,379
R-squared	0.073	0.096
Region fixed effects	Yes	Yes

Note: Data is from DHS 1999 and 2010. Column 1 sample comprises 0-5-year olds. Column 2 sample comprises 1-5-year olds. "treat 100" is a dummy variable that takes 1 for individuals within 100 km of a camp and 0 otherwise. "campphased" is a dummy that takes 1 if the nearest camp is in phase d and 0 otherwise. The reference group is campphase1 (recent camp). Specification includes a constant term, child age fixed effects and controls for child's gender, household head's gender, mother's age and education, household size, average annual temperature and rainfall in the cluster and survey round fixed effects. Robust standard errors in parentheses, clustered at cluster level. *** p<0.01, ** p<0.05, * p<0.1

Table A7: Long-term effect of camp exposure, by distance bands from nearest camp (By gender)

	HAZ Male	HAZ Female
(0-20km) * childopen	-0.351**	-0.383**
	(0.175)	(0.172)
(20-100km) * childopen	-0.141	-0.354**
	(0.145)	(0.155)
Observations	606	707
R-squared	0.187	0.134
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Notes: Adult sample. All models include a constant term. The reference group is those more than 100km away from a camp. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A8: Impact of camp exposure on younger generation, by distance bands from nearest camp (By gender)

	HAZ Male	HAZ Female
(0-20km) * C_i	-0.180	-0.475
	(0.261)	(0.296)
(20-100km) * C_i	-0.225	-0.364
	(0.247)	(0.245)
Observations	1,239	1,382
R-squared	0.173	0.159
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Notes: Younger generation sample. C_i is a dummy that takes 1 if the nearest camp was operational at some point in childhood and 0 otherwise. All models include a constant term. The reference group is those more than 100km away from a camp. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A9: Long-term effect of camp exposure, by distance bands from nearest camp (Non-migrants)

	HAZ Full	HAZ Non-migrants
(0-20km) * childopen	-0.362**	-0.654***
	(0.143)	(0.190)
(20-100km) * childopen	-0.234**	-0.593***
	(0.096)	(0.178)
Male	-0.316***	-0.216**
	(0.048)	(0.087)
Observations	1,313	534
R-squared	0.148	0.222
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Notes: Adult sample. All models include a constant term. The reference group is those more than 100km away from a camp. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A10: Impact of camp exposure on younger generation, by distance bands from nearest camp (Non-migrants)

	HAZ Full	HAZ Non-migrants
(0-20km) * C_i	-0.315	-0.361
	(0.276)	(0.289)
(20-100km) * C_i	-0.266	-0.350
	(0.234)	(0.249)
Male	-0.308***	-0.316***
	(0.0453)	(0.0495)
Observations	2,621	2,208
R-squared	0.154	0.155
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Notes: Younger generation sample. C_i is a dummy that takes 1 if the nearest camp was operational at some point in childhood and 0 otherwise. All models include a constant term. The reference group is those more than 100km away from a camp. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A11: Placebo test

	HAZ
Dist. Nearest camp	0.0011
	(0.001)
Dist. Nearest camp * Placeboafter	0.0002
	(0.001)
Male	-0.3054***
	(0.102)
Observations	305
R-squared	0.110
Birth year fixed effects	Yes
Region fixed effects	Yes

Note: The individuals in this sample are those in subgroup 1 in exposure Table 2.1. All models include a constant term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A12: Long-term impact of exposure to camps on adults who were children in 1993 (Controlling for baseline socioeconomic conditions)

	HAZ				
	(1)	(2)	(3)	(4)	(5)
treat20*childopen	-0.395**				
	(0.162)				
treat50*childopen		-0.280**			
		(0.127)			
treat80*childopen			-0.191*		
			(0.115)		
treat100*childopen				-0.236**	
				(0.106)	
treat170*childopen					-0.084
					(0.104)
Male	-0.315***	-0.313***	-0.309***	-0.310***	-0.316***
	(0.049)	(0.048)	(0.049)	(0.049)	(0.049)
Observations	1,276	1,276	1,276	1,276	1,276
R-squared	0.145	0.145	0.146	0.148	0.156
Birth year fixed effects	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes

Note: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. “treat d” is a dummy variable that takes 1 for individuals within d km of a camp and 0 otherwise. All models include a constant term and the base treatment term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A13: Balance test

	DID (treat20*Ci)	DID (treat50*Ci)	DID (treat80*Ci)	DID (treat100*Ci)	DID (treat170*Ci)
Household size	-0.635	0.0728	-0.353	-0.312	0.168
Child is male	-0.139*	-0.0308	-0.102**	-0.130***	0.0148
Rural household	0.0221	-0.0313	0.0503	-0.0485	0.0195
Education of household head	0.419	-0.28	-0.0967	-0.0519	0.0267
Age of household head	3.179	1.61	2.57	1.775	0.194
Household head is male	0.00914	0.127*	0.0949*	0.0834	0.0275
Household head's occupation is agriculture	-0.0774	-0.0498	-0.00529	0.00331	0.0652

Note: The individuals in this sample are those in subgroup 3 in exposure Table 2.1. All models include a constant term, treat d level effect, region fixed effects and year of birth fixed effects. Ci indicates if individual was born before or after camp closed. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A14: Long-term effect of camp exposure, by continuous distance to the nearest camp

	HAZ	
	(1)	(3)
Dist. Nearest camp	0.0015***	
	(0.001)	
Dist. Nearest camp*childopen	0.0008*	
	(0.000)	
log Dist. Nearest camp		0.1332**
		(0.053)
log Dist. Nearest camp *childopen		0.0540**
		(0.026)
Male	-0.3201***	-0.3165***
	(0.047)	(0.047)
Observations	1,313	1,313
R-squared	0.159	0.153
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Note: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. All models include a constant term. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A15: Impact of exposure to a camp in childhood on younger generation, by continuous distance to the nearest camp

	HAZ	
	(1)	(2)
Dist. Nearest camp	0.0004	
	(0.001)	
Dist. Nearest camp*C_i	0.0006**	
	(0.000)	
log Dist. Nearest camp		0.1074
		(0.083)
log Dist. Nearest camp*C_i		0.0128
		(0.014)
Male	-0.3030***	-0.3035***
	(0.045)	(0.045)
Observations	2,621	2,621
R-squared	0.153	0.154
Birth year fixed effects	Yes	Yes
Region fixed effects	Yes	Yes

Notes: The individuals in this sample are those in subgroup 3 in exposure Table 2.1. C_i is a dummy that takes 1 if the nearest camp was operational at some point in childhood and 0 otherwise. All models include a constant term. Robust standard errors in parentheses, clustered at village level. ***p<0.01, ** p<0.05, * p<0.1

Table A16: Long-term effect of camp exposure, by distance bands from nearest camp

	HAZ
(0-20km) * childopen	-0.362**
	(0.143)
(20-100km) * childopen	-0.234**
	(0.096)
Male	-0.316***
	(0.048)
Observations	1,313
R-squared	0.148
Birth year fixed effects	Yes
Region fixed effects	Yes

Notes: The individuals in this sample are those in subgroups 1 and 2 in exposure Table 2.1. All models include a constant term. The reference group is those more than 100km away from a camp. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Table A17: Impact of camp exposure on younger generation, by distance bands from nearest camp

	HAZ
(0-20km) * C_i	-0.315
	(0.276)
(20-100km) * C_i	-0.266
	(0.234)
Male	-0.308***
	(0.0453)
Observations	2,621
R-squared	0.154
Birth year fixed effects	Yes
Region fixed effects	Yes

Notes: The individuals in this sample are those in subgroup 3 in exposure Table 2.1. C_i is a dummy that takes 1 if the nearest camp was operational at some point in childhood and 0 otherwise. All models include a constant term. The reference group is those more than 100km away from a camp. Robust standard errors in parentheses, clustered at village level. *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

Predicting local economic activity and welfare from daytime satellite data

3.1 Introduction

Data gaps continue to be a challenge in developing countries (Beegle et al., 2016; Dang et al., 2019; Devarajan, 2013; Serajuddin et al., 2015). The scarcity of data is more acute at sub-national levels where regular data on measures of economic activity and welfare are usually lacking or collected infrequently. Yet, the ability to track poverty, design targeted policy responses, study the impacts of policies and events on welfare, spatial inequalities and other such outcomes, all hinge on the availability of regular sub-national data. Household surveys, the conventional data source on such outcomes, are expensive to collect (Kilic et al., 2017).

Satellite data has enormous potential for economics applications in such data-scarce contexts. Satellites collect panel data at low marginal-costs, have high spatial resolution and wide geographic coverage (Donaldson & Storeygard, 2016). Early applications of satellite data in economics introduced nightlights as a proxy for economic activity (Chen & Nordhaus, 2011; Henderson et al., 2012). Nightlights, however, have some limitations in developing countries where low levels of luminosity cannot be distinguished from background noise. Yet, it is precisely these poor areas for which data scarcity is a problem (Chen & Nordhaus, 2011).

A nascent line of enquiry investigates whether machine learning techniques applied to daytime satellite data could be better for predicting subnational economic well-being in developing countries (Engstrom et al., 2017; Goldblatt et al., 2019; Jean et al., 2016). This work generally finds that daytime data can better capture variation in economic well-being better than nightlight data. However, the literature is still inconclusive on whether the predictive power of daytime data holds in other contexts, for which outcomes, and whether prediction can be improved by combining with other routinely collected sources of data such as weather variables. In addition, most of the existing work focusses on predicting local economic well-being in a cross-sectional setting. There is a need for studies evaluating the ability to predict over time.

In this paper, I evaluate the ability to predict local welfare and economic activity in Tanzania from the Landsat satellite - the longest running earth observation satellite mission. Since Landsat data spans many years, it has potential for use to fill data gaps. Besides examining predictive ability for unseen locations at one snapshot in time, I evaluate the ability to predict across time i.e., to use a model trained in one time period to predict data for a time period where corresponding survey data may not be available. This would facilitate the study of economic outcomes in both areas and time periods where conventional data sources are not available. I find that features from Landsat can explain substantial variation in the share working in agriculture, as well as expenditure and assets from household surveys. In all instances, using a built-up variable extracted from the satellite imagery has higher predictive ability than using either spectral indices or flexible specifications that use characteristics of the distribution of the imagery. The model can explain 62% of the cluster level variation in the share working in agriculture, 45% of the variation in expenditure, and 51% of the variation in asset ownership in 2012. Similar results are obtained for 2008. Models trained on survey data from a specific year can also predict consumption expenditure in another time period reasonably well, but less so for agricultural occupation.

3.2 Literature review

In the absence of data on economic variables, one common approach is to adopt satellite data on night time lights as a proxy for local level economic activity or welfare (Hodler & Raschky, 2014; Mamo et al., 2019; Michalopoulos & Papaioannou, 2013, 2014). Some studies have empirically assessed the relationship between nightlights and economic activity at a national level. Chen and Nordhaus (2011) and Henderson et al., (2012) show that nightlights correlate with economic output at a country level. The assessment of the potential of nightlights to generate local wealth and welfare predictions for out-of-sample locations has received less attention (Weidmann & Schutte, 2017). However, at a sub-national level, nightlights tend to be less capable of picking up variation in economic activity for poor regions where the luminosity levels are usually very low (Chen & Nordhaus, 2011; Jean et al., 2016).

Advances in machine learning techniques have increased interest in predicting local level economic variables from other non-conventional data sources (Athey, 2019; Kleinberg et al., 2015). A growing body of literature investigates whether machine learning methods applied to non-conventional data sources could be used to predict poverty at the local level especially in developing countries. Prediction in this context refers to a setting where we have a dataset ("training" dataset) with covariates (or "features") and outcomes for a set of observations (e.g. villages). The goal is to use this data to build a model that obtains (i.e., predicts) the outcome for other unseen observations (Hastie et al., 2009: p.1). In the poverty example, the goal would be to obtain the poverty measure for other villages for which we do not have data on poverty. This type of work is important as it can facilitate analysis of economic activity and well-being at a local level (Glaeser et al., 2018).

Blumenstock et al., (2015) use data from mobile phone records to predict individual level socioeconomic status and wealth in sub-national localities in Rwanda. The applicability of this approach, however, is limited by its reliance on proprietary and confidential data. Findings from recent work suggest that daytime satellite imagery offers a promising alternative. Daytime imagery captures information about landscape features that are likely to be indicative of the variation in poverty in low-income countries (Goldblatt et al., 2019; Jean et al., 2016). Jean et al., (2016) train a convolutional neural network model to extract features from high resolution daytime satellite imagery and then use the extracted features in a second step to train ridge regression models to predict expenditure and assets from survey data. The model explains 37 to 55% of the variation in average household consumption from survey data and 55 to 75% of the variation in average household asset wealth in five African countries. Engstrom et al., (2017) extract features from satellite images of Sri Lanka and use the features to estimate village level poverty rates and average consumption. Estimated models explain about 60% of the variation in poverty and consumption across villages. In subsequent decomposition analysis, the authors find that measures of building density – built-up area and number of buildings, are particularly strong predictors of welfare. The authors also suggest that where values are not predicted with high accuracy, there is still value if the predictions preserve the rank order in the actual data. The authors explain that, this is because there are applications, such as for the purposes of targeting interventions, for which it may suffice to only know the rank ordering of locations by welfare. To examine to what extent the rank ordering of villages by predicted values corresponds to the rank ordering by actual values, the authors examine the correlation in the rank ordering between predicted and actual values. Observed and predicted values are ranked and the correlation between the rankings (the Spearman's rank-order correlation) is obtained. The Spearman's rank correlation coefficient indicates how well the predictions order the observed values, regardless of the accuracy of the predictions. A correlation coefficient of 1 would indicate that the model is able to order the data perfectly. The authors find a Spearman's ρ of between 0.67 and 0.7. However, the authors rely on costly commercial satellite imagery. If, as mentioned, one advantage of satellite data is reduced data collection costs, then the high costs of commercial satellite imagery remain prohibitive for developing countries. Additionally, both Jean et al., (2016) and Engstrom et al., (2017) use data that is limited in temporal availability and consequently, prediction was limited to cross-sectional settings.

Investigating such predictive capability across time necessitates that both the alternative data source and survey data for validation be available for multiple years. Goldblatt et al., (2019) propose using publicly available imagery data from an earth observation satellite - the Landsat - which has been monitoring the earth's surface since the early 1970s to date. The authors examine its potential for predicting the distribution of enterprises and household consumption across small geographic units (communes) in Vietnam. The authors conclude that Landsat has promising potential (it is able to explain 40-50% of the variation in economic activity and expenditure) but that its applicability in other contexts as well as for changes across time warrants continued investigation. This paper will continue this line of enquiry started by the authors.

Overall, the findings from the few existing studies raise important considerations, particularly whether results hold in other contexts and whether it is possible to predict over time. This paper seeks to contribute an understanding of these questions by examining the use of Landsat data to predict economic variables at a local level in Tanzania from two waves of household survey data. The variables of interest at the cluster level are the proportion of households having agriculture as the main economic activity, log per adult equivalent consumption expenditure, asset ownership and the proportion of households in extreme poverty.

A related strand of literature is concerned with the mapping of urban land cover. These works focus on delineating urban land from Landsat imagery (Baragwanath et al., 2019; Goldblatt, Deininger, et al., 2018; Goldblatt et al., 2016; Goldblatt, Stuhlmacher, et al., 2018). I will take built-up a predictive feature and evaluate its predictive ability comparing models that use it with specifications that use the characteristics of the imagery.

3.3 Data construction

I combine data from several sources: household surveys, Landsat imagery and georeferenced datasets on weather and geophysical characteristics.

3.3.1 Survey data

This analysis uses two waves of a geocoded nationally representative household panel survey – the 2008 and 2012 Tanzania Living Standards Measurement Survey (LSMS) to obtain cluster-level measures of economic activity and welfare that will be used to train and evaluate the accuracy of the models. Data collection for the first survey ran from October 2008 to

October 2009 and the second survey ran from October 2012 to October 2013. I will use images acquired from the satellite passes during the two time periods.

Since I will conduct analysis at the level of the cluster, I obtain cluster-level variables by taking the average over the households within the cluster. There are 409 clusters. In the 2012 wave, it is important to exclude from the clusters households that were tracked and found to no longer be in their 2008 locations. This ensures that in both survey waves, cluster averages are from households that are in that location at the time of the survey. Because of this, in 2012 clusters have on average fewer households than in 2008. The average cluster size is 8 households in 2008 and 7 households in 2012.

For a survey-based measure of the type of economic activity in the cluster, I calculate and use the proportion of households for whom agriculture is the main economic activity. For welfare, I use three different metrics. First, I use the log of daily per adult equivalent consumption expenditure. Consumption is the preferred measure of welfare since it has been found to be a more useful and accurate indicator of living standards than income (Deaton & Zaid, 2002; Haughton & Khandker, 2009). For each of the survey waves and for each household, I calculate the per adult equivalent consumption expenditure by dividing total household consumption expenditure by the number of adults in the household, and then average within the cluster. I then convert from the local currency to 2011 PPP dollars for consistency across the two years. I choose 2011 as the common currency year for ease of comparison to the international poverty line of $1.90 a day which is set at 2011 PPP prices. As a second welfare measure, I use the per adult equivalent consumption expenditure that I calculated in the preceding step to determine whether a household is above or below the international poverty line of $1.90 a day. I then calculate the proportion of households in a cluster that is in extreme poverty, defined as being below the international poverty line.

The third welfare measure of interest is a variable on asset ownership. Since LSMS surveys do not report an asset index, to construct it I follow a similar approach to that used by the demographic health surveys. Specifically, I use the households' responses to questions on assets owned and the reported value of the assets and apply principal components analysis. Households are categorized into wealth quintiles based on the distribution of the wealth score. A cluster is assigned the average of the quintile values of the households in the cluster. The 2008 survey did not include detailed questions on assets and as such, it is only possible to construct the asset variable for 2012. Summary statistics for these outcome variables of interest are available in Tables B1.1-B1.2 in the Appendix. These survey variables will be used to train and evaluate the model's accuracy.

Geographic coordinates of the cluster centroids are used to match the survey data to the Landsat data. To maintain the confidentiality of respondents, cluster centroids are reported as the average latitude and longitude of households in that cluster, plus an added random displacement typically between 2 to 5km. To ensure that the real cluster centroid is within the area used, I draw a 5km circular buffer around the cluster centroids; satellite variables will be obtained within the buffers.[9]

3.3.2 Landsat data

The Landsat mission, operated jointly by the National Aeronautics and Space Administration (NASA) and the United States Geological Survey agency (USGS), has been collecting data

[9] This approach is used by Weidmann & Schutte (2017) to match nightlight data onto DHS clusters.

on the earth's surface from 1972 to date. For the survey periods, the relevant satellite is the Landsat 7 (1999-present). It has a resolution of 30 metres. The Landsat technology relies on the fact that when sunlight hits the earth's surface, objects absorb and reflect light differently. The satellite detects different ranges of frequencies along the electromagnetic spectrum. The data is collected and is available as images in 8 spectral bands. Since Landsat has a revisit period of 16 days, there are on average 22 images a year. I use the image collection that has been calibrated by the USGS for sensor degradation and corrected for atmospheric interference. These are considered suitable for time series use (Chander et al., 2009; Earth Engine Data Catalogue, 2020; Thome et al., 1997; U.S. Geological Survey (USGS), 2019a, 2019b; Young et al., 2017). I extract several features from the imagery that will then be used as predictive variables in the model.

I merge the cluster buffers with the image collection from each of the time periods and then calculate cluster-level characteristics of the bands, namely their means, medians and standard deviations. I will use the band characteristics and their interactions as described in the methodology section below.

In addition, I calculate spectral indices. Spectral indices are derived from spectral bands and are transformations used to characterize specific land use categories. The established indices are the Normalized Vegetation Index (NDVI), the Normalized Built-up Index (NDBI) and the Normalized Water Index (NDWI). Each index was developed as a combination of bands that were found to capture the spectral signature of one of three classes - vegetation, built-up land and water respectively (Rouse et al., 1973; Young et al., 2017; Zha et al., 2003). These indices are calculated as:

$$NDBI = \frac{band5 - band4}{band4 + band5} \quad NDVI = \frac{band4 - band3}{band4 + band3} \quad NDWI = \frac{band2 - band4}{band2 + band4}$$

Their values range from -1 to 1 with a higher index value indicating more built-up, vegetation or water respectively. Obtaining the value of the indices for a cluster involves several steps. For each pixel in the cluster, I calculate the median value of each band from the images for that period. I use the median of band values to calculate the indices. Finally, to get the cluster value, I take the average of all pixels falling within each cluster.

The last predictive feature from the Landsat is a variable on built-up. I obtain this variable from the World Settlement Footprint (WSF) dataset of the German Aerospace Centre (DLR). The WSF is a map of built-up land at a global scale that has been produced from Landsat imagery. To classify land as built-up or not, the DLR used samples from a variety of locations globally to train classifiers, and ensure data was representative of diverse settlement patterns. Built-up was taken to include buildings, building lots, roads or paved surfaces (Marconcini et al., 2019). The WSF map labels pixels as built-up or not. I merge the WSF map with the cluster locations and calculate the size of the built-up area in sq. km. in each cluster for the two years of interest.

3.3.3 Additional data (weather and geophysical characteristics)

I also use rainfall data and the slope and elevation of the cluster as they are indicators of the type of economic activity possible in an area and, therefore, likely to improve the ability to predict economic activity and wealth in the clusters. Since I do the analysis for all of Tanzania, I also include a variable for whether the cluster is on the mainland or in the Zanzibar archipelago. Rainfall data is taken from the University of Idaho's *TerraClimate* dataset. I calculate the average annual rainfall for a 5-year window preceding each survey period i.e., the average rainfall between 2004-2008 and 2008-2012. Elevation and slope are obtained from the NASA Shuttle Radar Topography Mission (SRTM) dataset.

3.4 Methodology

The objective is to use the satellite features to predict the survey variables of interest, namely the proportion of households having agriculture as the main economic activity, log per adult equivalent consumption expenditure, asset index and the proportion of households in extreme poverty.

3.4.1 Cross-section

I begin with a cross-sectional setting, to estimate the outcomes in the 2008 survey using satellite features from the same time period. I compare several models. In all cases, I use weather and geophysical characteristics. What differs across models is the information from the Landsat that is used. First, I begin with the built-up area. Recall from the literature review that built-up area was found to be a particularly strong predictor of welfare in other contexts. I examine how well it can predict in this context. Second, I use spectral indices. Spectral indices can be thought of as striking a balance between using variables that requires extensive ground truth data to classify – built-up area in this case – and raw Landsat bands which are difficult to interpret (Goldblatt et al., 2019). However, as the authors aptly note, one may hypothesise that, since the Landsat indices are a combination of some and not all Landsat bands, the predictive accuracy could be improved by using all available information i.e., using all bands. In a similar manner, I investigate whether more flexible models that use all Landsat bands and characteristics of the distribution of the band values and their interactions have better predictive accuracy. I use means, medians and standard deviations of all bands, interactions and polynomials resulting in flexible but high dimensional models. Since the interest is in prediction, standard statistical methods would not be suitable. For example, OLS tends to produce predictions with low bias but large variance thereby providing good within-sample prediction but poor out-of-sample prediction. Additionally, in high dimensional settings where the number of predictor variables exceed the number of observations (p > N), it is not possible to use least squares. Shrinkage or regularization methods can handle these issues. Prediction accuracy is improved by shrinking or setting some coefficients to zero. By doing so, one trades off bias to reduce the variance of the predicted values, and hence improve the overall prediction accuracy. I follow the approach in the literature and use the least absolute shrinkage and selection operator (LASSO) and cross-validation to select variables and optimise out-of-sample prediction. The LASSO minimises the residual sum of squares subject to a constraint on the absolute size of coefficient estimates and a penalty term that penalises the size of the model (Belloni et al., 2014; Hastie et al., 2009; Tibshirani, 1996).

The LASSO estimator is defined as:

$$\hat{\beta} = \underset{\beta}{arg\ min} \left[\sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right]$$

Where y_i is the cluster-level outcome variable of interest, $\lambda > 0$ is the lasso penalty term (“tuning parameter”) and x_{ij} are the potential covariates. As λ increases the magnitude of estimated coefficients is shrunk toward zero. Covariates with estimated coefficients of zero are thus excluded.

I proceed as follows: I randomly split the LSMS clusters into a training (286 clusters i.e., 70% of the data) and validation data set (123 clusters i.e., 30% of the data). The training data is used to estimate the model parameters while the validation data is used to evaluate the model's out-of-sample predictive accuracy. To determine the LASSO parameters that minimise out-of-sample prediction error (the 'tuning' parameters), I use 5-fold cross-validation. In this procedure, the training data is randomly divided into five folds of approximately equal sizes. In the first instance, fold 1 is treated as the validation data set while folds 2-5 constitute the training data. The model is estimated using the training data (folds 2-5) and predictive performance assessed using fold 1. In the next step, fold 2 is treated as the validation data while the other folds are the training data. The process is repeated each time using a different fold as the validation data and the remaining folds as the training data. The LASSO parameter that minimises the mean-squared prediction error is then identified and used to estimate the model.[10]

3.4.2 Over time

After examining the ability for Landsat to predict welfare and economic activity in a cross-sectional setting, I turn to examining whether the models can generalise across time. Survey data is often collected infrequently or after long time gaps. If a model trained and evaluated on a year when corresponding survey data was available could be used to obtain predictions for other time periods where survey training data is not available, this could be insightful in filling such data gaps. To examine this, I use both rounds of the survey. The 2008 data is used to generate the prediction model which is then applied to the 2012 Landsat data using the estimated parameters from 2008, in order to obtain estimates of the outcome variable that are compared to the actual variables from the 2012 survey to evaluate accuracy. Then I do the reverse i.e., train the model in 2012 and use it to predict backwards in 2008. The accuracy of predictions from models trained in one year and applied in another year will indicate how

[10] A discussion of the k-fold cross validation method is available in (Hastie et al., 2009)

well the model is able to generalise across time. Douidich and Ezzrari (2016) and Mathiassen (2013) apply this forward and backward approach in cross-survey imputation using data from one survey to predict poverty rate from other surveys. However, since their approach relies on survey data for prediction, it still requires the availability of survey data in the target year – this would not be possible in situations where there was no survey data at all. I use publicly available and routinely collected satellite data for the prediction.

3.5 Results and Discussion

Before presenting the results, it is helpful at this juncture to recall the objective of prediction tasks compared to standard inference tasks. In inference, one is interested in understanding the relationship between an outcome and predictor variables; interpretability is thus highly desirable. On the other hand, in a machine learning context, where we are mainly interested in obtaining predictions for new or unseen observations, the goal is to obtain the model that gives the highest accuracy (lowest test set error) regardless of the form of inputs or of the interpretability of the model. There can, thus, be a trade-off between interpretability and predictive accuracy. It is for this reason that a criticism of machine learning techniques is that they are "black boxes" – able to predict well but at the cost of interpretability. In this case, an advantage of the LASSO method is that it allows us to retain some interpretability. Specifically, by selecting only the best predictors, it produces parsimonious models and thus lends insight into what variables are important predictors.

To compare the predictive accuracy across models of the same dependent variable, I use the Root Mean Squared Error (RMSE). The RMSE is a measure of the average deviation of the predictions from the observed values. Lower values of RMSE are preferable as they indicate higher predictive accuracy. Additionally, an important consideration for model assessment is "overfitting", that is, a situation where a model captures patterns in the data very well because it also treats random noise as informative. Flexible models are particularly prone to overfitting. While a flexible model may fit the data well, the fit obtained may not yield accurate estimates of the outcome variable for new observations that were not part of the training data set. As such, the metric of interest for model comparison for predictive purposes, is the test set error (out-of-sample RMSE in this case) as it indicates how the model performs on new data. In other words, for a well-fitted model, the RMSE in the training data (in-sample) should be similar to that in the test data (out-of-sample RMSE).

I now turn to the cross-sectional results first followed by the results for models trained on satellite data from another year.

3.5.1 Cross-section

Tables 3.1A-3.1D below show the comparison of results from the various models for each outcome variable of interest. Weather and geophysical characteristics are included as predictors in all models. Column 1 shows the results from using built-up area whereas column 2 shows the results from using the spectral indices. Column 3 shows the results from using the means of all Landsat bands while columns 4-6 show the results from using all bands, interactions of band characteristics and polynomials.

Below, unsurprisingly, the out-of-sample RMSE is higher than the in-sample RMSE, indicating some overfitting. However, the overfitting is worse in the more flexible models. Comparing the out-of-sample RMSEs across the columns, the results indicate that the models using built-up area as a predictive feature (column 1 in all Tables) perform best (i.e., have the lowest RMSE). The more flexible models have a higher out-of-sample error. This suggests that the predictive information from Landsat is encompassed in the built-up area feature. Although the indices yield more accurate predictions than using all the information in the Landsat bands, they are also outperformed by using the built-up area feature. The model with the lowest test set error explains about 62% of the cluster level variation in the share working in agriculture (Table 3.1A), 45% of the variation in log per adult equivalent expenditure (Table 3.1B) and 51% of the variation in asset ownership (Table 3.1C) and 22% of the proportion of households in extreme poverty (Table 3.1D) in the test set for the 2012 cross-section.

Table 3.1A: Predicting share working in agriculture (2012)

Dependent variable	**Proportion of hhs in the cluster for whom agriculture is the main economic activity**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	4	5	10	16	16	4
In-sample RMSE	0.234	0.291	0.259	0.250	0.249	0.309
Out-of-sample RMSE	0.238	0.317	0.280	0.273	0.275	0.316
In-sample R-squared	0.671	0.488	0.596	0.623	0.627	0.424
Out-of-sample R-squared	0.622	0.331	0.478	0.505	0.495	0.330
Penalty parameter (λ)	0.018	0.007	0.000	0.001	0.006	0.108

Table 3.1B: Predicting expenditure (2012)

Dependent variable	**Per adult equivalent consumption expenditure (log)**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	5	5	11	18	18	3
In-sample RMSE	0.329	0.387	0.349	0.338	0.337	0.403
Out-of-sample RMSE	0.366	0.416	0.391	0.387	0.387	0.447
In-sample R-squared	0.585	0.425	0.534	0.562	0.563	0.376
Out-of-sample R-squared	0.446	0.282	0.366	0.379	0.381	0.175
Penalty parameter (λ)	0.003	0.012	0.0001	0.001	0.005	0.155

Table 3.1C: Predicting assets (2012)

Dependent variable	**Assets**				
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257
No. of variables selected by LASSO	4	4	11	14	18
In-sample RMSE	0.704	0.830	0.750	0.736	0.694
Out-of-sample RMSE	0.710	0.888	0.829	0.834	0.837
In-sample R-squared	0.535	0.354	0.472	0.492	0.548
Out-of-sample R-squared	0.509	0.231	0.330	0.322	0.317
Penalty parameter (λ)	0.008	0.051	0.0002	0.009	0.012

Table 3.1D: Predicting proportion of households in extreme poverty (2012)

Dependent variable	**Proportion of households in extreme poverty**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	3	3	11	14	20	13
In-sample RMSE	0.179	0.189	0.178	0.172	0.167	0.174
Out-of-sample RMSE	0.192	0.205	0.193	0.203	0.202	0.211
In-sample R-squared	0.305	0.225	0.318	0.362	0.401	0.344
Out-of-sample R-squared	0.217	0.110	0.206	0.123	0.132	0.054
Penalty parameter (λ)	0.006	0.014	0.000	0.002	0.003	0.008

Similar results are obtained when using the 2008 survey wave (Tables 3.2A-3.2C). Using built-up area, weather and geophysical characteristics (column 1) yields the highest accuracy for all outcomes predicted. As an example, for expenditure (Table 3.2B) the RMSE is 0.37 compared to 0.49 in the more flexible models. In the subsequent discussion section, I will address what we can conclude from the RMSEs obtained.

Table 3.2A: Predicting share working in agriculture (2008)

Dependent variable	**Proportion of cluster for whom agriculture is the main economic activity**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	2	6	9	16	15	16
In-sample RMSE	0.208	0.247	0.226	0.219	0.219	0.220
Out-of-sample RMSE	0.210	0.246	0.234	0.228	0.228	0.230
In-sample R-squared	0.472	0.252	0.377	0.415	0.411	0.407
Out-of-sample R-squared	0.480	0.290	0.355	0.387	0.389	0.378
Penalty parameter (λ)	0.010	0.002	0.001	0.002	0.005	0.007

Table 3.2B: Predicting expenditure (2008)

Dependent variable	**Per adult equivalent consumption expenditure (log)**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	5	3	11	10	10	11
In-sample RMSE	0.323	0.413	0.377	0.371	0.368	0.365
Out-of-sample RMSE	0.370	0.430	0.410	0.409	0.418	0.415
In-sample R-squared	0.605	0.353	0.463	0.477	0.488	0.497
Out-of-sample R-squared	0.422	0.219	0.291	0.292	0.261	0.273

Table 3.2C: Predicting proportion of households in extreme poverty (2008)

Dependent variable	**Proportion of households in extreme poverty**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	5	6	9	8	12	6
In-sample RMSE	0.170	0.180	0.175	0.174	0.170	0.176
Out-of-sample RMSE	0.198	0.208	0.202	0.209	0.206	0.205
In-sample R-squared	0.225	0.139	0.182	0.193	0.225	0.176
Out-of-sample R-squared	0.134	0.045	0.103	0.032	0.065	0.071
Penalty parameter (λ)	0.002	0.001	0.000	0.006	0.005	0.013

3.5.2 Over time

Tables 3.3A-3.3C show the result when data from 2008 is used for training and validation and the estimated parameters applied to the 2012 Landsat data to generate predictions which are evaluated against the actual 2012 survey values. First, a comparison of the in-sample performance (i.e., training on 2008 and evaluating on 2008) to the out-of-sample performance (i.e., training on 2008 and evaluating on 2012) shows that models trained on a specific year's data are generally less predictive of welfare and economic activity in another year. This is less of a surprise as out-of-sample predictions are typically less accurate than in-sample predictions. A striking result is that when comparing the RMSE across models of the same dependent variable the highest accuracy is, as in the cross-sectional case, from the model using built-up area as a predictive feature. As in the cross-sectional case, this suggests that most of the predictive information can be found in the built-up area feature.

Secondly, although, as noted above, predicting across time results in less accurate predictions than in the same year, when built-up area is used as a predictor (column 1), models are able to predict over time without much loss in accuracy. When using built-up area, models trained on a different time period are still able to explain high variation in the outcome variables in the test set. For example, a model trained on 2008 data, when applied to 2012, can explain 51% of the variation in log expenditure with similar in-sample and out-of-sample RMSE (see column 1 of Table 3.3B). The noteworthy exception is the share working in agriculture.

When doing the reverse i.e., predicting backwards, the pattern of results remains similar (see Tables 3.4A-3.4C). For example, a model trained on 2012 data when applied to 2008 can predict 53% of the variation in log expenditure with similar in-sample and out-of-sample RMSE (see column 1 of Table 3.4B). Notable here also is that the model trained in 2012 to predict the cluster level share working in agriculture in 2008, performs poorly.

Table 3.3A: (Predicting across time) Predicting the share working in agriculture in 2012 using model trained on 2008 data

Dependent variable	**Proportion of cluster for whom agriculture is the main economic activity**					
	(1)	(2)	(3)	(4)	(5)	(6)
Variables included	Built-up area, weather & geophysical characteristics	Indices, weather & geophysical characteristics	Band means, weather & geophysical characteristics	Band means, medians, stdevs weather & geophysical characteristics	Band interactions, weather & geophysical characteristics	Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	2	4	11	18	16	27
In-sample RMSE	0.207	0.246	0.228	0.215	0.221	0.213
Out-of-sample RMSE (2012)	0.376	0.413	0.401	0.402	0.402	0.400
In-sample R-squared	0.473	0.252	0.356	0.428	0.397	0.439
Out-of-sample R-squared (2012)	0.123	-0.056	0.0008	-0.0023	-0.0017	0.0082
Penalty parameter (λ)	0.009	0.014	0.0001	0.001	0.006	0.004

Table 3.3B: (Predicting across time) Predicting expenditure in 2012 using model trained on 2008 data

Dependent variable	**Per adult equivalent consumption expenditure (log)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Variables included	Built-up area, weather & geophysical characteristics	Indices, weather & geophysical characteristics	Band means, weather & geophysical characteristics	Band means, medians, stdevs weather & geophysical characteristics	Band interactions, weather & geophysical characteristics	Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	3	5	11	21	30	26
In-sample RMSE	0.335	0.428	0.394	0.374	0.357	0.356
Out-of-sample RMSE (2012)	0.355	0.414	0.384	0.388	0.388	0.391
In-sample R-squared	0.578	0.311	0.414	0.474	0.521	0.522
Out-of-sample R-squared (2012)	0.505	0.326	0.423	0.408	0.409	0.399
Penalty parameter (λ)	0.008	0.011	0.0004	0.001	0.003	0.004

Table 3.3C: (Predicting across time) Predicting proportion of households in extreme poverty using model trained on 2008 data

Dependent variable	**Proportion of households below the poverty line**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	2	7	8	7	7	9
In-sample RMSE	0.180	0.191	0.186	0.186	0.185	0.185
Out-of-sample RMSE (2012)	0.192	0.200	0.197	0.203	0.202	0.199
In-sample R-squared	0.198	0.0980	0.144	0.150	0.153	0.159
Out-of-sample R-squared (2012)	0.211	0.139	0.164	0.112	0.126	0.151
Penalty parameter (λ)	0.008	0.0007	0.001	0.008	0.009	0.010

Across time (Predicting backwards)

Table 3.4A: (Predicting across time) share working in agriculture in 2008 using model trained on 2012 data

Dependent variable	**Proportion of cluster for whom agriculture is the main economic activity**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2025
No. of variables selected by LASSO	4	4	9	17	3	8
In-sample RMSE	0.242	0.306	0.268	0.254	0.322	0.281
Out-of-sample RMSE (2008)	0.360	0.364	0.416	0.420	0.381	0.411
In-sample R-squared	0.645	0.433	0.563	0.608	0.372	0.522
Out-of-sample R-squared (2008)	-0.570	-0.606	-1.090	-1.138	-0.752	-1.044
Penalty parameter (λ)	0.006	0.013	0.001	0.001	0.092	0.044

Table 3.4B: (Predicting across time) Predicting expenditure in 2008 using model trained on 2012 data

Dependent variable	**Per adult equivalent consumption expenditure (log)**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2025
No. of variables selected by LASSO	5	7	10	14	30	11
In-sample RMSE	0.331	0.396	0.360	0.360	0.324	0.377
Out-of-sample RMSE (2008)	0.347	0.428	0.420	0.432	0.449	0.565
In-sample R-squared	0.581	0.401	0.506	0.505	0.600	0.458
Out-of-sample R-squared (2008)	0.532	0.286	0.313	0.274	0.217	-0.240
Penalty parameter (λ)	0.001	0.001	0.001	0.004	0.003	0.041

Table 3.4C: (Predicting across time) Predicting proportion of households in extreme poverty in 2008 using model trained on 2012 data

Dependent variable	**Proportion of households in extreme poverty**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2025
No. of variables selected by LASSO	3	7	10	15	19	27
In-sample RMSE	0.168	0.179	0.167	0.163	0.152	0.146
Out-of-sample RMSE (2008)	0.187	0.196	0.197	0.210	0.203	0.328
In-sample R-squared	0.320	0.228	0.328	0.358	0.438	0.485
Out-of-sample R-squared (2008)	0.124	0.0388	0.0276	-0.101	-0.0249	-1.677
Penalty parameter (λ)	0.006	0.0003	0.0008	0.002	0.002	0.002

3.5.3 Discussion

Some key insights emerge from the results. First, models using the built-up area feature perform better than directly using imagery characteristics. Second, models trained in a specific time period can predict expenditure as well as in the cross-sectional case, but the ability to predict agricultural activity across time is poor. It is worth considering potential explanations for these results and their implications.

First, why do models using the built-up feature predict better than those using raw imagery? Recall from the data section (3.3.2) that built-up feature was produced by the DLR from the satellite data by training models on ground truth data. Therefore, the prediction models that use built-up implicitly involve two steps – the first step (done by the DLR) is extracting the built-up feature and in the second step, I use the resulting feature in prediction models to predict survey data. In short, the analysis can be viewed as comparing two different approaches, namely one which uses a secondary product from the satellite data and a second approach of directly using imagery characteristics. The results suggest that the first approach performs best.

This finding can shed some light on how satellite data and possibly other similar types of high dimensional data are used. On one hand, as (Goldblatt et al., 2019) aptly note, producing features such as built-up from the satellite data is a computationally intensive process that also demands ground truth data – this is often a costly endeavour for individual researchers. While using imagery characteristics is less costly, these results suggest that approaches that directly use imagery characteristics are likely to be suboptimal. Indices offer an intermediate approach but are still outperformed by the approach in which the data is subject to a first step of extracting features such as a built-up. Therefore, one can conclude that there is value to be gained from efforts that focus on reducing high dimensional satellite imagery into features that are predictive of economic variables. Such features may be more useful for analysis than raw data. Determining which features are predictive of economic variables will be context-specific and can be guided by theory and existing empirical work. In this case, models using built-up area can predict the spatial variation in expenditure. This aligns with the fact that what constitutes built-up, such as infrastructure, are well-established correlates of economic activity and living standards. However, one could also think of contexts in which such a feature may not be best suited. In slums or densely built urban areas, for example, rather than built-up area per se, extracting features on the quality of the built-up itself such as type of roofing material may be more predictive of welfare.[11]

One could hypothesise that the poor predictive capacity across time may, in part, be attributed to the fact that cluster size changes between the waves. Specifically, as explained in the data section, the 2012 clusters have on average fewer households than in 2008 because of households moving to different locations. The results presented used all the clusters regardless of the number of households in the cluster in 2012. This means some clusters in 2012 were small – 89% of the clusters are composed of five or more households in 2012 whereas in 2008 all clusters had seven or eight households. I repeat the analysis excluding for 2012 the clusters that had five or fewer households. This results in a smaller sample in 2012 but I still obtain a pattern of results similar to the full sample (See Tables B2.1-B2.6 in the Appendix).

A more plausible explanation for the poor performance in predicting agricultural activity across time could be that agricultural activity is more variable over time than other indicators. Notice

[11] For example, see (Marx et al., 2019a) who find that in a Kenyan slum, imagery on roof quality from high resolution daytime satellite correlates with socioeconomic welfare of the slum households.

from the descriptive statistics (Tables B1.1-B1.3 in the Appendix) that although there is a statistically significant difference in the means of all variables between the two years, the difference in means between the two years for agricultural activity is substantial. The average difference in share of agricultural activity is about 78% of the 2008 mean, whereas the change in log expenditure is about 4.7% of the 2008 mean; in other words, agricultural activity appears to have changed more than expenditure between the survey rounds. As such, a model trained on data from a specific time period may not 'learn' well enough to predict in other time periods when agricultural activity is markedly different. Additionally, in this analysis, when predicting across time, data was trained only on one year's data. It is possible that the ability to predict across time may be improved by using multiple years' data for training. This could be a potential avenue for future research.

As discussed above, comparing RMSE of the models of the same dependent variable was able to show which model has the lowest error rate i.e., the most accurate model. In each case, this was the model that used the built-up feature. A pertinent question, however, is, how good is this best choice? Are these errors acceptable? It bears noting that there is no universally acceptable RMSE value since it is dependent on the scale of the outcome variable. To obtain a more intuitive understanding of the model performance, one can normalise the error by dividing it by the mean value of the dependent variable to obtain the Normalized Root Mean Square Error (NRMSE). The NRMSE is 0.3 for log expenditure and 0.4 for agricultural activity in the 2012 cross-section. This implies that the error for the best model is 30% of the mean value of log expenditure and 40% of the mean value of agricultural activity. For predictions across time, when predicting outcomes in 2012 from models trained on 2008 data, it is 30% of the mean of log expenditure and 60% of the mean of share in agricultural activity. When predicting backwards, for agriculture the error is equivalent to the mean. Seen in this way, these errors appear to be non-negligible. One can, therefore, conclude that, although the models can capture variation in the test data reasonably well, the errors in the predicted values are high. However, it is also important to bear in mind that due to the small cluster sizes, the survey variables are inherently noisy, making prediction a greater challenge. The results should, therefore, also be evaluated in light of the demanding nature of the prediction task.

Another way to assess the model's performance is to compare the results obtained to the result one would get from predicting past survey outcomes, simply using survey outcomes from another year. This means finding out how well using the Landsat compares to simply using the survey outcome of a location in a certain year to predict the outcome for the same location in another year. To examine this, I proceed as before and split the clusters into a training and test set. I train the model to predict the 2008 survey outcomes using the 2012 survey outcomes as well as weather and geophysical characteristics, as predictors. As before, I use cross-validation to minimise out-of-sample prediction error.

Tables 3.5A and 3.5B show the result. Results from the most accurate corresponding Landsat model are also provided in Panel B of the Tables. Using a survey from another year performs slightly better than the approach of using Landsat as indicated by the lower NRMSEs. Of course, this approach of predicting using survey outcomes from another year necessitates having that survey data available in the first place. This would not be feasible for locations without survey data. But this does suggest that the Landsat approach would be more informative for the more data deprived contexts and less so for those where other survey data is already available.

Overall, the findings point to an important trade-off. The costs associated with conventional face-to-face survey data collection imply that it is usually only feasible to obtain detailed data for a few locations and at infrequent intervals. Non-conventional secondary sources such as the Landsat satellite, collect granular data, at low marginal costs and at regular intervals. However, one may necessarily need to trade off some accuracy, especially to fill gaps in the most data deprived contexts.

Furthermore, the data has some limitations. While the built-up feature performed better than raw imagery, we cannot infer from the data aspects such as the condition or quality of the built-up. Yet, the quality rather than simply the area, is likely to be more correlated with welfare indicators used in this application, namely expenditure and assets, and therefore potentially offer better prediction. Additionally, since the use of refined features involves a first step of extracting these features, their predictive performance is ultimately tied to how these features are obtained. Specifically, a possible limitation of a globally trained dataset is that reference training data comes from different albeit representative locations. It is possible that prediction may be improved if the features are extracted using reference data obtained only from Tanzania. Finally, the predictive performance could be outperformed by higher resolution imagery, notwithstanding the costs of the of the latter.

With these caveats in mind, Table 3.6 shows predictions for other years for which LSMS data on the clusters is not available. Based on these predictions, average cluster level expenditure increased by 5% between 1990 and 2005. How may we use these predictions given the aforementioned limitation of poor accuracy? As suggested by Engstrom et al., (2017), where accuracy of predicted values is low, if predictions preserve rank accuracy reasonably well, predictions may still be informative for applications where it suffices to know rank ordering. There is no LSMS data for these periods to assess predictive accuracy or the ranking performance. However, the Spearman's rank correlation coefficient between the rankings of observed log expenditure in 2008 and predicted values from using a model trained in 2012 to predict in 2008 is 0.62, suggesting that there is likely to be movement in the rank order of clusters across time.

Table 3.5A: Predicting outcomes in 2012 survey

	Per adult equivalent consumption expenditure (log)	Proportion of cluster having agriculture as main economic activity	Proportion of households below the poverty line
Panel A: LSMS to LSMS			
In-sample NRMSE	0.193	0.539	0.899
Out-of-sample NRMSE	0.210	0.563	0.846
In-sample R-squared	0.758	0.617	0.353
Out-of-sample R-squared	0.715	0.586	0.300
Panel B: Using Landsat			
In-sample NRMSE	0.268	0.399	0.863
Out-of-sample NRMSE	0.304	0.404	0.944
In-sample R-squared	0.585	0.671	0.305
Out-of-sample R-squared	0.446	0.622	0.217

Table 3.5B: Predicting outcomes in 2008 survey

	Per adult equivalent consumption expenditure (log)	Proportion of cluster having agriculture as main economic activity	Proportion of households below the poverty line
Panel A: LSMS to LSMS			
In-sample NRMSE	0.204	0.392	0.794
Out-of-sample NRMSE	0.211	0.413	0.843
In-sample R-squared	0.759	0.681	0.411
Out-of-sample R-squared	0.734	0.605	0.375
Panel B: Using Landsat			
In-sample NRMSE	0.247	0.632	0.984
Out-of-sample NRMSE	0.300	0.631	0.942
In-sample R-squared	0.605	0.472	0.225
Out-of-sample R-squared	0.422	0.480	0.134

Table 3.6: Predicting for other years

	1990	1995	2000	2005
Log per adult equivalent consumption expenditure (2011 USD)	1.23	1.25	1.27	1.28
Observations	409	409	409	409

3.6 Conclusion

Sub-national data on welfare and economic activity is often lacking or collected at infrequent intervals. While findings from an emerging literature suggest that non-conventional sources have potential to fill such data gaps, their external validity is not yet understood. In this paper, I have examined the ability of remote sensing data from the Landsat to predict local level outcomes. I examined the case where we may have sub-national survey data in one time period and are interested in obtaining estimates of the share working in agriculture, wealth and consumption expenditure for other locations in the same time period. I have also examined the case where we may have survey data in a certain year and are interested in obtaining predictions for another year for which survey data is not available. I find that that Landsat data, combined with weather and geophysical characteristics, can explain some variation in sub-national agricultural activity, wealth and consumption for new or unseen locations. The finding that using a built-up area variable that was extracted from the Landsat has better predictive ability than using spectral indices or all Landsat bands lends some support to the possible merits of first using high dimensional satellite imagery to produce secondary datasets of measures that are predictive of economic outcomes but also intuitively easier to interpret than the raw data.

However, the predictive accuracy found in this context is low; errors are about a third of the mean of expenditure and close to half of the mean of agricultural activity. Additionally, the ability to generalise over time is poor. Indeed, the results suggest that for locations where we have survey data from another year, the approach tested here is unlikely to do any better than predictions using that survey. The poor predictive performance suggests that such an approach may complement but not supplant ground-level data collection for the purposes of policy decisions which require more precise estimates.

It is possible that the use of more flexible machine learning methods would have yielded better predictive performance. However, these methods require copious amounts of data for training; the limited survey sample size constrained their application in this paper. Additionally, the ability to predict across time is likely to be improved by using multiple years' data for training –a key area for ongoing research efforts. Overall, as non-conventional data sources and machine learning techniques become increasingly available, it remains as equally important to assess the viability of approaches in different contexts before they are widely adopted.

3.7 Appendix

B1. Descriptive statistics

Table B1.1: Summary Statistics (2008)

	mean	sd	min	max
Survey variables				
Log per adult equivalent consumption expenditure (2011 USD)	1.29	0.51	0.30	3.03
Share of households in extreme poverty	0.18	0.20	0.00	0.88
Share of households with agriculture as main economic activity	0.33	0.29	0.00	1.00
Non-survey data				
Elevation(m)	690.20	613.12	6.81	2369.70
Slope	4.75	3.30	1.32	20.25
NDBI	-0.05	0.10	-0.30	0.20
NDVI	0.32	0.14	-0.04	0.66
NDWI	-0.37	0.13	-0.65	-0.02
Annual precipitation(mm)	1044.72	270.43	387.75	1854.21
Built-up area (km sq.)	13.78	21.51	0.00	72.84
Observations	409			

Table B1.2: Summary Statistics (2012)

	mean	sd	min	max
Survey variables				
Log per adult equivalent consumption expenditure (2011 USD	1.22	0.51	0.21	2.69
Share of households in extreme poverty	0.21	0.22	0.00	1.00
Share of households with agriculture as main economic activity	0.59	0.40	0.00	1.00
Asset score	3.23	1.03	1.00	5.00
Non-survey data				
Elevation(m)	690.20	613.12	6.81	2369.70
Slope	4.75	3.30	1.32	20.25
NDBI	-0.05	0.10	-0.26	0.20
NDVI	0.32	0.13	-0.04	0.64
NDWI	-0.37	0.13	-0.61	-0.02
Annual precipitation(mm)	966.02	238.35	389.30	1999.44
Built-up area (km sq.)	14.32	21.74	0.00	73.01
Observations	409			

Table B1.3: Differences in survey outcomes between 2008 and 2012

	Mean difference	t	p > |t|
Log per adult equivalent consumption expenditure (2011 USD	-0.06	-4.81	0.00
Share of households in extreme poverty	0.02	2.33	0.02
Share of households with agriculture as main economic activity	0.26	20.57	0.00
Observations	409		

B2. Excluding small clusters

Table B2.1: Predicting share working in agriculture in 2012 using model trained on 2008 data

Dependent variable	**Proportion of cluster for whom agriculture is the main economic activity**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2026
No. of variables selected by LASSO	4	7	11	19	11	14
In-sample RMSE	0.208	0.243	0.226	0.213	0.219	0.216
Out-of-sample RMSE (2012)	0.384	0.422	0.413	0.410	0.415	0.413
In-sample R-squared	0.483	0.297	0.392	0.457	0.429	0.443
Out-of-sample R-squared (2012)	-0.016	-0.222	-0.171	-0.156	-0.181	-0.174
Penalty parameter (λ)	0.0031	0.0005	0.0001	0.007	0.0102	0.009

Table B2.2: Predicting expenditure in 2012 using model trained on 2008 data

Dependent variable	**Per adult equivalent consumption expenditure (log)**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2026
No. of variables selected by LASSO	3	7	11	8	15	8
In-sample RMSE	0.329	0.411	0.379	0.374	0.360	0.366
Out-of-sample RMSE (2012)	0.338	0.402	0.368	0.396	0.374	0.391
In-sample R-squared	0.569	0.329	0.400	0.437	0.440	0.440
Out-of-sample R-squared	0.605	0.383	0.475	0.488	0.524	0.508

(2012)						
Penalty parameter (λ)	0.467	0.247	0.367	0.270	0.348	0.286

Table B2.3: Predicting proportion of households in extreme poverty using model trained on 2008 data

Dependent variable	**Proportion of households in extreme poverty**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2028
No. of variables selected by LASSO	5	5	9	6	11	16
In-sample RMSE	0.177	0.185	0.180	0.180	0.177	0.178
Out-of-sample RMSE (2012)	0.198	0.208	0.205	0.209	0.207	0.207
In-sample R-squared	0.225	0.157	0.200	0.194	0.223	0.217
Out-of-sample R-squared (2012)	0.170	0.0903	0.111	0.0801	0.0955	0.0941
Penalty parameter (λ)	0.0002	0.001	0.001	0.007	0.005	0.01

Table B2.4: (Predicting across time) share in agriculture in 2008 using model trained on 2012 data

Dependent variable	**Proportion of cluster for whom agriculture is the main economic activity**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2024
No. of variables selected by LASSO	5	6	9	18	23	20
In-sample RMSE	0.228	0.291	0.259	0.246	0.233	0.236
Out-of-sample RMSE (2008)	0.360	0.368	0.421	0.448	0.423	0.546
In-sample R-squared	0.643	0.420	0.540	0.585	0.627	0.617

Out-of-sample R-squared (2008)	-0.563	-0.635	-1.144	-1.433	-1.168	-2.610
Penalty parameter (λ)	0.001	0.002	0.001	0.001	0.003	0.005

Table B2.5: (Predicting across time) Predicting expenditure in 2008 using model trained on 2012 data

Dependent variable	**Per adult equivalent consumption expenditure (log)**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2024
No. of variables selected by LASSO	5	4	11	15	17	21
In-sample RMSE	0.327	0.367	0.339	0.332	0.324	0.312
Out-of-sample RMSE (2008)	0.347	0.440	0.431	0.453	0.427	0.480
In-sample R-squared	0.498	0.366	0.459	0.483	0.507	0.544
Out-of-sample R-squared (2008)	0.531	0.248	0.277	0.200	0.289	0.104
Penalty parameter (λ)	0.0007	0.010	0.001	0.003	0.006	0.006

Table B2.6: (Predicting across time) Predicting proportion of households in extreme poverty in 2008 using model trained on 2012 data

Dependent variable	**Proportion of households in extreme poverty**					
Variables included	(1) Built-up area, weather & geophysical characteristics	(2) Indices, weather & geophysical characteristics	(3) Band means, weather & geophysical characteristics	(4) Band means, medians, stdevs weather & geophysical characteristics	(5) Band interactions, weather & geophysical characteristics	(6) Polynomials of band characteristics, weather & geophysical characteristics
No. of variables in the model	6	8	12	26	257	2024
No. of variables selected by LASSO	5	6	10	11	18	24
In-sample RMSE	0.192	0.199	0.192	0.189	0.180	0.171
Out-of-sample RMSE (2008)	0.187	0.198	0.200	0.222	0.215	0.255
In-sample R-squared	0.242	0.186	0.242	0.268	0.337	0.399

Out-of-sample R-squared (2008)	0.132	0.0181	0.00459	-0.229	-0.153	-0.624
Penalty parameter (λ)	0.0003	0.003	0.001	0.003	0.003	0.003

Chapter 4

Do refugee camps have an urbanizing effect? –A bird's eye view from Tanzania

4.1 Introduction

Urban growth and forced displacement are among the defining demographic trends of the 21st century (United Nations Department of Economic and Social Affairs [UNDESA], 2019; United Nations Economic Commission for Africa [UNECA], 2017; United Nations High Commissioner for Refugees [UNHCR], 2019). Africa is experiencing the fastest rate of urban growth in the world. It is projected that by 2050 urban land cover will be four times higher than in 2000 (OECD 2020). While some of that urban growth is driven by expansion of existing cities and urban areas, a distinctive feature of the urbanisation process in African countries is the transformation of previously rural areas. The emergence of small towns and urban clusters in previously rural areas has challenged the singular focus on rural-urban migration and has led to a growing research interest in understanding other drivers of urbanisation in Africa. There is a growing recognition that transformation of rural settlements may play a far more important role in explaining urbanisation than hitherto acknowledged (Berdegué et al., 2014; Fox, 2017; OECD, 2020; Tacoli & Agergaard, 2017).

At the same time, although forced displacement is a global phenomenon, Africa bears a disproportionately large share of the displaced – over one third of the global forcibly displaced population. A significant number of them live in camps which have been the conventional policy response for hosting displaced populations (UNHCR, 2019a). These camps are usually envisioned as an emergency temporary response to cater for the needs of the displaced while they await eventual return and are typically located in areas previously of relatively low population density. However, the camps often last for several years and become nuclei of humanitarian aid, infrastructure investments and local trade in previously remote areas. Countries such as Tanzania, Kenya, Uganda and the Democratic Republic of Congo have been marked by regional forced displacement.

These influxes were large and sudden enough to alter local socieconomic outcomes. Previous micro-level analyses has found that proximity to camps has an effect on the local labour market, household consumption and wealth, human capital, and the environment in hosting regions (Alix-Garcia et al., 2013; Alix-Garcia & Saah, 2010; Baez, 2011; Maystadt et al., 2020; Maystadt & Duranton, 2019; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015, 2016). However, the extent to which the camps, as centres of resource inflows, such as humanitarian aid, infrastructural investments and local trade, have an urbanising effect on surrounding rural areas, has received little attention.

Studies of urbanisation in Africa have typically focused on emergence of urban settlements around a nexus such as a port, road or railway with consequent rural-urban migration as the primary driver of urban growth (Jedwab & Moradi, 2016; Lall, 2017). Natural resource endowments such as mines and the locations of government and administrative services are also other important catalysts for the emergence of urban settlements in sub-Saharan Africa (Agergaard et al., 2019; Lazaro et al., 2019; Stren & Halfani, 2001; Tacoli & Agergaard, 2017; Wolff et al., 2020). Research has pointed to a distinction between production-led urbanization which-occurs alongside industrialization and consumption-led urbanization which is characterized primarily by nontradables. In the latter process, secondary towns can form as "consumption cities" where farmers sell their products and buy services, rather than forming as manufacturing centres as is typically assumed in models of urban growth in the industrialized world (Gollin et al., 2016; Henderson & Turner, 2020).

The interactions that develop around refugee camps could contribute to such consumer-led urbanization processes. The increased population provides a ready local market for agricultural products. Additional trading opportunities develop as refugees exchange or sell relief items such as clothing, blankets, construction material and cooking implements for items that are not included in their aid baskets. Foodstuffs that are not matched to refugees' usual diets in the home countries are typically sold for more preferable or familiar foodstuffs. These exchanges are conducted in the camp market and markets in surrounding villages which are bustling with restaurants, coffee shops and services such as tailoring, radio repair and hair dressing. The influx of relief workers also creates demand for a variety of products and services to cater to their tastes leading to the mushrooming of small businesses. The vibrant market opportunities attract outsiders who move in to take advantage of these new business opportunities as they are more able to source merchandize from outside the region. Locals also hire refugees as a source of cheap labour for neighbouring villages. In addition to business and trade, the relief operations generate employment opportunities for hosts as some members of the local community are employed by NGOs and in the homes of aid workers as drivers or domestic helps. Locals also benefit from services such as water pumps and health services provided in the camps (Jacobsen, 2005; Veney, 2007). Thus, refugee camps, by providing an enhanced demand for agricultural producers, a source of cheap labour for neighbouring villages, and multiple sources of enhanced demand arising from the influx of expatriate workers and investment in infrastructure, can in principle transform the environment in hosting regions.

While some research documents spatial expansion of rural villages and “mushrooming of new boomtowns” and highlights how conflict dynamics, such as forced displacement reshape the local landscape, limited data has proven to be a considerable challenge for researchers (Büscher & Mathys, 2019: p.55). This paper aims to fill this gap in the literature, using high resolution satellite data to examine the urbanising effect of refugee camps in Tanzania in the period 1993-2015. In so doing, it contributes to a deeper understanding of the effects of forced displacement beyond individual and household level outcomes to wider processes such as urban growth. Understanding and quantifying the effects of forced displacement on urban growth is crucial to effectively managing the rapid pace of urban growth.

To overcome prevailing data challenges, the paper uses an indicator of non-agricultural economic activity and settlement (built-up area), to investigate whether proximity to camps has had an urbanising effect on surrounding areas. By urbanising effect, I mean an increase in settlement and non-agricultural economic activity as indicated by built-up land area. I will examine the ‘spillover’ effect on areas beyond the camps themselves. The empirical approach follows a difference-in-difference strategy using locality level panel data on built-up area. The paper finds a small positive effect of camps on built-up area. The built-up area within localities increased on average by 0.006 sq. km., about 6% from the average pre-camp level of built-up. The effect is found only within rural localities. I find no effect on localities that were already more urbanised.

The remainder of the paper is structured as follows: section 4.2 provides an overview of the pertinent literature, section 4.3 describes the data used, and section 4.4 outlines the empirical methodology. The results are discussed in section 4.5. Section 4.6 provides some robustness checks, while section 4.7 concludes.

4.2 Literature review

Urban growth, the increase in the number of people living in urban areas, is acknowledged to be explained by natural urban population growth, migration from rural areas to urban areas, and the transformation of rural settlements to urban (Agergaard et al., 2019; Dyson, 2011; Fox, 2017; Tacoli & Agergaard, 2017; UNECA, 2017).[12] Among these factors, early attention focused on rural-urban migration. The wage gap between urban and rural areas attracts migrants from rural areas (Harris & Todaro, 1970). Attention has increasingly shifted, however, away from a singular focus on rural-urban migration in explaining Africa's urbanisation process, as the experience of African countries demonstrates that other factors may play a far more important role than previously acknowledged (Berdegué et al., 2014; Dyson, 2011; Fay & Opal, 2000; Fox, 2014; Jedwab et al., 2014; Tacoli & Agergaard, 2017; UNECA, 2017). A unique feature of the urbanisation process in African countries is the emergence of urban centres in areas previously considered to be rural (Berdegué et al., 2014; Fox, 2017; Güneralp et al., 2017; Lazaro et al., 2019; OECD, 2020; Tacoli & Agergaard, 2017). Some villages have even metamorphosed into settlements of urban scale (OECD, 2020).

Büscher and Mathys (2019) and Büscher (2018) provide compelling qualitative evidence that conflict dynamics – including the settlement of forcibly displaced populations in an area – can be

[12] While urbanisation is commonly loosely used to refer to the increase in the number of people living in urban areas, in its strictest definition, it is the increase in the proportion of a population living in urban areas in comparison to rural areas. In other words, urbanisation refers to changes in population ratios while urban growth refers to changes in population sizes (Fox, 2017).

powerful drivers of urban growth. In Büscher and Mathys (2019), the authors provide qualitative evidence that some rural villages in the Democratic Republic of Congo (D.R.C.) mushroomed into boom towns that did not exist prior to the war, a phenomenon they refer to as "boomtown urbanisation" (Büscher and Mathys, 2019: p.55). They discuss how the population concentration created demand and opportunities for the emergence of commercial transactions. However, the context that these authors examine predominantly involves internally displaced individuals. For example, they study the case of a town that emerged as a result of being an epicentre of internally displaced people (IDPs) who could self-settle. Refugee camps exist in a different context since refugees face greater restrictions on their activities, including the right to work, own property or self-settle. Nevertheless, the authors aptly observe that such 'hidden' processes of urban growth have largely escaped academic and policy inquiry and that little is known about them, including the scope of the geographic expansion, demographic evolution or of the social and economic networks and political characteristics. Alix-Garcia et al. (2013) show evidence of dramatic changes in land-use in Darfur, Sudan due to civil war, but these changes primarily concern the abandonment of land or more intensive use of agriculture, rather than a shift between rural and urban environments. Similarly, Maystadt et al. (2020) show how land use is affected by the presence of refugee camps in Africa, but their focus is on the effect of camps on deforestation and its replacement by cropland. Hence the literature on the longer run effect of refugee camps on urbanization is limited.

These observations invite us to consider the channels through which the presence of refugee camps could affect urban growth. Some qualitative works examine whether refugee camps, as initially empty spaces, transform into settlements that could be classed as urban (Agier, 2002; Jansen, 2016; Montclos & Kagwanja, 2000; Oka, 2011). This literature finds that camps have the potential to be urban centres because they have the typical requisite characteristics – high population density, cosmopolitan population, concentration of infrastructure and facilities (roads, schools, health centres and sanitation infrastructure), and are loci of trading activities and labour markets. Additionally, of the camp inhabitants that have some form of employment, most are engaged in non-agricultural activities, either employed by the humanitarian organisations providing some of the camp services, or in construction work or local trade (Montclos & Kagwanja, 2000; Oka, 2011). It is important to bear in mind here that the definition of "urban" varies considerably. (This point is discussed fully in the data section). In any case, this literature focusses on examining the camp itself and does not explore how these urban-like features of the camp could affect the surrounding areas. The question arising is, whether the camp, having the aforementioned characteristics, precipitates transformation of surrounding areas.

If camps can be categorised as urban settlements (or at least have characteristics of urban settlements), then the potential for camps to have an urbanising effect can be approached through the lens of the literature on proximity to urban centres and its impact on rural areas. It has been shown that agriculture-consumption linkages lead to the development of the rural non-farm economy (Diao et al., 2019; Haggblade et al., 1989, 2010; Otsuka, 2007). Increased demand for agricultural produce can lead to adoption of better inputs such as fertilisers or seeds of better quality, thus improving land and labour productivity. Increased labour productivity in the agricultural sector releases workers who can then undertake non-farm activities. At the same time, part of the increased farm income from productivity increases can be invested outside agriculture. The higher income also enables households to afford more non-food items (Berdegué et al., 2014; Haggblade et al., 2007). Besides increased demand for agricultural produce, other channels through which proximity to urban centres affects rural households and businesses include greater access to services, increased investments in infrastructure and connectivity, and human capital spill overs (Berdegué et al., 2015).

From the aforementioned literature, it is plausible that proximity to camps would affect the transformation of hosting regions. First, the increased population provides demand for agricultural produce. Second, although refugee movement was restricted, the refugee population did still provide an inexpensive source of cheap agricultural labour for neighbouring villages (Whitaker, 1999). Lower wages due to the increased labour supply may drive locals from employment activities outside the home to working in household farms as was the case in Tanzania (See Ruiz & Vargas-Silva, 2015, 2016). Lower labour costs, however, may also provide small-scale farmers with opportunities to expand agricultural production and to subsequently shift to other forms of income-generating activities. In addition, the influx of international organisations may provide employment opportunities with the organisations directly, as domestic staff for expatriate workers or opportunities for entrepreneurship to meet demand for new goods to cater to the taste of the new populations (Landau, 2004; Whitaker, 1999). In Tanzania, the forced migration shock impacted the host regions' labour market and goods markets (Ruiz & Vargas-Silva, 2015, 2016). Roads in the hosting regions were upgraded in order to service the camps and locals benefitted from access to facilities such as health and water facilities provided to camp populations (Whitaker, 1999). Other studies also found that locals around refugee camps experienced an increase in consumption spending and assets (Alix-Garcia & Saah, 2010; Maystadt & Verwimp, 2014; Ruiz & Vargas-Silva, 2015). In summary, camps are associated with new employment opportunities, increased trade and cheap labour supply – factors which may transform surrounding regions.

Urban growth is taken to be a process of both a concentration of people and non-agricultural economic activity. It is, however, important to bear in mind that, while urban growth is understood to involve such a structural transformation or reorganization of activities away from agriculture, areas may be classified as urban demographically (based on high population density) without such a transformation. At a national level, the phenomenon of urbanisation without structural transformation has been documented in Africa (Fay & Opal, 2000; Fox, 2017; Gollin et al., 2016; Potts, 2013).

It is worth pointing out that some factors associated with camps may lead one to hypothesise that the camps may induce outward migration by locals or render these areas less attractive. In Tanzania, there was pressure on natural resources (increased demand for firewood and water) and reports of outbreaks of diseases in local areas at the outset (Whitaker, 1999; Green 1995; Eriksson et al., 1996; Berry 2008). In addition, competition for jobs may lead to outward migration of those displaced from the labour force by refugee labour. There is not, however, much evidence to support this hypothesis. On the contrary, there is evidence, at least for the Kagera region, that refugee presence was associated with lower likelihood of outward migration (Maystadt & Verwimp, 2014).

It is also worth taking into consideration that rural transformation is mediated by institutional frameworks (Berdegué et al., 2014). Tacoli and Agergaard (2017) specifically examine the emergence of small urban centres in Tanzania and find that their potential for emerging within rural regions depends on the quality of transport and communication links to other regions, and the structural conditions prevailing at the national and local level. Additionally, where rural dwellers move from subsistence agriculture but into low-paying low-productivity non-farm work, then the process of rural transformation may be constrained (Berdegué et al., 2014; Haggblade et al., 2007). In such cases, small urban settlements may emerge but lack sufficient stimulus to grow significantly. (Berdegué et al., 2014; Haggblade et al., 2007).

The existence of institutional constraints is worth bearing in mind because, while the qualitative literature finds that refugee camps have the typical characteristics of urban settlements, at the same time, it emphasises that camps are superficial urban centres. The camp is viewed to exist in a state of “temporary permanence” (Jansen, 2016: p.2 citing Picker & Pasquetti, 2015). Most of the economic activity remains informal and unsanctioned, even if it is regular, and there is reluctance by government authorities to integrate it into the wider economy because that engenders perceptions of refugee permanence (Agier, 2002; Montclos & Kagwanja, 2000; Oka, 2011).

This institutional context in which refugee camps exist suggests that they may differ in their urbanising influence. The urbanising effect of camps is, therefore, not straightforward. While it has been found that camps in Tanzania stimulated local trade, attracted infrastructural investments and created increased demand for agriculture and for other products, in the presence of the aforementioned institutional constraints, these effects may not necessarily lead to the rural transformation process discussed above. Landau (2004) argues that despite the influx giving rise to conditions (higher population density, improved infrastructure, abundant labour supply) that have elsewhere led to rural transformation, the influx did not lead to a fundamental transformation of economic practices, but simply increased the likelihood of reliance on subsistence agriculture and localised trade. The authors’ conclusions are, however, drawn from a qualitative comparison of two Tanzanian districts and should, therefore, be treated with caution. Thus, a key motivation for this paper is to fill a gap in the literature by examining whether the camps had an urbanising effect using a measurable indicator of non-agricultural economic activity and settlement (built-up area) for which I obtain longitudinal data.

4.3 Data

Lack of longitudinal data remains a major constraint to understanding urban growth processes, especially for developing countries where data limitations are most acute. Studying whether rural areas transition to urban and/or whether existing urban areas expand, for example, requires a measure of what is urban, but obtaining such a measurable attribute over time is challenging. Not only does the definition of what is 'urban' vary from country to country, but to complicate matters further, even within the same country, the definition often varies in surveys from year to year, hindering the ability to study changes over time (Angel et al., 2005, 2011; Cohen, 2006; European Commission et al., 2020; OECD, 2020; Tacoli & Agergaard, 2017; UNFPA, 2011). Tanzania is a case in point – the definition of an urban area has been inconsistent in its censuses (1967, 1978, 1988, 2002, 2012). Wenban-Smith (2014) traces the evolution of urban designation in Tanzania. In the early censuses (1967 and 1978), the definition was based on administrative boundaries. By 1988, the designation of localities as urban was left to the discretion of the district administration. There were no central criteria and district authorities differed in their practices. The subsequent censuses of 2002 and 2012 do not offer any clarity, but rather confirm that urban designations are assigned non-uniformly. The 2002 census, for example, states that "The urban areas are defined as the localities that are identified as urban areas by the district authority. There is no clear and uniform definition applied by the various districts in the country." (Wenban-Smith, 2014: p.4-5; National Bureau of Statistics, 2014).

Forced displacement amplifies these challenges in already data-scarce environments. Büscher and Mathys (2019), highlight some of the enormous data challenges in these contexts – lack of census data, maps and cadastral data, or where they do exist, finding different and contradictory versions. Researchers faced with these circumstances are often forced to abandon the enterprise altogether, or to resort to observations and narrative accounts of histories and changes, to piece together patterns and trends.

To overcome these data challenges, I use built-up area – a measurable indicator of economic activity and settlement, as a definition of 'urban growth'. The expansion of the built-up environment to house populations and non-agricultural activities is a fundamental feature of urban growth, since rural-urban transformation is typically accompanied by conversion of land from agriculture to other uses (Bren d'Amour et al., 2017). Built-up area is highly positively correlated with the distribution of population and non-agricultural economic activity (Bagan & Yamagata, 2015; Goldblatt, Deininger, et al., 2018; Sudhira et al., 2004; Yin et al., 2005).

To obtain built-up, I use the World Settlement Footprint Evolution (WSF), developed by the German Aerospace Center (DLR). The WSF is the first dataset showing the evolution of built-up globally at both a high temporal (annual) and spatial resolution (30m). I exploit these two unique characteristics of the dataset – high temporal frequency and spatial resolution. To date, only a few datasets outline settlement or built-up area. However, these datasets are limited in the time span they cover or in their temporal frequency. Most of the data is available only for one year. The most frequent dataset is available only every 15 years and at a low resolution. Hence, they do not facilitate continuous monitoring or study of urbanisation patterns (Goldblatt, Deininger, et al., 2018; Marconcini et al., 2019; Palacios-Lopez et al., 2019). Additionally, the ability of the datasets to detect small settlements in rural areas, such as much of Tanzania, and scattered suburban areas is limited due to the fact that images are acquired from a single date which hampers classification accuracy (Marconcini et al., 2019). To overcome these limitations, the DLR team combined an extensive archive of Landsat images from 1985-2015 and iterative machine learning techniques, to map built-up on a yearly basis. Because the Landsat satellite revisits the same location multiple times within a year, multiple images of the same location are available which improves upon the accuracy of the aforementioned datasets. Samples from a variety of locations globally, including locations characterised by small settlements (akin to those found in Tanzania), were used in training classifiers, to ensure data is representative of diverse settlement patterns. Built-up includes buildings, building lots, roads or paved surfaces (Marconcini et al., 2019).

My units of analysis are grid-cells (localities) of 0.05 by 0.05 degrees (approximately 5.5km by 5.5km or 30.25 sq. km.). I based the choice of the size of the grid cell on information about the population density and average population in Western Tanzania. As an example, the population density of Kagera region is about 97/km^2 and the average village has a population of about 3,500 people: therefore, an average village in Kagera would be about 36/km^2; roughly the size of the grid cell I use (Tanzania National Bureau of Statistics, 2012 population and housing census).[13] I use these grids (rather than administrative units) as they are exogenous by construction whereas administrative boundaries tend to be defined by geographic, demographic and political characteristics (Angel et al., 2011; UNFPA, 2011). The grid-level approach has been used in other research contexts to investigate subnational outcomes.[14] I merge the WSF data with the localities. I then calculate the total number of pixels that were classified as built-up in each year and from that calculate the size of the built-up area (in sq. km.) in each locality for each year over the period 1985-2015. Figure C1 in the Appendix shows that the built-up area is positively correlated with population density in Tanzania.

[13] In section 4.5.2, I will also provide analysis separately for localities that at the time were already more densely populated and those that were more rural as the effect may vary by pre-camp density.
[14] See for example (Mamo et al., 2019; Michalopoulos & Papaioannou, 2013, 2014)

Nevertheless, the built-up variable has its limitations. While it is used as proxy for settlement and non-agricultural activity, it is not possible to precisely pin down what the built-up is used for. One concern is that it could also be picking up built structures that are used for agricultural processes and hence may overstate the transition from agricultural activity. However, this is likely to be a concern in more developed countries and less so for the context at hand given the relatively rural landscape and less sophisticated agricultural sector. Additionally, it is not possible to infer the quality of the structures.[15] These limitations constrain the ability to tease out the nature of the economic activity induced by camps.

Information on refugee camps is compiled from various sources. The geographic coordinates are obtained from Maystadt and Verwimp (2014), Zhou (2014) and the UNHCR field office in Tanzania. The opening and closing dates of camps are obtained from Zhou (2014) and the UNHCR field office in Tanzania. The total camp sample comprises 22 camps. As the interest is in the built-up of localities surrounding camps rather than within the camps themselves, I exclude the area of the camps. I further limit the sample to localities within 500km of a camp. Table 4.1 provides some summary statistics on the outcome variable (built-up area) for all localities for all years in the sample. The level of built-up is low (0.1km^2) which is not surprising as most of the locations are rural.

Table 4.1: Summary statistics of Built-up area (all localities, 1985-2015)

	mean	s.d.	min	Max
Built-up area (sq.km.)	0.10	0.60	0.00	24.40
Observations	415789			

[15] For an example of an application where high resolution imagery can distinguish the quality of housing stock see (Marx et al., 2019b). However, such high-resolution imagery is usually costly or limited in time span and/or temporal frequency.

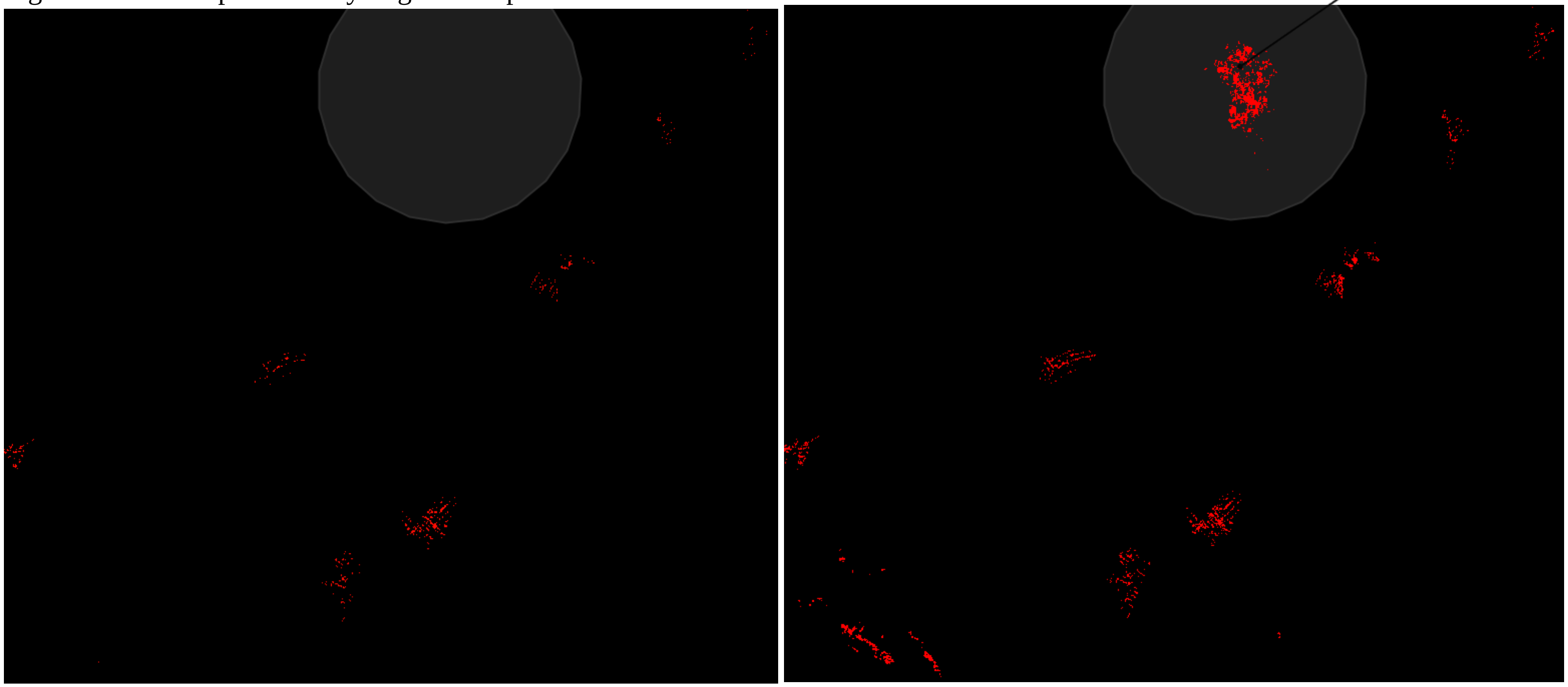
To estimate the effect of proximity to camps on settlement and non-agricultural economic activity in surrounding areas as proxied by built-up area, I will compare localities within 100km of a camp to those farther than 100km but within 500km. In the subsequent section, I provide a discussion on the validity of this comparison. Camps opened between 1993 and 1999. 19% of all localities had a camp within 100km by 1999 (Table 4.2).

Table 4.2: Number of localities with a camp (within 100km)

	Year	No. of localities		
	1993	155		
	1994	1296		
	1995	1425		
	1996	1759		
	1997	2515		
	1999	2758		
No. of localities that ever had a camp		2758	19%	of the sample
No. of localities that never had a camp		11654	81%	of the sample
Total sample		14412		

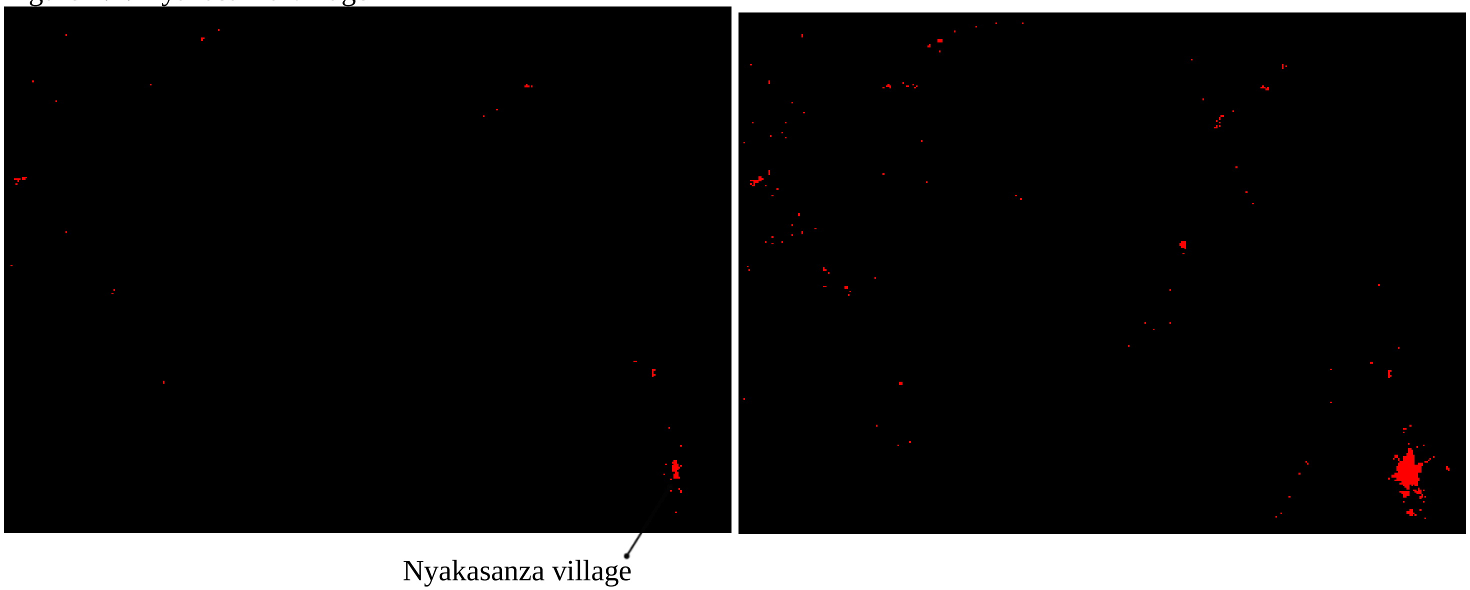
For illustrative purposes, Figure 4.1 shows the areas around Nyarugusu camp which opened in 1996. Panel A shows the built-up in 1990 before the camp opened and Panel B in 2000. Changes in built-up in locations beyond the camp are visually apparent. Figure 4.2 zooms into Nyakasanza village before and after the nearest camp (Lukole) opened. Some expansion of the village is noticeable. While these illustrations provide some evidence of built-up changes, the magnitude of the change and whether it can be attributed to camps, that is, whether camps are associated with an effect above-and-beyond the ‘natural’ increase in built-up areas associated with population growth which would have happened even in the absence of camps, is a matter of empirical inquiry which this paper seeks to examine.

Figure 4.1: Built-up around Nyarugusu camp



Notes: The left panel is from 1990 before the camp. The right panel is in 2000, four years after Nyarugusu camp opened. The circular ring denotes the camp location. Built-up changes are visually apparent beyond the camp.

Figure 4.2: Nyakasanza village



Notes: The left panel shows built-up around Nyakasanza village in 1990 (bottom right-hand corner). The nearest camp (Lukole) opened in 1993. The right panel shows the same village in 2000.

4.4 Empirical methodology

To estimate the effect of proximity to camps on built-up area, I use the following specification:

$$y_{ijt} = \beta(Camp100_{it}) + \theta_i + \delta_t + + \gamma_{jt} + e_{it} \qquad (4.1)$$

y_{ijt}, is the built-up area in sq.km. for grid i in region j in year t. $Camp100_{it}$ is a dummy indicating whether grid i is within 100 km of a camp in year t. For a grid that never has a camp within 100 km, the dummy always takes the value 0. For a grid that has a camp within 100 km, it takes the value 0 for all years before the camp opens, 1 for all the years the camp is open and 0 again for all years after it closes. I exclude all post-closure years in order to examine the effect in the period during which the camps were operational. In section 4.5, I also examine the effect in the post-closure period. I also exclude all grids with less than two observations i.e., grids that have only one grid-year. δ_t are year fixed effects capturing common shocks, θ_i are grid fixed effects representing time-invariant grid characteristics while γ_{jt} are linear region-specific time trends.[16] e_{it} is the idiosyncratic error term. Since grid and time fixed effects control for time-invariant spatial heterogeneity among grids and time-variant shocks that simultaneously affect all the grids respectively, this approach reduces endogeneity concerns.

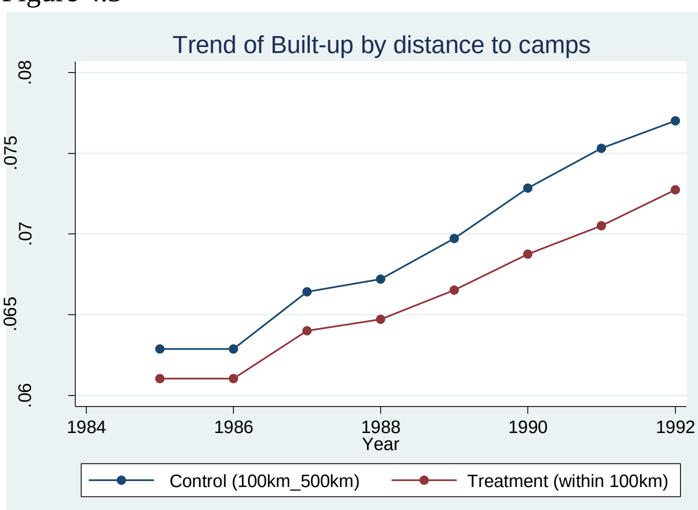
The coefficient of interest, β, identifies the change in built-up area associated with a change in a grid's status from having no camp within 100 km to having a camp within 100 km. There is enough within variation on the treatment variable – 2758 grids (or 19% of the sample) switch from having no camp to being within 100km of a camp.

[16] There are on average 1200 grids in a region.

Identification comes from both temporal variation (within the localities that switch from having no camp to having a camp) and spatial variation (in proximity to camps). By using localities that are within 100km-500km as the control group, the validity of this strategy rests on the assumption of parallel trends prior to the treatment This relies on the two groups of grid-years having the same pre-camp trends and I formally test this assumption in the sensitivity analysis.

Figure 4.3 shows the evolution of built-up pre-1993, before any camps opened and, therefore, the cleanest universal pre-camp period for all localities. Though localities farther from the eventual campsites are more built-up, the trends appear parallel; I test this proposition formally later in the robustness section.

Figure 4.3



Notes: Graph showing the evolution of the built-up area within grids for the treated and control localities before1993 (before any camps opened).

4.5 Results and Discussion

4.5.1 Main Result

The results from the baseline specification (4.1) are presented in Table 4.3. I find a positive and statistically significant effect of camp presence on built-up area. Proximity to a refugee camp is associated with a 0.006 sq. km. increase in built-up area within 100km of the camp, approximately a 6% increase from the average pre-camp built-up level.

I also show results from using different distance thresholds (20km and 50km). The results from using narrower treatment thresholds suggest that the effect is stronger closer to the camps. The estimates from these narrower thresholds should, however, be treated with caution, firstly because narrower thresholds come at the expense of lesser within variation since the number of control grids increases at the expense of treated, and secondly, because the hypothesis of common pre-trends is rejected. While the influx was unanticipated and choice of camp sites limited (see context), the government still wanted to restrict refugee movement. Additionally, the UNHCR emergency guidelines for site location included that the possible friction between local inhabitants and refugees that may arise from siting camps near urban areas should be considered. This can explain why using shorter thresholds results in a treated group that was urbanising at a different (slower) rate from the control.

Table 4.3: Main Result – effect of proximity to camps on built-up 1985-2015

	Built-up area (sq. km.)		
	(1)	(2)	(3)
camp20	0.023***		
	(0.007)		
camp50		0.013***	
		(0.004)	
camp100			0.006***
			(0.002)
Constant	0.102***	0.102***	0.102***
	(0.000)	(0.000)	(0.000)
Observations	415,789	415,789	415,789
R-squared	0.966	0.966	0.966
Year fixed effects	Yes	Yes	Yes
Grid fixed effects	Yes	Yes	Yes
Region specific trend	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at grid level.
*** p<0.01, ** p<0.05, * p<0.1

4.5.2 Subsample analysis

Looking at the entire sample may obscure where the effect is happening i.e., whether in rural areas or in areas that were already more built-up (a case of densification of areas that were already more built-up). To investigate the effect further, I next estimate the regression separately for rural localities and for those that were more built-up pre-camp.

I split the sample into rural and urban subsamples based on their pre-camp level of urban. Given the lack of clarity on the criteria used to designate localities as urban in Tanzania (see the earlier discussion in the data section), I do not rely on existing administrative designations of what is ‘urban’. Instead, I adopt a more systematic approach to identify localities that could be considered to have been urban pre-camp. I follow the method and thresholds jointly recommended by several agencies recently, in response to the widely acknowledged lack of consistent definition of what constitutes cities, towns and rural areas.[17] Specifically, I consider a grid to have been an urban cluster in 1990 if it had a population density greater than 300 persons per sq. km and a total population greater than 5000.

Wenban-Smith (2014) examines the applicability of such a density-based definition of urban in the context of Tanzania and concludes that it offers the advantage of consistency. The author acknowledges that population density criteria do not sufficiently recognise the complexities of urban development in developing countries such as Tanzania, where areas could be demographically urban but not functionally urban in terms of economic activity or access to facilities. Nevertheless, given the lack of clear criteria and sparse data for Tanzania, pre-camp population density does offer an objective and accepted way to identify localities that were urban pre-camp.[18]

Following this approach of identifying which areas were urban pre-camp, I conclude that 214 localities could be considered to have been urban prior to camps. On average, these localities do indeed have a higher level of built-up pre-camp (1.4 sq. km. compared to 0.1 sq.km in the rural sample). After splitting the samples, the validity of the identification assumption was assessed again for each subsample (See Tables C1-C2 and Figures C2-C3 in the Appendix).

Table 4.4 shows the result for the subsample of rural localities. The effect is positive, statistically significant and close in magnitude to the result in the full sample. Within the urban subsample,

---

[17] This method referred to as the "degree of urbanisation" was endorsed in March 2020 by the UN Statistical Commission. It was jointly developed by The European Commission, International Labour Organization (ILO), Food and Agriculture Organization (FAO), Organisation for Economic Co-operation and Development (OECD), UN-Habitat and the World Bank).

[18] This may raise the question of why I do not use this density-based measure as the outcome variable. While such an exercise would be useful at the very least to compare with the results from using the built-up measure, subnational population data is not available for all the years in the period of study. This is one of the advantages offered by panel data on built-up at such granular level.

however, I do not find evidence that camps had an urbanising effect (Table 4.5). The result (which implies a 1-standard error confidence interval of [-0.35, -0.11] and a 2-standard error confidence interval of [-0.47, 0.01]), does not provide strong evidence against a zero effect.

Table 4.4: Effect of proximity to camps on built-up 1985-2015
(Rural sample)

	Built-up area (sq. km.)		
	(1)	(2)	(3)
camp20	0.020***		
	(0.004)		
camp50		0.011***	
		(0.002)	
camp100			0.007***
			(0.001)
Constant	0.076***	0.076***	0.076***
	(0.000)	(0.000)	(0.000)
Observations	410,021	410,021	410,021
R-squared	0.952	0.952	0.952
Year fixed effects	Yes	Yes	Yes
Grid fixed effects	Yes	Yes	Yes
Region specific trend	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at grid level
*** p<0.01, ** p<0.05, * p<0.1

Table 4.5: Effect of proximity to camps on built-up 1985-2015
(Sample of existing urban areas)

	Built-up area (sq. km.)		
	(1)	(2)	(3)
camp20	-0.052		
	(0.145)		
camp50		-0.097	
		(0.102)	
camp100			-0.231*
			(0.120)
Constant	1.967***	1.973***	1.993***
	(0.004)	(0.008)	(0.015)
Observations	5,768	5,768	5,768
R-squared	0.976	0.976	0.976
Year fixed effects	Yes	Yes	Yes
Grid fixed effects	Yes	Yes	Yes
Region specific trend	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at grid level
*** p<0.01, ** p<0.05, * p<0.1

4.5.3 Long run effects

To examine the long run effect and specifically whether the effect persists after closure, as an alternative approach, I consider built up averages at two time periods – the 'universal' pre-camp and post-camp periods. I use the average built-up area in the period 1990-1992 (just before the refugee influx) and 2013-2015 (after all camps closed). I then difference over the two periods. Time-invariant factors drop out. The estimated equation is:

$$\Delta y_{ij} = \beta\left(\Delta Treat100_{ij}\right) + \gamma_j + e_{it} \qquad (4.2)$$

Where Δy_{ij} is the change in built-up area for grid i in region j between the two time periods. $Treat100_{ij}$ is the change in treatment status for grid i in region j between the two time periods and takes the value 1 for grids that were treated between the two periods and 0 for untreated. I also include a region fixed effect γ_j which controls for any unobserved region-level trends.

Identification is thus from within region-variation, allaying concerns of time-trending unobservables at the region level.

In the overall sample, I find a positive but not statistically significant effect. (See Table 4.6 column 3.) I also present results from using the 20km and 50km thresholds with the aforementioned caveat (see the earlier discussion in 6.1). The results from these narrower treatment thresholds (columns 1 and 2), suggest that the effect persisted closer to the camps.

In the rural sample, I find evidence that the positive urbanising effect is detectable even after the camps closed (Table 4.6B). Overall, this long-run comparison suggests that the effect persisted for rural areas and closest to camps. This is not surprising given that in section 4.5.2, it was found that the effect is within rural localities.

Table 4.6: Long run results (Long difference over pre-camp and post closure periods)

	Δ Built-up area		
	(1)	(2)	(4)
ΔTreat20	0.057***		
	(0.016)		
ΔTreat50		0.037***	
		(0.010)	
ΔTreat100			0.005
			(0.009)
Constant	0.013***	0.013***	0.013***
	(0.003)	(0.003)	(0.003)
Observations	14,395	14,395	14,395
R-squared	0.035	0.035	0.034
Region fixed effects	Yes	Yes	Yes

Robust standard Robust standard errors in parentheses, clustered at grid level.
*** p<0.01, ** p<0.05, * p<0.1

Table 4.6B: Long run results (Long difference over pre-camp and post closure periods), Rural

	Δ Built-up area		
	(1)	(2)	(4)
ΔTreat20	0.056***		
	(0.012)		
ΔTreat50		0.037***	
		(0.006)	
ΔTreat100			0.010**
			(0.005)
Constant	0.013***	0.013***	0.013***
	(0.003)	(0.003)	(0.003)
Observations	14,181	14,181	14,181
R-squared	0.050	0.050	0.048
Region fixed effects	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at grid level.
*** p<0.01, ** p<0.05, * p<0.1

4.5.4 Discussion

The results indicate that the refugee camps were associated with an increase in built-up area, but not in areas that were already more urban. It is insightful to contextualise this effect in population terms. In many different contexts, it has been found that built-up land cover grows faster than population (Angel et al., 2011; Bren d'Amour et al., 2017; European Commission et al., 2020; Goldblatt, Deininger, et al., 2018; Güneralp et al., 2020; Seto et al., 2010; Stow et al., 2016; Sudhira et al., 2004; Yin et al., 2005). Since the population in Tanzania was growing (population growth rate of 3.4% in 1992 (World Bank, 2020)), it is expected that built-up land would have increased, even in the absence of camps. What the result shows is that camps *were* associated with a 6% increase in built-up above and beyond the 'natural' increase in built-up associated with indigenous population growth.

This increase in built-up above and beyond the expected is indicative of a population increase over and above the indigenous or existing population growth. Recall that this cannot be attributed to the refugee population itself since built-up of camps themselves was excluded. The view that there was a population increase in these areas is consistent with reports of inward migration by entrepreneurs seeking to profit from new opportunities associated with the camps. While we may be concerned that camps induced outward migration of some locals, there is evidence, at least for Kagera, that refugee presence was associated with an overall lower likelihood of outward migration (Maystadt & Verwimp, 2014). Still, we cannot completely rule out out-migration of some locals. Maystadt& Verwimp (2014), for example, found that households self-employed in agriculture were more likely to migrate than other households and so, to date, the net effect has not been documented. The finding in this paper suggests that overall, the outward migration was outweighed by in-migration to exploit economic opportunities.

I turn next to economic activity, since built-up is not only an indicator of settlement but also non-agricultural activities. We can examine whether the result obtained is consistent with changes in employment or occupational mobility, given that the development of a rural non-farm economy plays a role in rural transformation (see literature review). For this purpose, it is worthwhile revisiting here some empirical findings in this regard. Maystadt & Verwimp (2014) finds that for Kagera, refugee presence did not substantially affect occupational mobility. Ruiz and Vargas-Silva (2015) note that it resulted in a higher likelihood of engaging in subsistence agricultural activities, a lower likelihood of being employed outside the household, and that it did not affect diversification of economic activities. Fieldwork from Landau (2004) reveals that opportunities for direct employment with the humanitarian organizations were few relative to the population (and more often were taken up by foreigners and more qualified Tanzanians from other parts of the country). Further, while there was secondary employment for locals hired to work for aid workers in capacities such as domestic staff, overall wages from these were not enough for residents to make substantial investments in land or productive equipment or to exit from subsistence activities. Landau (2004) also observes that most of the entrepreneurial activity was driven by Tanzanians coming from other regions and, therefore, was less likely to have a profound effect on the economic practices of the locals. From a sample of Kagera residents that excludes in-migrants, Maystadt and Verwimp (2014) notably find that while there was an increase in welfare overall, the self-employed in non-agricultural activities had a decline in welfare in contrast with reports of 'business boom'. The authors reconcile this finding by concluding that the entrepreneurial activity was by in-migrants. One can make two pertinent conclusions from existing empirical findings. First, the reports of boom in local entrepreneurship seem to be driven by outsiders. Second, while camp presence attracted some entrepreneurship and spurred local trade, on the whole there appears to have been no fundamental transition from agriculture to non-agricultural activities.

The small urbanising effect I find in this paper is consistent with these existing findings. In the robustness section, I probe further into the nature of economic activity using night-time lights. I fail to find substantial variation over time using this data. Although the limitations of night-light data are acknowledged, this lends some support to the view that the dominant nature of new economic activity associated with camps appears to have been localised trade. This may, in turn, explain why we observe no effect in the more urban areas. A plausible explanation is that such opportunities would be set up nearest to the point of demand – the more urban areas (sample 2 in the subsample analysis) were farther from the camps. This is also consistent with the fact that refugee movement was restricted; such trading opportunities could, therefore, not expand much further.

The potential role of institutional constraints in explaining the finding

The effect I find seems modest in light of the reports of booming economic activity and also the 'boomtown' urban growth described in other contexts (see literature review). One possible explanation is that institutional constraints limited the kind of rural transformation that may have been possible. The status of the camp and its inhabitants mean that it and the activity associated with it cannot be fully (or formally) integrated into the wider economy (Agier, 2002; Montclos & Kagwanja, 2000; Oka, 2011). No matter how long refugee camps last, camps are viewed as temporary – waiting zones – where refugees await either eventual repatriation to their countries of origin or resettlement in asylum countries. Indeed, hosting governments are typically wary of creating situations that may mean the refugees will settle permanently. It is only recently that there has been some recognition that this type of policy response is not consistent with the reality of the protracted nature of forced displacement. In the case of Tanzania, refugees had no right to work and to own property. Although some managed to leave camps, their movement was heavily curtailed. It is, therefore, conceivable that the type of economic activity that develops under this fettered environment would be localised trade dominated by non-tradeable goods and services that are not known to facilitate the more durable economic activities necessary for structural transformation to take place (Jedwab and Gollin, 2016).

A close examination of other contexts where camps are reported to be associated with high urban growth lends further support to the role of institutional constraints. The context in DRC in Büscher and Mathys (2019), for example, largely involves internally displaced individuals (IDPs). Although not in their areas of origin, IDPs experience less impediments to economic opportunities like access to work or the right to own property- opportunities that are usually unavailable to refugees in cross-border displacement situations like those in Tanzania. In short, camps for the IDPs have a different institutional context to refugee camps.

But it is important to note here that even among refugee camps, the context can vary from country to country. Using night-time lights, for example, Alix-Garcia et al., (2018) find increased economic activity around a Kenyan refugee camp. While the Kenyan Government shifted from a strategy of integration to encampment in the early 1990s, refugee networks are sufficiently strong and expansive. The Kakuma and Daadab camps have developed very strong commercial and trade linkages far into the interior of the country including right into the capital. Although this can be partly explained by the location of the camps on major trade routes to Sudan and Somalia which stimulated the growth of road transport links with the rest of the country, the ethnic ties between the Somali camp population and the indigenous Kenyan Somali population have played an instrumental role in the development of refugees' networks (Montclos & Kagwanja, 2000; Oka, 2011; Omata & Sterck, 2018; Veney, 2007). For example, the Hawala money transfer system practiced by the Somali community acts as an enforcement exchange mechanism which facilitates large volumes of trade even in the absence of formal contract enforcement. It is a centuries' old informal system of transferring money across long distances (Alix-Garcia et al., 2019; Ismail, 2007; Schaeffer, 2008). Thus, in Kenya, deep trade networks were able to develop despite the encampment policy because of age-old and shared cultural institutions such as the Hawala money transfer system.

The view that institutional frameworks are key forces for mediating rural transformation has solid basis (Berdegué et al., 2014). More broadly, the important role of institutions in determining economic outcomes is well established (Acemoglu, 2012; North, 1991). Institutions can be understood as "the rules of the game in a society or, more formally, the humanly devised constraints that shape human interaction" (North, 1990). Property rights and institutions connected with contractual engagements are particularly important in this context. First, secure property rights enable individuals to make investments in physical and human capital. Empirical evidence shows that secure property rights create incentives for long-term productivity enhancing investments that are conducive for structural transformation.[19] Secondly, institutions that reduce uncertainty and increase trust are important because they reduce transaction costs and thereby facilitate the expansion of trade and increased specialization, division of labour and technological improvements. Such institutions are believed to have played a role in the transformation of societies from rural to urban (North, 1991).

In the context of Tanzania, Tacoli & Agergaard (2017) note that besides transport and communication links, land-owning structures influence the expansion of small towns. The authors examine two villages that transformed into emerging urban centres following the country's liberalisation of agricultural production in the 1990s and highlight that the ability of migrant labor to acquire land has played a role in the development of Madizini town. Specifically, the privatization of the state-owned plantation and factory stimulated sugarcane production which led to the development of an out-grower scheme. Migrant labor that moved to the village were able to participate in the out-grower scheme, settled bought and rented land and this contributed to the demographic growth and economic development of the town. This contrasts with the camp setting. Unlike local migrant labor, restricted property rights have implications on refugees' ability to make long-term investments. This fosters confinement to localized trade and constrains the ability to invest in productive assets since security of assets is limited.

[19] For examples see (Besley, 1995; Deininger et al., 2014; Do & Iyer, 2008; Galiani & Schargrodsky, 2010; Holden et al., 2009; Jacoby et al., 2002).

Furthermore, refugees' restricted rights potentially increase transaction costs. As noted above, expansion of trade is facilitated by institutions that reduce uncertainty and increase trust. In early stages of trade, transaction costs are low as exchanges are governed by social networks and informal rules. However, geographical expansion of trade is associated with agency and enforcement problems (North, 1991). These require more effective institutions to reduce transaction costs which is unlikely to be fostered by camp settings. Curtailment of refugee mobility limits their ability to establish networks especially distant networks although institutions such as the Hawala system described in the case of the Kenyan camps above can be mediating factors. Additionally, the fact that refugees' situation is viewed as temporary may create a barrier to trust since there are no guarantees of their length of stay. This underscores the importance of institutional contexts and suggests potential avenues through which policies could improve the welfare of refugees while at the same time enabling host countries to benefit from camps.

4.6 Robustness checks

In this section, I assess the validity of the results in a variety of ways.

4.6.1 Testing for pre-trends

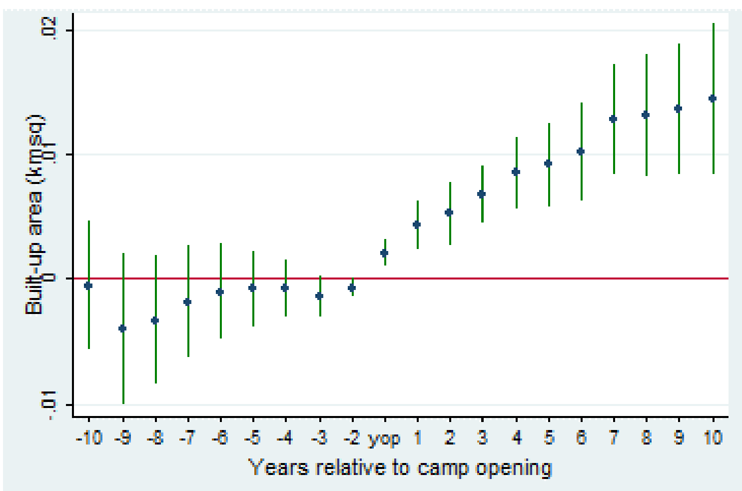
The identification assumption in the main specification implies that in the absence of camps, trends in built-up area would not have differed between the control and comparison group. Although this assumption of parallel trends is not directly testable, a test for trends in the pre-treatment period can lend confidence to its validity. To probe the validity of the identification assumption, therefore, I test the hypothesis that, prior to camps, there were no differential trends by proximity to future camp sites.

First, I estimate a variant of the main specification that includes leads and lags. Specifically, I interact treatment with dummies for the number of years since camp opened. The omitted dummy is the dummy for the year prior to camp opening (t=-1). This approach allows for a partial test of the identifying assumption. If the trends in built-up area between treatment and control grids are similar, then the coefficients for t<0 should not be statistically significant i.e., the difference-in-differences should not be statistically significantly different between the two groups in the pre-treatment period. It is important to note, however, that this may not be expected to hold for all pre-treatment years because once the influx happened in 1993 and the first wave of camps opened in 1993, it is not unrealistic to expect that there could be anticipatory effect in grids whose nearest camps were set up in subsequent years (1994,1995,1996,1997,1999). In this case, we may observe the effect just a few years before camp opening. In other words, since the 1993 influx was sudden and unexpected, once it happened, it may have caused anticipatory effects in grids that were treated in subsequent years. This approach allows me to check the timing and evolution of the effect. The main specification does not provide an understanding of the evolution of the effect such as how soon after camp opening is the effect and whether it is more pronounced in certain years.

The results are reported in Table 4.6 and Figure 4.4 shows the plot of the coefficients on the time since camp opening indicators. The coefficients are positive and statistically significant after the camps open. There is a clear rise in built-up after the camps open. Camp presence is associated with an increase in built-up area of about 0.005km^2 in the first two years rising to about 0.014km^2 after ten years (Figure 4.4). As a guide to the magnitude of the coefficients, the mean value of the built-up area (km^2) before 1993 was 0.1km^2, thus the effect would be an increase of 5% to 14% of the average pre-1993 level of built-up area.

All the pre-camp coefficients, by contrast, are individually not statistically significant. Note however, the hypothesis that they are jointly zero is rejected. As discussed above, this points to an anticipatory effect within grids that were treated later. As an additional robustness check, I consider the period before 1993 which is before any camps opened and, therefore, the cleanest and universal pre-camp period for all grids. There should be no anticipatory effects in this period. I restrict the data to this period and follow a placebo strategy to test whether there are any effects in this pre-camp period. The data for this covers 1985-1992. I assign placebo treatment considering 1985-1988 as pre-treatment and 1989-1992 as post-treatment. I report the result in Table 4.7. The coefficient for the placebo treatment is not statistically significant, thus providing evidence that prior to the camps, trends in built-up area did not depend on proximity to future camp sites. I then repeat the analysis using treatment year as 1993 for all grids. The results remain similar (See Table C3 in the Appendix).

Figure 4.4: Effect of camp presence within 100km on Built-up area (sq.km)



Note: The graph shows plots of coefficients and 95% CI for the regression in table 4.6. (As a guide on the magnitude of the coefficients, the mean value of the built-up area (sq.km) pre-1993 was 0.1 sq.km)

Table 4.6: Including leads and lags (treatment within 100km)

VARIABLES	Built-up area (sq. km)
Treat100*t==-10	-0.001
	(0.003)
Treat100*t==-9	-0.004
	(0.003)
Treat100*t==-8	-0.003
	(0.003)
Treat100*t==-7	-0.002
	(0.002)
Treat100*t==-6	-0.001
	(0.002)
Treat100*t==-5	-0.001
	(0.002)
Treat100*t==-4	-0.001
	(0.001)
Treat100*t==-3	-0.001*
	(0.001)
Treat100*t==-2	-0.001*
	(0.000)
Treat100*t==0	0.002***
	(0.001)
Treat100*t==1	0.004***
	(0.001)
Treat100*t==2	0.005***
	(0.001)
Treat100*t==3	0.007***
	(0.001)
Treat100*t==4	0.009***
	(0.001)
Treat100*t==5	0.009***
	(0.002)
Treat100*t==6	0.010***
	(0.002)
Treat100*t==7	0.013***
	(0.002)
Treat100*t==8	0.013***
	(0.002)
Treat100*t==9	0.014***
	(0.003)
Treat100*t==10	0.014***
	(0.003)
Observations	408,685
R-squared	0.967
Year fixed effects	Yes
Grid fixed effects	Yes
Region specific trend	Yes

Model includes a constant term. Robust standard errors in parentheses, clustered at grid level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4.7: Test for pre-treatment trends (placebo treatment on pre-camp periods)

VARIABLES	Built-up area (sq. km.)
treat100xplacebopost	-0.001
	(0.001)
Observations	115,160
R-squared	0.997
Year fixed effects	Yes
Grid fixed effects	Yes
Region specific trend	Yes

Note: The sample is all grids within 500km of a camp. Data is restricted to the pre-camp years 1985-1992. Model includes a constant term. Robust standard errors in parentheses, clustered at grid level.
*** p<0.01, ** p<0.05, * p<0.1

4.6.2 Investigating changes in economic activity further

To further investigate the nature and scale of economic activity, I examine changes in nightlight data, a commonly used proxy for economic activity. I maintain the same units of observation (grid cells) and examine luminosity level over time.

The nightlight data covering 1992-2012 is from the US Defence Meteorological Satellite Program (DMSP-OLS). It detects luminosity from outdoor and indoor use of light, fires and gas flares, on a scale of 0 to 63, where higher values imply higher night-time light intensities. I examine the stable lights which exclude ephemeral lights such as fires and background noise and only show lights from places with persistent lighting. Table C4 in the appendix shows the number of grid cells as a percentage of the total grid cells falling within select intervals on the 0-63 scale. I find little variation over time – the majority of the grid cells (about 97%) have zero luminosity over the entire time period.

A caveat with this data is, of course, that the absence of reported stable nightlights does not necessarily imply absence of nightlights, or absence of economic activity (Chen & Nordhaus, 2011; Henderson et al., 2012). In areas of low development, the low intensity of man-made lights, may not be distinguished from background noise, given the low resolution of the satellite. It should also be noted that the night-time satellites have a much lower resolution (approximately 1km) compared to the built-up data used in this paper (30m).

The absence of variation does, however, lend some support to the conclusion about the nature of activity – we can rule out that there were drastic changes, or at least changes sizeable enough to be detected with this measure.

4.7 Conclusion

This paper sits at the nexus of two pressing socio-economic phenomena – urban growth and forced displacement. The emergence of new urban agglomerations in previously rural areas has motivated a need to examine other drivers of urbanisation beyond the focus placed on rural exodus to existing urban areas. While evidence from some contexts suggest that an inflow of displaced populations could hasten urban growth in hosting regions, lack of data has meant that this channel has largely been unexamined. Understanding how forced displacement affects the urban growth process is critical to effectively managing the rapid pace of urban growth.

Using high resolution panel data on built-up area as a proxy measure for settlement and non-agricultural economic activity, this paper examined whether proximity to refugee camps has an urbanizing effect. This paper takes a panoramic view by examining whether the various effects documented to be associated with refugee camps (employment, health, infrastructure and service provision, new economic opportunities that attracted people or pushed out existing residents), could coalesce in the aggregate and be reflected in the transformation of surrounding areas.

This paper finds that camps were associated with a small urbanising effect in rural localities. Proximity to camps is associated with 0.006 sq.km. increase in built-up area, a magnitude comparable to a 3% increase in population. The camps do not appear to have had an urbanising effect on localities that were more urban. I have shown that the results are consistent with existing findings on the labour market, camp-induced migratory behaviour and reports on the nature of economic activity. Specifically, the findings are consistent with the explanation that while camp presence undoubtedly spurred local trade, affected the labour and goods markets and attracted some in-migrants, there appears to have been no fundamental shift away from subsistence agricultural activities – the type of drastic structural transformation that could hasten rural-urban transformation. The dominant nature of new economic activity seems to have been localised trade.

This paper has proffered the consideration that institutional constraints could explain why despite engendering conditions that could be conducive to rural transformation (ready markets for agricultural produce, improved transport infrastructure and low labour costs), we do not observe drastic effects. Data constraints presently limit the ability to conclusively demonstrate this robustly. Further detailed study of institutional constraints applied to the context of refugee camps would be a promising avenue for future research.

The analysis in this paper yields useful policy insights. Camps have been used a mechanism by host countries for the management of forcibly displaced populations even in protracted contexts. However, this analysis shows that the view that forcibly displaced persons can be siloed is not congruent with empirical findings. Camps have a transformative effect on hosting regions and these effects arise despite institutional and locational restrictions on the activity of their inhabitants. In fact, it is likely that by imposing restrictions hosting countries may be inadvertently fostering activity dominated by non-tradeable goods and services that are not known to facilitate structural transformation. Policy efforts should thus focus on how best to harness such potential. This is particularly relevant if countries are to avoid the trap of urbanizing without growth. Therefore, addressing such constraints is likely to improve the welfare of both the displaced and facilitate economic development of hosting regions.

4.8 Appendix

Figure C1

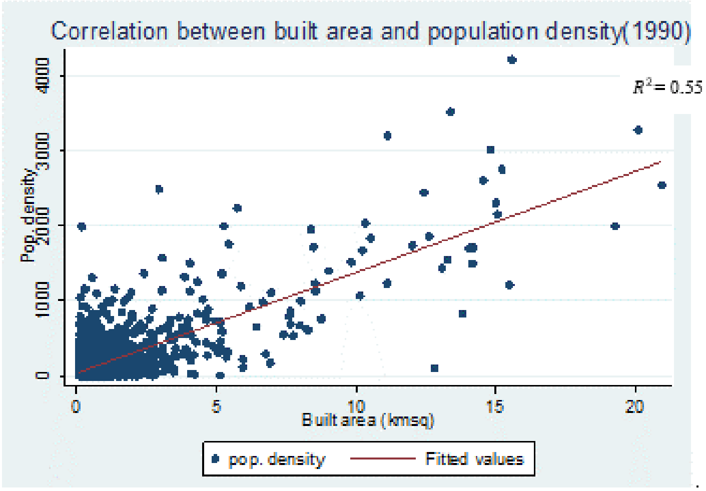


Figure C2: Urban sample

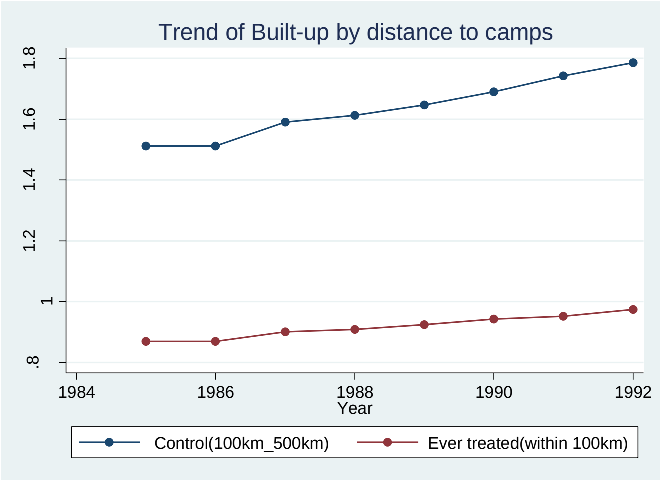
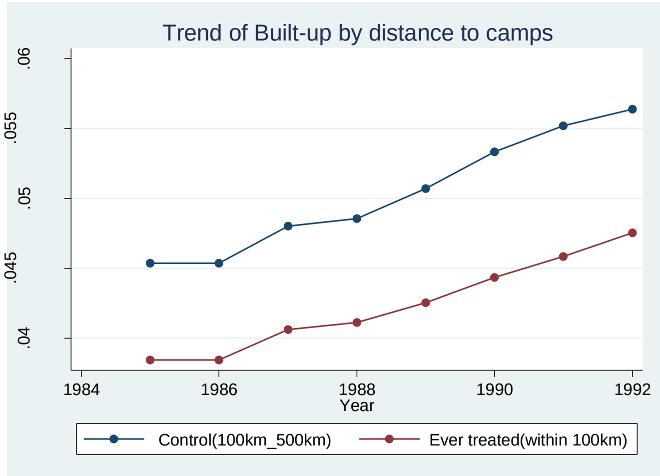


Figure C3: Rural sample



Note: Fig. C2 and C3 show the evolution of built-up area within grids for treated and control localities pre-1993 (before any camps opened) for urban and rural samples respectively.

Table C1: Test for pre-treatment trends (placebo treatment on pre-camp periods)
Sample of existing urban areas

VARIABLES	Built-up area (sq. km.)
treat100xplacebopost	-0.061
	(0.043)
Observations	1,712
R-squared	0.997
Year fixed effects	Yes
Grid fixed effects	Yes
Region specific trend	Yes

Note: The sample is all grids within 500km of a camp that were urban prior to camps. Data is restricted to the pre-camp years 1985-1992. Model includes a constant term. Robust standard errors in parentheses, clustered at grid level.
*** p<0.01, ** p<0.05, * p<0.1

Table C2: Test for pre-treatment trends (placebo treatment on pre-camp periods)
Rural sample

VARIABLES	Built-up area (sq. km.)
treat100xplacebopost	-0.001
	(0.001)
Observations	113,448
R-squared	0.995
Year fixed effects	Yes
Grid fixed effects	Yes
Region specific trend	Yes

Note: The subsample of grids within 500km of a camp that were rural prior to camps. Data is restricted to the pre-camp years 1985-1992. Model includes a constant term. Robust standard errors in parentheses, clustered at grid level.
*** p<0.01, ** p<0.05, * p<0.1

Table C3: Built-up and proximity to camps 1985-2015 (during camp operations)

	Built-up area (sq.km.)		
	(1)	(3)	(5)
treat20xpost1993	0.022***		
	(0.006)		
treat50xpost1993		0.013***	
		(0.004)	
treat100xpost1993			0.006**
			(0.003)
Observations	415,789	415,789	415,789
R-squared	0.966	0.966	0.966
Year fixed effects	Yes	Yes	Yes
Grid fixed effects	Yes	Yes	Yes
Region specific trend	Yes	Yes	Yes

Robust standard errors in parentheses, clustered at grid level
*** p<0.01, ** p<0.05, * p<0.1

Table C4: Number of grid cells falling within specified luminosity values on the 0-63 scale as a % of total observations (Stable lights)

	% of total observations							
Luminosity value	1992	1995	1997	2000	2005	2009	2010	2012
Missing Data	0.31%	0.31%	0.31%	0.31%	0.31%	0.31%	0.31%	0.31%
0	99.11%	98.41%	98.22%	98.03%	97.89%	97.29%	96.27%	96.64%
> 0 & ≤ 2	0.37%	0.85%	1.20%	1.06%	1.28%	1.70%	2.01%	1.86%
>2 & ≤ 6	0.10%	0.23%	0.15%	0.35%	0.33%	0.44%	0.91%	0.72%
>6 &≤ 10	0.10%	0.12%	0.10%	0.12%	0.15%	0.12%	0.23%	0.23%
>10 & ≤ 63	0.02%	0.08%	0.02%	0.14%	0.04%	0.14%	0.27%	0.25%
N	5173	5173	5173	5173	5173	5173	5173	5173

Chapter 5

Conclusion

This thesis examined three themes that are relevant to the economic development of refugee hosting regions – health, data and urban growth. It combines conventional and non-conventional data sources to address gaps in the literature and uncover long-term impacts of refugee camp presence on hosting regions.

The first paper of the thesis examined the long-term health effect of childhood exposure to refugee camps. While an existing study (Baez, 2011), finds a negative effect of camp exposure on children's health, long term impacts were unaddressed in the literature. I provide evidence for the first time in the literature that the effect persisted into adulthood. The effect is non-trivial and can be translated into a 2.9% to 5.9% reduction in adult hourly earnings for an individual that was three years of age at the time of exposure. I also find that there were differential effects by age and duration of exposure. The effect was larger for age groups 0-3 and 9-12, consistent with the established view that these are periods of greater sensitivity and strongest growth. Further analysis showed that the negative effect on child health was likely due to disease outbreaks in early camp phases.

For the subsequent generation, however, I find no difference in health outcome between those exposed to the camps and those born in the post-camp era and thus never exposed to the camps. This can be explained by the fact that disease outbreaks tended to be severe in the immediate aftermath of camp establishment. Over the years and by the time camps closed, these outbreaks had been brought under control.

One of the hypotheses in the paper had been that the closure of camps could have had an adverse health effect on children born in the post camp era because relief efforts cease and local economic opportunities that were hitherto generated by the camps shrink. I find that this is not the case − the cohorts born after the camps closed are not worse off than those cohorts born when the camps were in operation. I showed that this is consistent with evidence that the initial negative effect was likely due to outbreak of disease that subsided over time, such that there is little difference when the camp closes. Data limitations, however, meant it was not possible to entirely rule out that that the camp rehabilitation and reconstruction programme, established to fill the gap left by the withdrawal of humanitarian agencies, may have mitigated the adverse effects from the loss of benefits due to camp closures. The efficacy of the programme is an important consideration but is deferred to future research, and would be informative on how to avoid gaps in service delivery and losses of economic opportunities when humanitarian agencies withdraw from host regions.

Since forced migration is often at the centre of political discourse, it is important to caution against a narrow interpretation that solely focuses on negative health effects of early childhood exposure to refugee camps. Such interpretations of the results would miss a very important policy insight of the paper − a timely response to displacement situations is crucial. Specifically, implementation of public health measures as early as possible can mitigate adverse effects for camp residents and their hosting localities. Finally, due to data limitations, this paper focused on hosting communities. An equally important but under researched area which future efforts can address is the health of refugees.

The second paper of the thesis examined whether remote sensing imagery from a daytime satellite (the Landsat) can be used to predict local level outcomes using machine learning methods. The motivation of this paper was the prevailing shortage of sub-national data on welfare and economic activity which limits research on forced displacement and on developing countries more broadly.

First, I examined the case where a sub-national survey data in one time period can be used to predict agricultural occupation, wealth and consumption expenditure for other locations in the same time period. I found that that Landsat data, combined with weather and geophysical characteristics, can explain high variation in sub-national agricultural activity, wealth and consumption for new or unseen locations. Out-of-sample predictive accuracy, however, is low – errors are about a third of the mean of expenditure and close to half of the mean of agricultural occupation.

Second, I examined whether survey data in a certain year can be used to obtain predictions for another year for which survey data is not available. The results showed that the ability to generalise over time is mixed. Models trained on survey data from a specific year can predict consumption expenditure in another time period as well as in the cross-section, but not so for agricultural occupation. This can be explained by the fact that agricultural occupation is more variable over time than these other indicators.

This paper contributes to a growing literature that investigates whether machine learning techniques applied to non-conventional data sources can be used for predicting subnational economic well-being in developing countries. An important conclusion of the paper is that there is value to be gained from efforts that focus on reducing high dimensional satellite imagery into features that are interpretable and still predictive of economic variables. Such features may be more useful for analysis than raw data. The findings of the study also point to an important trade-off. The costs associated with conventional face-to-face survey data collection imply that it is usually only feasible to obtain detailed data for a few locations and at infrequent intervals. Non-conventional secondary sources such as satellites collect granular data, at low marginal costs and at regular intervals; one may, however, necessarily need to trade off some accuracy, especially to fill gaps in the most data deprived contexts. This is an active research area and future research will need to continue to address this trade off. One limitation of this study was the small sample size. It is possible that more flexible machine learning methods may have yielded higher predictive accuracy. However, these methods require large amounts of data for training and the limited survey sample size constrained their application in this paper's context. Furthermore, more flexible methods imply a trade-off between predictive power and model interpretability.

In the third and final paper, I examined forced displacement as a potential exogenous shock that could provide impetus to urbanisation. The study was motivated by the emergence of small towns and urban clusters in previously rural areas which has suggested the need to examine additional avenues outside of the conventional channel of rural-urban migration as a driver of urban growth. There is limited empirical evidence on whether refugee camps, as centres of resource inflows such as humanitarian aid, infrastructure investment, and local trade, propel urban growth − a study area that is known to be particularly data constrained.

Using satellite derived data on built-up area as a proxy for settlement and non-agricultural economic activity, I found that despite engendering conditions that may hasten urban growth, camp presence is associated with a small increase in built-up area above and beyond the 'natural' increase in built-up associated with indigenous population growth. The effect is only in the rural localities. I find no strong evidence that camp presence was associated with densification of localities that were already urban pre-camp. The magnitude of the effect is a striking finding that contrasts with reports of booming economic activity and the type of ‘boomtown’ urban growth described in other contexts. Further analysis with a different dataset on nightlights, however, confirmed there was no drastic effect on economic activity. I suggested that the key to understanding this finding is the institutional constraints within which camps exist. These constraints did not facilitate deep economic linkages with the rest of the country. These could explain why despite engendering conditions that can lead to rural transformation − ready markets for agricultural produce, improved transport infrastructure and low labour costs − we do not observe drastic effects in the case of Tanzania. Due to data constraints, however, it was not possible to conclusively demonstrate this robustly. Further detailed study of institutional constraints applied to the context of refugee camps would be a promising avenue for future research.

Although I showed that built-up area correlates with population density, one may still question whether built-up is an appropriate measure of urban growth. The study of whether rural areas transition to urban and/or whether existing urban areas expand, requires a measure of what is urban. I demonstrated that in the case of Tanzania, as with many other countries, existing administrative designations of what is ‘urban’ are fraught with inconsistencies and are arbitrarily determined. Data for urban metrics are likewise not regularly collected. I do not see any other remedy available at present, given these data constraints.

Overall, this thesis provides important insights into the current climate of unprecedented refugee flows. These insights are also relevant in view of the climate-induced forced displacement that is predicted to occur in the future. A major challenge currently confronting several countries, and one that is likely to remain relevant, is how to sustainably support host regions. The findings in this thesis provide several policy insights. First, the onset of forced displacement flows is a critical window of operations response to prevent negative effects from taking hold. This is particularly important as the conditions at the onset could have persistent effects. Second, efforts to address data shortages in these contexts are extremely important and can be partly addressed by non-conventional data sources that are collected more frequently and at higher granularity. Finally, understanding the way in which policy responses may interact with wider processes and development outcomes, not only in the short-term but also in the long-term, is important. The empirical evidence in this thesis can be informative for policies to facilitate economic development of hosting regions for the mutual benefit of both hosting communities and displaced populations.

Bibliography

Acemoglu, D. (2012). *Why nations fail: The origins of power, prosperity, and poverty*. Profile Books.

Adisa, J. (1996). *The Comfort of Strangers: The Impact of Rwandan Refugees in Neighbouring Countries*. UNCHS (Habitat).

Agergaard, J., Tacoli, C., Steel, G., & Ørtenblad, S. B. (2019). Revisiting Rural–Urban Transformations and Small Town Development in Sub-Saharan Africa. *The European Journal of Development Research*, *31*(1), 2–11. https://doi.org/10.1057/s41287-018-0182-z

Agier, M. (2002). Between War and City. *Ethnography*, *3*(3), 317–341.

Ahsan, M. N., & Maharaj, R. (2018). Parental human capital and child health at birth in India. *Economics & Human Biology*, *30*, 130–149. https://doi.org/10.1016/j.ehb.2018.06.006

Akresh, R., Bhalotra, S., Leone, M., & Osili, U. O. (2017). *First and Second Generation Impacts of the Biafran War* (Issue w23721)). National Bureau of Economic Research.

Akresh, R., Lucchetti, L., & Thirumurthy, H. (2012). Wars and child health: Evidence from the Eritrean–Ethiopian conflict. *Journal of Development Economics*, *99*(2), 330–340.

Akresh, R., Verwimp, P., & Bundervoet, T. (2011). Civil war, crop failure, and child stunting in Rwanda. *Economic Development and Cultural Change*, *59*(4), 777–810.

Alderman, H., Hoddinott, J., & Kinsey, B. (2006). Long term consequences of early childhood malnutrition. *Oxford Economic Papers*, *58*(3), 450–474.

Alix-Garcia, J., Bartlett, A., & Saah, D. (2013). The landscape of conflict: IDPs, aid and land-use change in Darfur. *Journal of Economic Geography*, *13*(4), 589–617. https://doi.org/10.1093/jeg/lbs044

Alix-Garcia, J., & Saah, D. (2010). The Effect of Refugee Inflows on Host Communities: Evidence from Tanzania. *The World Bank Economic Review*, *24*(1), 148–170. https://doi.org/10.1093/wber/lhp014

Alix-Garcia, J., Walker, S., Bartlett, A., Onder, H., & Sanghi, A. (2018). Do refugee camps help or hurt hosts? The case of Kakuma, Kenya. *Journal of Development Economics*, *130*, 66–83. https://doi.org/10.1016/j.jdeveco.2017.09.005

Alix-Garcia, J., Walker, S., & Bartlett, A. (2019). Assessing the direct and spillover effects of shocks to refugee remittances. *World Development*, *121*, 63–74. https://doi.org/10.1016/j.worlddev.2019.04.015

Angel, S., Chabaeva, A., Gitlin, L., Kraley, A., Parent, J., & Perlin, M. (2005). *The dynamics of global urban expansion* (No. 35563; pp. 1–205). The World Bank. http://documents.worldbank.org/curated/en/138671468161635731/The-dynamics-of-global-urban-expansion

Angel, S., Parent, J., Civco, D. L., & Blei, A. M. (2011). *Making room for a planet of cities*. Lincoln Institute of Land Policy.

Athey, S. (2019). The Impact of Machine Learning on Economics. In A. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The Economics of Artificial Intelligence: An Agenda* (pp. 507–547). University of Chicago Press.

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

Bagan, H., & Yamagata, Y. (2015). Analysis of urban growth and estimating population density using satellite images of nighttime lights and land-use and population data. *GIScience & Remote Sensing*, *52*(6), 765–780. https://doi.org/10.1080/15481603.2015.1072400

Baragwanath, K., Goldblatt, R., Hanson, G., & Khandelwal, A. K. (2019). Detecting urban markets with satellite imagery: An application to India. *Journal of Urban Economics*, 103173. https://doi.org/10.1016/j.jue.2019.05.004

Becker, S. O., & Ferrara, A. (2019). Consequences of forced migration: A survey of recent findings. *Labour Economics*, *59*, 1–16. https://doi.org/10.1016/j.labeco.2019.02.007

Beegle, K., Christiaensen, L., Dabalen, A., & Gaddis, I. (2016). *Poverty in a Rising Africa*. The World Bank.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-Dimensional Methods and Inference on Structural and Treatment Effects. *Journal of Economic Perspectives*, *28*(2), 1–23.

Berdegué, J. A., Carriazo, F., Jara, B., Modrego, F., & Soloaga, I. (2015). Cities, Territories, and Inclusive Growth: Unraveling Urban–Rural Linkages in Chile, Colombia, and Mexico. *World Development*, *73*, 56–71. https://doi.org/10.1016/j.worlddev.2014.12.013

Berdegué, J. A., Rosada, T., & Bebbington, A. J. (2014). The Rural Transformation. In B. Currie-Alder, R. Kanbur, D. M. Malone, & R. Medhora (Eds.), *International*

Development: Ideas, Experience, and Prospects (pp. 463–478). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199671656.003.0028

Besley, T. (1995). Property Rights and Investment Incentives: Theory and Evidence from Ghana. *The Journal of Political Economy*, *103*(5), 903–937.

Betts, A., Chaara, I., Naohiko, O., & Sterck, O. (2019). *Refugee Economies in Uganda: What Difference Does the Self-Reliance Model Make?* Refugee Studies Centre, University of Oxford.

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, *350*(6264), 1073–1076. https://doi.org/10.1126/science.aac4420

Braun, S., & Mahmoud, T. O. (2014). The employment effects of immigration: Evidence from the mass arrival of German expellees in postwar Germany. *Journal of Economic History*, *74*(1), 69–108.

Bren d'Amour, C., Reitsma, F., Baiocchi, G., Barthel, S., Güneralp, B., Erb, K.-H., Haberl, H., Creutzig, F., & Seto, K. C. (2017). Future urban land expansion and implications for global croplands. *Proceedings of the National Academy of Sciences*, *114*(34), 8939–8944. https://doi.org/10.1073/pnas.1606036114

Büscher, K. (2018). African cities and violent conflict: The urban dimension of conflict and post conflict dynamics in Central and Eastern Africa. *Journal of Eastern African Studies*, *12*(2), 193–210. https://doi.org/10.1080/17531055.2018.1458399

Büscher, K., & Mathys, G. (2019). War, Displacement and Rural–Urban Transformation: Kivu's Boomtowns, Eastern D.R. Congo. *European Journal of Development Research*, *31*(1), 53–71. https://doi.org/10.1057/s41287-018-0181-0

Calderón, V., & Ibáñez, A. M. (2009). *Labor Market Effects of Migration-related Supply Shocks: Evidence from Internally Displaced Populations in Colombia*. Universidad de los Andes, Facultad de Economía, CEDE.

Case, A., & Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, *116*(3), 499–532.

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, *113*(5), 893–903. https://doi.org/10.1016/j.rse.2009.01.007

Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, *108*(21), 8589–8594. https://doi.org/10.1073/pnas.1017031108

Cline, M. G., Meredith, K. E., Boyer, J. T., & Burrows. (1989). Decline of height with age in adults in a general population sample: Estimating maximum height and distinguishing birth cohort effects from actual loss of stature with aging. *Human Biology*, 415–425.

Cogneau, D., & Jedwab, R. (2012). Commodity price shocks and child outcomes: The 1990 cocoa crisis in Cote d'Ivoire. *Economic Development and Cultural Change*, *60*(3), 507–534.

Cohen, B. (2006). Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society*, *28*(1), 63–80. https://doi.org/10.1016/j.techsoc.2005.10.005

Coulson, A. (2013). Tanzania: A Political Economy. In *Tanzania* (2nd ed.). Oxford University Press. https://www-oxfordscholarship-com.ezproxy.sussex.ac.uk/view/10.1093/acprof:oso/9780199679966.001.0001/acprof-9780199679966-chapter-23

Dang, H., Jolliffe, D., & Carletto, C. (2019). Data gaps, data incomparability, and data imputation: A review of poverty measurement methods for data-scarce environments. *Journal of Economic Surveys*, *33*(3), 757–797. https://doi.org/10.1111/joes.12307

De Onis, M., Garza, C., Onyango, A. W., & Borghi, E. (2007). Comparison of the WHO child growth standards and the CDC 2000 growth charts. *Journal of Nutrition*, *137*(1), 144–148.

Deaton, A., & Zaid, S. (2002). Guidelines for Constructing Consumption Aggregates for Welfare Analysis. *World Bank Publications*, *135*. https://openknowledge.worldbank.org/bitstream/handle/10986/14101/multi0page.pdf?sequence=1&isAllowed=y

Deininger, K., Jin, S., Xia, F., & Huang, J. (2014). Moving Off the Farm: Land Institutions to Facilitate Structural Transformation and Agricultural Productivity Growth in China. *World Development*, *59*, 505–520. https://doi.org/10.1016/j.worlddev.2013.10.009

Devarajan, S. (2013). Africa's Statistical Tragedy. *Review of Income and Wealth*, *59*(S1), S9–S15. https://doi.org/10.1111/roiw.12013

Diao, X., Magalhaes, E., & Silver, J. (2019). Cities and rural transformation: A spatial analysis of rural livelihoods in Ghana. *World Development*, *121*, 141–157. https://doi.org/10.1016/j.worlddev.2019.05.001

Do, Q., & Iyer, L. (2008). Land Titling and Rural Transition in Vietnam. *Economic Development and Cultural Change*, *56*(3), 531–579. https://doi.org/10.1086/533549

Donaldson, D., & Storeygard, A. (2016). The View from Above: Applications of Satellite Data in Economics. *Journal of Economic Perspectives*, *30*(4), 171–198. https://doi.org/10.1257/jep.30.4.171

Douidich, M., Ezzrari, A., Van der Weide, R., & Verme, P. (2016). Estimating Quarterly Poverty Rates Using Labor Force Surveys: A Primer. *The World Bank Economic Review*, *30*(3), 475–500. https://doi.org/10.1093/wber/lhv062

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, *91*(4), 795–813.

Dyson, T. (2011). The Role of the Demographic Transition in the Process of Urbanization. *Population and Development Review*, *37*, 34–54. https://doi.org/10.1111/j.1728-4457.2011.00377.x

Earth Engine Data Catalogue. (2020). *USGS Landsat 7 Surface Reflectance Tier 1*. https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LE07_C01_T1_SR

Engstrom, R., Hersh, J., & Newhouse, D. (2017). Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being. *Policy Research Working Paper 8284. World Bank*. https://doi.org/10.1596/1813-9450-8284

Eriksson, J., Adelman, H., Borton, J., Christensen, H., Kumar, K., & Suhrke, A. (1996). *The international response to conflict and genocide: Lessons from the Rwanda experience: Synthesis report.* Steering Committee of the Joint Evaluation of Emergency Assistance to Rwanda.

European Commission, ILO, FAO, OECD, UN-Habitat, & World Bank. (2020). *A recommendation on the method to delineate cities, urban and rural areas for international statistical comparisons*. https://ec.europa.eu/eurostat/cros/system/files/bg-item3j-recommendation-e.pdf

Fay, M., & Opal, C. (2000). Urbanization without Growth: A Not-So-Uncommon Phenomenon. *Policy Research Working Paper 2412. World Bank.*

Fox, S. (2014). Urbanization as a Global Historical Process: Theory and Evidence from sub-Saharan Africa. *Population and Development Review*, *38*(2), 285–310.

Fox, S. (2017). Mortality, migration, and rural transformation in Sub-Saharan Africa's urban transition. *Journal of Demographic Economics*, *83*(1), 13–30. https://doi.org/10.1017/dem.2016.29

Galasso, E., Wagstaff, A., Naudeau, S., & Shekar, M. (2016). *The Economic Costs of Stunting and How to Reduce Them*. World Bank.

Galiani, S., & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, *94*(9–10), 700–729. https://doi.org/10.1016/j.jpubeco.2010.06.002

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2018). Big data and big cities: The promises and limitations of improved measures of urban life. *Economic Inquiry*, *56*(1), 114–137. https://doi.org/10.1111/ecin.12364

Goldblatt, R., Deininger, K., & Hanson, G. (2018). Utilizing publicly available satellite data for urban research: Mapping built-up land cover and land use in Ho Chi Minh City, Vietnam. *Development Engineering*, *3*, 83–99. https://doi.org/10.1016/j.deveng.2018.03.001

Goldblatt, R., Heilmann, K., & Vaizman, Y. (2019). Can Medium-Resolution Satellite Imagery Measure Economic Activity at Small Geographies? Evidence from Landsat in Vietnam. *Policy Research Working Paper 9088. World Bank.*

Goldblatt, R., Stuhlmacher, M. F., Tellman, B., Clinton, N., Hanson, G., Georgescu, M., Wang, C., Serrano-Candela, F., Khandelwal, A. K., Cheng, W.-H., & Balling, R. C. (2018). Using Landsat and nighttime lights for supervised pixel-based image classification of urban land cover. *Remote Sensing of Environment*, *205*, 253–275. https://doi.org/10.1016/j.rse.2017.11.026

Goldblatt, R., You, W., Hanson, G., & Khandelwal, A. K. (2016). Detecting the Boundaries of Urban Areas in India: A Dataset for Pixel-Based Image Classification in Google Earth Engine. *Remote Sensing*, *8*(8), 634. https://doi.org/10.3390/rs8080634

Gollin, D., Jedwab, R., & Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth*, *21*(1), 35–70. https://doi.org/10.1007/s10887-015-9121-4

Green, R. H. (1995). *Making Whole Again: Human Development Consequences of Hosting Refugees in Kagera Region, Tanzania.*

Güneralp, B., Lwasa, S., Masundire, H., Parnell, S., & Seto, K. C. (2017). Urbanization in Africa: Challenges and opportunities for conservation. *Environmental Research Letters*, *13*(1), 015002. https://doi.org/10.1088/1748-9326/aa94fe

Güneralp, B., Reba, M., Hales, B. U., Wentz, E. A., & Seto, K. C. (2020). Trends in urban land expansion, density, and land transitions from 1970 to 2010: A global synthesis. *Environmental Research Letters*, *15*(4), 044015. https://doi.org/10.1088/1748-9326/ab6669

Haggblade, S., Hazell, P. B. R., & Reardon, T. (2007). *Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World*. International Food Policy Research Institute.

Haggblade, S., Hazell, P., & Brown, J. (1989). Farm-nonfarm linkages in rural sub-Saharan Africa. *World Development*, *17*(8), 1173–1201. https://doi.org/10.1016/0305-750X(89)90232-5

Haggblade, S., Hazell, P., & Reardon, T. (2010). The Rural Non-farm Economy: Prospects for Growth and Poverty Reduction. *World Development*, *38*(10), 1429–1441. https://doi.org/10.1016/j.worlddev.2009.06.008

Han, M. C. (2009). *Dynamic and Lasting Impacts: Socioeconomic Effects of Protracted Refugee Camps on Host Communities in Tanzania*.

Harris, J. R., & Todaro, M. P. (1970). Migration, Unemployment and Development: A Two-Sector Analysis. *The American Economic Review*, *60*(1), 126–142. JSTOR.

Harttgen, K., Lang, S., & Seiler, J. (2019). Selective mortality and the anthropometric status of children in low- and middle-income countries. *Economics & Human Biology*, *34*, 257–273. https://doi.org/10.1016/j.ehb.2019.04.001

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition* (2nd ed.). Springer-Verlag. https://doi.org/10.1007/978-0-387-84858-7

Haughton, J., & Khandker, S. R. (2009). *Handbook on Poverty and Inequality*. World Bank. https://openknowledge.worldbank.org/handle/10986/11985

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, *102*(2), 994–1028. https://doi.org/10.1257/aer.102.2.994

Henderson, J. V., & Turner, M. A. (2020). Urbanization in the Developing World: Too Early or Too Slow? *Journal of Economic Perspectives*, *34*(3), 150–173. https://doi.org/10.1257/jep.34.3.150

Hirvonen, K. (2014). Measuring catch-up growth in malnourished populations. *Annals of Human Biology*, *41*(1), 67–75.

Hoddinott, J., Maluccio, J., Behrman, J. R., Martorell, R., Melgar, P., & Quisumbing, A. R. (2011). *The consequences of early childhood growth failure over the life course*. International Food Policy Research Institute Discussion Paper 1073.

Hodler, R., & Raschky, P. A. (2014). Regional Favoritism. *The Quarterly Journal of Economics*, *129*(2), 995–1033. https://doi.org/10.1093/qje/qju004

Holden, S. T., Deininger, K., & Ghebru, H. (2009). Impacts of Low-Cost Land Certification on Investment and Productivity. *American Journal of Agricultural Economics*, *91*(2), 359–373. https://doi.org/10.1111/j.1467-8276.2008.01241.x

Ismail, A. A. (2007). Lawlessness and economic governance: The case of *hawala* system in Somalia. *International Journal of Development Issues*, *6*(2), 168–185. https://doi.org/10.1108/14468950710843415

Jacobsen, K. (2005). *The economic lives of refugees*. Kumarian Press.

Jacoby, H. G., Li, G., & Rozelle, S. (2002). Hazards of Expropriation: Tenure Insecurity and Investment in Rural China. *American Economic Review*, *92*(5), 1420–1447. https://doi.org/10.1257/000282802762024575

Jansen, B. J. (2016). The protracted refugee camp and the consolidation of a 'humanitarian urbanism.' *International Journal of Urban and Regional Research*.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.

Jedwab, R., Christiaensen, L., & Gindelsky, M. (2014). Rural Push, Urban Pull and... Urban Push? New Historical Evidence from Developing Countries. *Institute for International Economic Policy Working Paper Series*.

Jedwab, R., & Moradi, A. (2016). The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa. *Review of Economics and Statistics, 98*(2), 268–284. https://doi.org/10.1162/REST_a_00540

Kenya says go home. (2016, May). *The Economist*. https://www.economist.com/middle-east-and-africa/2016/05/14/kenya-says-go-home

Kilic, T., Serajuddin, U., Uematsu, H., & Yoshida, N. (2017). *Costing Household Surveys for Monitoring Progress Toward Ending Extreme Poverty and Boosting Shared Prosperity*. World Bank. https://doi.org/10.1596/1813-9450-7951

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction Policy Problems. *American Economic Review, 105*(5), 491–495. https://doi.org/10.1257/aer.p20151023

Lall, S. V. (2017). Renewing expectations about Africa's cities. *Oxford Review of Economic Policy, 33*(3), 521–539. https://doi.org/10.1093/oxrep/grx038

Landau, L. B. (2004). Challenge without Transformation: Refugees, Aid and Trade in Western Tanzania. *Journal of Modern African Studies, 42*(1), 31–59.

Lazaro, E., Agergaard, J., Larsen, M. N., Makindara, J., & Birch-Thomsen, T. (2019). Urbanisation in Rural Regions: The Emergence of Urban Centres in Tanzania. *European Journal of Development Research, 31*(1), 72–94. https://doi.org/10.1057/s41287-018-0185-9

Lupala, J. M. (2015). The effects of closure of refugee camps on livelihood activities to the surrounding villages: A case of Mtabila Camp in Tanzania. *Case Studies Journal*, *4*(11).

Mamo, N., Bhattacharyya, S., & Moradi, A. (2019). Intensive and extensive margins of mining and development: Evidence from Sub-Saharan Africa. *Journal of Development Economics*, *139*, 28–49. https://doi.org/10.1016/j.jdeveco.2019.02.001

Marconcini, M., Metz-Marconcini, A., Üreyen, S., Palacios-Lopez, D., Hanke, W., Bachofer, F., Zeidler, J., Esch, T., Gorelick, N., Kakarla, A., & Strano, E. (2019). *Outlining where humans live – the world settlement footprint 2015*. arXiv preprint arXiv:1910.12707.

Marx, B., Stoker, T. M., & Suri, T. (2019). There Is No Free House: Ethnic Patronage in a Kenyan Slum. *American Economic Journal: Applied Economics*, *11*(4), 36–70. https://doi.org/10.1257/app.20160484

Mathiassen, A. (2013). Testing Prediction Performance of Poverty Models: Empirical Evidence from Uganda. *Review of Income and Wealth*, *59*(1), 91–112. https://doi.org/10.1111/roiw.12007

Maystadt, J. F., & Duranton, G. (2019). The development push of refugees: Evidence from Tanzania. *Journal of Economic Geography*, *19*(2), 299–334. https://doi.org/10.1093/jeg/lby020

Maystadt, J. F., Mueller, V., Van Den Hoek, J., & van Weezel, S. (2020). Vegetation changes attributable to refugees in Africa coincide with agricultural deforestation. *Environmental Research Letters*, *15*(4), 044008. https://doi.org/10.1088/1748-9326/ab6d7c

Maystadt, J. F., & Verwimp, P. (2014). Winners and Losers among a Refugee-Hosting Population. *Economic Development and Cultural Change*, *62*(4), 769–809. https://doi.org/10.1086/676458

Michalopoulos, S., & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, *81*(1), 113–152. https://doi.org/10.3982/ECTA9613

Michalopoulos, S., & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, *129*(1), 151–213. https://doi.org/10.1093/qje/qjt029

Micklewright, J., & Ismail, S. (2001). What can child anthropometry reveal about living standards and public policy? An illustration from Central Asia. *Rev. Income Wealth*, *47*(1), 65–80.

Montclos, M.-A. P. D., & 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), 205–222. https://doi.org/10.1093/jrs/13.2.205

Moradi, A. (2010). Nutritional status and economic development in sub-Saharan Africa, 1950–1980. *Economics & Human Biology*, *8*(1), 16–29.

National Bureau of Statistics. (2014). *2012 Population and Housing Census (PHC): Basic demographic and socio-Economic Profile, Tanzania Zanzibar*.

NBS. (2000). *Tanzania Demographic and Health Survey 1999*. National Bureau of Statistics (NBS) (Tanzania) and Macro International Inc.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

North, D. C. (1991). Institutions. *Journal of Economic Perspectives*, *5*(1), 97–112.

OECD. (2020). *Africa's urbanisation dynamics 2020: Africapolis, mapping a new urban geography*. OECD. https://doi.org/10.1787/b6bccb81-en

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).

Omata, N., & Sterck, O. (2018). *Refugee Economies in Kenya*. Refugee Studies Centre, University of Oxford.

Ongpin, P. (2008). Refugees in Tanzania – Asset or Burden? *Journal of Development and Social Transformation*, *5*, 11.

Osafo-Kwaako, P. (2012). *Essays in economic history and development*. Doctoral dissertation, Harvard University.

Otsuka, K. (2007). The rural industrial transition in East Asia: Influences and implications. In *S. Haggblade, P. Hazell, & T. Reardon (Eds.), Transforming the rural nonfarm economy – Opportunities and Threats in the Developing World*. John Hopkins University Press.

Palacios-Lopez, D., Bachofer, F., Esch, T., Heldens, W., Hirner, A., Marconcini, M., Sorichetta, A., Zeidler, J., Kuenzer, C., Dech, S., Tatem, A. J., & Reinartz, P. (2019). New Perspectives for Mapping Global Population Distribution Using World Settlement Footprint Products. *Sustainability*, *11*(21), 6056. https://doi.org/10.3390/su11216056

Perez-Heydrich, C., Warren, J. L., Burgert, C. R., & Emch, M. E. (2013). *Guidelines on the use of DHS GPS data* (No. 8; Spatial Analysis Reports). ICF International.

Picker, G., & Pasquetti, S. (2015). Durable camps: The state, the urban, the everyday: Introduction. *City*, *19*(5), 681–688. https://doi.org/10.1080/13604813.2015.1071122

Potts, D. (2013). Rural-Urban and Urban-Rural Migration Flows as Indicators of Economic Opportunity in Sub-Saharan Africa: What Do the Data Tell Us? *Migrating out of Poverty RPC Working Paper 9. Migrating out of Poverty Consortium, University of Sussex*.

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. *Proceedings of the 3rd ERTS Symposium*, *1*, 309–317.

Ruiz, I., & Vargas-Silva, C. (2015). The Labor Market Impacts of Forced Migration. *American Economic Review*, *105*(5), 581–586. https://doi.org/10.1257/aer.p20151110

Ruiz, I., & Vargas-Silva, C. (2016). The labour market consequences of hosting refugees. *Journal of Economic Geography*, *16*(3), 667–694. https://doi.org/10.1093/jeg/lbv019

Rutinwa, B. (2017). The end of Asylum? The Changing Nature of Refugee Policies in Africa. In H. Lambert (Ed.), *International Refugee Law* (1st ed., pp. 35–64). Routledge. https://doi.org/10.4324/9781315092478-3

Sanghi, A., Onder, H., & Vemuru, V. (2016). *" Yes" in My Backyard? The Economics of Refugees and Their Social Dynamics in Kakuma, Kenya*. World Bank.

Schaeffer, E. C. (2008). Remittances and Reputations in Hawala Money-Transfer Systems: Self-Enforcing Exchange on an International Scale. *Journal of Private Enterprise*, *24*(1).

Schultz, T. P. (2002). Wage gains associated with height as a form of health human capital. *American Economic Review*, *92*(2), 349–353.

Serajuddin, U., Uematsu, H., Wieser, C., Yoshida, N., & Dabalen, A. (2015). Data Deprivation: Another Deprivation to End. *Olicy Research Working Paper 7252. World Bank*. https://doi.org/10.1596/1813-9450-7252

Seto, K. C., Sánchez-Rodríguez, R., & Fragkias, M. (2010). The New Geography of Contemporary Urbanization and the Environment. *Annual Review of Environment and Resources*, *35*(1), 167–194. https://doi.org/10.1146/annurev-environ-100809-125336

Silwal, A. R. (2016). *Three essays on agriculture and economic development in Tanzania*. Doctoral dissertation, University of Sussex.

Stow, D. A., Weeks, J. R., Shih, H., Coulter, L. L., Johnson, H., Tsai, Y.-H., Kerr, A., Benza, M., & Mensah, F. (2016). Inter-regional pattern of urbanization in southern Ghana in the first decade of the new millennium. *Applied Geography*, *71*, 32–43. https://doi.org/10.1016/j.apgeog.2016.04.006

Strauss, J., & Thomas, D. (1998). Health, Nutrition, and Economic Development. *Journal of Economic Literature*, *36*(2), 766–817.

Stren, R., & Halfani, M. (2001). The Cities of Sub-Saharan Africa: From Dependency to Marginality. In R. Paddison, *Handbook of Urban Studies* (pp. 466–485). SAGE Publications Ltd. https://doi.org/10.4135/9781848608375.n29

Sudhira, H. S., Ramachandra, T. V., & Jagadish, K. (2004). Urban sprawl: Metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation*, *5*, 29–39. https://doi.org/10.1016/j.jag.2003.08.002

Tacoli, C., & Agergaard, J. (2017). *Urbanisation, rural transformations and food systems:The role of small towns*. International Institute for Environment and Development (IIED).

Tanzanian Affairs. (1994). *Benaco – Tanzania's second city?* https://www.tzaffairs.org/1994/09/benaco-tanzanias-second-city/

Thomas, D., Lavy, V., & Strauss, J. (1996). Public policy and anthropometric outcomes in the Cote d'Ivoire. *Journal of Public Economics*, *61*(2), 155–192.

Thomas, D., & Strauss, J. (1997). Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, *77*(1), 159–185.

Thome, K., Markham, B., Barker, J., Slater, P., & Biggar, S. (1997). Radiometric Calibration of Landsat. *Photogrammetric Engineering & Remote Sensing*, *63*(7), 853–858.

Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

UNDP. (2009). *A second life for refugee camps in north-western Tanzania*. United Nations Development Programme. https://www.undp.org/content/undp/en/home/news-centre/news/2009/12/17/a-second-life-for-refugee-camps-in-north-western-tanzania.html

UNFPA. (2011). *State of World Population 2011: People and Possibilities in a World of 7 Billion*. UN. https://doi.org/10.18356/cbd2a655-en

UNHCR. (2011). *UNHCR Global Report: United Republic of Tanzania*. https://www.unhcr.org/afr/4fc880a20.pdf

UNHCR. (2019a). *Global Trends: Forced Displacement in 2018*. United Nations High Commission for Refugees.

UNHCR. (2019b). *UNHCR Population Statistics Database*. http:// popstats.unhcr.org.

UNHCR. (2020). *Global Trends: Forced Displacement in 2019*. United Nations High Commission for Refugees.

UNHCR. (2021, March). *UNHCR Statement on the Government of Kenya's intention to close Dadaab and Kakuma refugee camps*. UNHCR Kenya.

https://www.unhcr.org/ke/19849-unhcr-statement-on-the-government-of-kenyas-intention-to-close-dadaab-and-kakuma-refugee-camps.html

United Nations Department of Economic and Social Affairs [UNDESA]. (2019). *World urbanization prospects: The 2018 revision*.

United Nations Economic Commission for Africa [UNECA]. (2017). An overview of urbanization and structural transformation in Africa. In United Nations Economic Commission for Africa, *Economic Report on Africa 2017* (pp. 64–92). UN. https://doi.org/10.18356/a80a5e9d-en

U.S. Geological Survey (USGS). (2019a). *Landsat 7 (L7) Data Users Handbook*. https://www.usgs.gov/core-science-systems/nli/landsat/landsat-7-data-users-handbook

U.S. Geological Survey (USGS). (2019b). *Landsat Collection 1 Level 1 Product Definition*. https://www.usgs.gov/media/files/landsat-collection-1-level-1-product-definition

Veney, C. (2007). *Forced migration in Eastern Africa: Democratization, structural adjustment, and refugees*. Palgrave Macmillan.

Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., & Sachdev, H. S. (2008). Maternal and child undernutrition: Consequences for adult health and human capital. *Lancet*, *371*(9609), 340–357.

Weidmann, N. B., & Schutte, S. (2017). Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, *54*(2), 125–140. https://doi.org/10.1177/0022343316630359

Wenban-Smith, H. (2014). *Urbanization in Tanzania, Phase one.* London School of Economic and Political Science.

Whitaker, B. E. (1999). Changing opportunities: Refugees and host communities in western Tanzania. *Journal of Humanitarian Assistance, 4*(1), 1–23.

WHO. (1995). *Physical status: The use and interpretation of anthropometry. Report of a WHO expert committee.* (Vol. 854, pp. 1–452) [Tech. Rep. Ser.]. World Health Organization.

WHO. (2005). *The World Health Report: 2005: Make Every Mother and Child Count*. World Health Organization.

Wolfcarius, E. (2008). *UNHCR Prepares Transition Ahead While Camps Are Closing Retrieved from*. https://www.unhcr.org/news/latest/2008/6/4857de172/

Wolff, E., Grippa, T., Forget, Y., Georganos, S., Vanhuysse, S., Shimoni, M., & Linard, C. (2020). Diversity of urban growth patterns in Sub-Saharan Africa in the 1960–2010 period. *African Geographical Review*, *39*(1), 45–57. https://doi.org/10.1080/19376812.2019.1579656

World Bank. (2017). *Forcibly displaced: Toward a development approach supporting refugees, the internally displaced, and their hosts.*

World Bank. (2020). *Population growth (annual %)—Tanzania*. https://data.worldbank.org/indicator/SP.POP.GROW

Yin, J., Stewart, D., Bullard, S., & MacLachlan, J. (2005). Changes in urban built-up surface and population distribution patterns during 1986-1999: A case study of Cairo, Egypt. *Computers, Environment and Urban Systems*, *29*, 595–616. https://doi.org/10.1016/j.compenvurbsys.2005.01.008

Young, N. E., Anderson, R. S., Chignell, S. M., Vorster, A. G., Lawrence, R., & Evangelista, P. H. (2017). A survival guide to Landsat preprocessing. *Ecology*, *98*(4), 920–932. https://doi.org/10.1002/ecy.1730

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*(3), 583–594. https://doi.org/10.1080/01431160304987

Zhou, Y. Y. (2014). *How Refugees Drive Preferences for Citizenship Exclusion: Evidence from Sub-Saharan Africa*. APSA 2014 Annual Meeting Paper. https://ssrn.com/abstract=2454409