IZA

IZA DP No. 3468

Civil Wars beyond their Borders: The Human Capital and Health Consequences of Hosting Refugees

Javier E. Baez

April 2008

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Civil Wars beyond their Borders: The Human Capital and Health Consequences of Hosting Refugees

Javier E. Baez
*Syracuse University
and IZA*

ABSTRACT

Civil Wars beyond their Borders:
The Human Capital and Health Consequences of Hosting Refugees[*]

Between 1993 and 1994, extremist militia groups carried out the extermination of ethnic Tutsis and moderate Hutus in the genocides of Burundi and Rwanda. Nearly one million people were killed and thousands were forcibly uprooted from their homes. Over the course of a few months, Kagera – a region in northwestern Tanzania – received more than 500,000 refugees from these wars. This region is home to a series of geographic natural barriers, which resulted in variation in refugee intensity. I exploit this variation to investigate the short and long run causal effects of hosting refugees on the outcomes of local children. Reduced-form estimates offer evidence of adverse impacts almost 1.5 years after the shock: a worsening of children's anthropometrics of 0.3 standard deviations, an increase of 15 to 20 percentage points in the incidence of infectious diseases and an increase of roughly 7 percentage points in mortality for children under five. I also exploit intra- and inter-cohort variation and find that childhood exposure to this massive arrival of refugees reduced height in early adulthood by 1.8 cm (1.2%), schooling by 0.2 years (7.1%) and literacy by 7 percentage points (8.6%). Designs using the distance from the village to the border with Rwanda as an alternative instrumental strategy for refugee intensity support the findings. The estimates are robust across a variety of samples, specifications and estimation methods and provide evidence of a previously undocumented indirect effect of civil wars on the well-being of children and subsequent economic growth in refugee-hosting communities.

JEL Classification: O10, O12, O15

Keywords: civil conflicts, refugees, children, human capital, health, Africa

Corresponding author:

Javier E. Baez
Department of Economics
Maxwell School of Citizenship and Public Affairs
Syracuse University
Syracuse, NY 13244-1020
USA
E-mail: jebaez@maxwell.syr.edu

1. Introduction

The human suffering and economic consequences inflicted by internal unrest and civil conflicts are evident to all. Wars not only produce large death tolls and disrupt human and physical capital accumulation but damage the environment, weaken institutions, limit political governance and erode civil liberties. And their horrors uproot entire populations from their lands. These displacements are common and on the rise in sub-Saharan Africa, the poorest region of the world. Figures from the United Nations High Commission of Refugees (UNHCR, 2005) confirm this trend as they show a more than a threefold increase in the number of forced migrants in the region during the last two decades.

Despite the enormous importance of assessing the socioeconomic impacts of wars on refugees, this paper focuses attention on other individuals that are also affected by these displacements and often do not receive the same public coverage: the host communities. Massive population shocks such as those triggered by most civil conflicts in sub-Saharan Africa can threaten the well-being of permanent residents. The risks include disease outbreaks, food and land scarcity, unsafe drinking water, overburdened school and health care facilities, environmental degradation, crime and other security concerns. In the particular case of young children, the combined effect of these mechanisms can affect early childhood development. On the other hand, the arrival of refugees can increase the welfare of their hosts by bringing resources from international humanitarian assistance into the community and, thus, increasing economic activity and, perhaps, public infrastructure. Since economic theory does not provide an unambiguous prediction as to how these opposite mechanisms interact, the answer to which one dominates is at the end an empirical subject.

Notwithstanding the growing number of refugee crises, the effects of civil conflicts on both refugees and local inhabitants remain largely unexplored in the empirical literature. One possible explanation to this is the difficulty in isolating the true effect of such crises because of the large number of omitted factors that might produce unconvincing associations between refugees or permanent resident status and post-conflict socioeconomic outcomes. In addition, a more plausible restriction to test a causal pathway on specific groups of people stems from the lack of rich data at the household level. Finding a natural experiment can overcome these limitations. In this paper I

investigate the short and long run causal effects of hosting refugees on children's outcomes in rural Tanzania by exploiting the population flows from the genocides in Burundi (1993) and Rwanda (1994) as a natural experiment.

On October, 1993, not long after being democratically elected, the Hutu president of Burundi was assassinated and the country entered a period of civil strife. The Hutu community responded to the assassination by killing thousands of Tutsi civilians; the Burundian army (still controlled by Tutsis) extended the cycle of revenge with a mass-killing of Hutu civilians. It was reported at the time that between 100,000 and 150,000 people died in the genocide and approximately 800,000 Burundians were forced to flee their homes. The violent ethnic rivalry spread out months later to Rwanda. On April 6, 1994, the airplane carrying Juvénal Habyarimana and Cyprien Ntyamina (presidents of Rwanda and Burundi, respectively) was shot down as it approached Kigali –the capital and largest city of Rwanda– and both presidents died in the plane crash. After this episode, extremist militia groups started the extermination of ethnic Tutsis and moderate Hutus. Over the course of three months, between 800,000 and one million people died (more than one tenth of Rwanda's population) and at least two million people fled into bordering countries.

Tanzania has historically been welcoming displaced people for decades and is the host to the largest refugee populations in Africa. Kagera, a region located in the northwestern corner of the country, was particularly affected by the enormous influx of refugees from the conflicts in Burundi and Rwanda described above. It was estimated that 250,000 Burundians fled into the Kigoma and Kagera regions during the first wave of immigrants between the end of 1993 and the beginning of 1994. However, the largest flow began a few months later (April 28, 1994), when more than 250,000 Rwandans forming lines of up to 12 kilometers long flooded into Kagera in less than 24 hours, the largest and fastest refugee movement ever according to officials from UNHCR (1997). Overall, the Kagera region hosted a population of over 500,000 refugees from the Rwandan genocide. Together with the influx of Burundians, refugees outnumbered local citizens of the two main recipient provinces, with populations of roughly 320,000 and 250,000 inhabitants at the pre-war period.

The empirical investigation of the human capital consequences of this forced migration on locals is the main focus of this study. However, a key identification concern in this type of studies is the potentially non-random nature of the population inflows into host communities. Traditional

3

estimation methods might be biased away from the true impact if areas relatively poorer and with unobservable disadvantages were systematically more or less affected by the waves of displaced persons. In order to address this and other endogeneity problems, this study exploits a series of regionally distinctive topographic characteristics and natural barriers that appear to produce extensive exogenous refugee variation across provinces in Northwestern Tanzania. Anecdotal accounting and qualitative evidence indicate that refugee settlements (both organized and self-settled) were largely concentrated in a subset of provinces of Kagera: Ngara and Karagwe. My identification strategy benefits from two main geographic features of the region under study. First, Ngara and Karagwe border the west side of a chain of mountainous terrains as well as swamps, lakes and large forest and game reserves that create a band of unoccupied areas separating these districts from other provinces of Kagera. The placement of these types of soils and bodies of water right in the middle of the region – restricting population movements – and the division it creates into western (treatment) and eastern villages (control) can be viewed as a natural experiment. Second, Tanzania is home to Lake Victoria, the largest lake in Africa. Since Kagera is situated on the southwest shore of the lake and it has a width of about 240 kms, two largely rural regions (Mara and Mwanza) bordering the southeastern shoreline of the lake provide additional natural control groups that are thought to be immune to the arrival of refugees. In addition to research designs that employ binary variables to identify these sources of geographical variation, I also make use of geo-referenced data to exploit the distance from the village to the border with Rwanda as an alternative instrumental design for village-level cross-sectional refugee intensity.

A double and triple difference analysis that exploits variation in village-level refugee intensity and cohort exposure together with pooled cross-sectional surveys and longitudinal household level data is employed to implement my identification strategies. Although the well-being of several groups of people may be well affected by large inflows of forced and very poor migrants, I focus on the net short and long run impacts on local children in particular. Reduced-form results offer robust evidence of adverse impacts nearly 1.5 years after the two genocides: a worsening of children's anthropometrics (a shift to the left of the overall distribution of nearly 0.3 standard deviations), an increase of 15 to 20 percentage points in the incidence rate of infectious diseases and a dramatic increase of approximately 7 percentage points in under age five mortality rates. I also exploit intra- and inter-cohort variation and find that childhood exposure to the massive arrival of refugees

4

reduced height in early adulthood (i.e. ten years after the genocides) by 1.8 cm (1.2%), schooling by 0.2 years (7.1%) and literacy by 7 percentage points (8.6%).[1] I also investigate the robustness of these findings to several confounding factors and find that they do not seem to be driven by a declining trend in the health status of children in affected areas prior to the two historical episodes, by endogenous migration, attrition and humanitarian assistance, by misspecification biases from the linearity assumption, and the selection implied by the determinants of survival.

Even that these findings are consistent across a variety of samples, specifications, estimation methods and robustness checks, they differ from the results obtained in one of the few previous related works. Alix-Garcia (2006) showed that proximity to refugee camps was associated with an increase in trade within the village and some indicators of welfare (i.e. electricity, televisions, refrigerators and vehicles) among local residents in Western Tanzania. However, these latter findings can be reconciled with negative impacts on specific groups of people arising from shortages of food, water, health care or from environmental damage and new epidemics –which were all concerns at the post-conflict and migration time– that can be hidden behind more active local markets.

The remainder of the paper proceeds as follows. The next section gives background that reviews the previous relevant literature and briefly documents the genocides and their forced inflows of population into northwestern Tanzania. Section 3 describes the datasets used for the empirical analysis in the paper. Section 4 presents the identification strategies employed and reports summary statistics, regression results examining the short and long run effects of the shocks as well as a set of robustness exercises. Section 5 discusses the potential mechanisms that could explain the direction and magnitude of the impacts. Finally, Section 6 concludes.

2. Background

2.1 Existing Evidence

The number of refugees in the world has increased in the last three decades. Surprisingly, there has not been a corresponding increase in the attention to this issue paid by economists and,

---

[1] Percentage changes in long run outcomes variables calculated as a proportion of the average of control group in 2004.

thus, little research on the welfare impacts of forced migration on either displaced persons or receiving communities is available. Despite this, there are two main lines of analysis in the literature that provide useful background and intuition for the subject of this study. One of these lines documents the main political, socioeconomic, demographic and ethnic determinants of forced migration in the developing world (Collier and Hoeffler, 1998; Hatton and Williamson, 2002 and 2004; Engel and Ibanez, 2007). The second strand literature –more informative for the goals of this paper– has itself focused extensively on two main aspects of migration, mostly non-forced migration. One group of studies examines its impact on a wide range of immigrant outcomes associated with their economic assimilation to the host region (Borjas, 1987; Hansen and Lofstrom, 2004; Bolesta, 2006; Bevelander and Lundh, 2007). Other research efforts have addressed the effects of migration on the labor market performance of native populations (Card, 1990; Card and Altonji, 1991; Lalonde and Topel, 1991; Pischke and Velling, 1994; Borjas, Freeman and Katz, 1997; Friedberg, 2001; Fairlie and Meyer, 2003). Both lines of research have arrived, in general, at mixed results. That is, migrants can affect both positively and negatively the outcomes of recipient populations in different domains.

A review of the existing evidence reveals few empirical papers examining the impact of refugees on the well-being of the groups of people in receiving regions. More recently, however, two papers have devoted attention to this subject. Alix-Garcia (2006) investigates the impact of refugee inflows from the genocides of Burundi (1993) and Rwanda (1994) on host districts in Western Tanzania. Although very close in the nature of the population shock and geographical focus of my paper, this study employs a different research design and examines other issues such as changes in household-level expenditures and proxies of welfare (i.e. dirt floors, electricity, television, refrigerators and vehicles) in receiving regions. Her study initially shows that, perhaps as a result of more economic activity, local prices of some agricultural commodities in villages close to refugee camps exhibited higher volatility. Next, a double difference analysis is used to suggest the existence of beneficial effects on local residents through a fall in household's expenditures on food products and a positive change in some of their welfare indicators. In contrast, a second related work by Montalvo and Reynal-Querol (2007) finds that the arrival of refugees has had harmful impacts on health and the burden of death in recipient countries by increasing their transmission and persistence of malaria. The authors gathered macro data to construct a long panel for 135 countries

and performed several specifications and robustness checks (including instrumental variable estimates) to argue that each 1,000 additional refugees were connected to between 2,000 and 2,700 new cases of malaria in host countries.

A few case studies based largely on qualitative data compiled by the UNCHR in refugee camps and nearby areas from several parts of the world also offer insights into the different sorts of effects arising from the contacts between forced migrants and domestic inhabitants. Yet, both this sort of approach and the more formal analysis on the consequences of such interactions have reached mixed conclusions. Furthermore, the long run effects of these population shocks on locals remain largely unexplored empirically.

My paper exploits topographic characteristics and other extensive natural barriers as plausible sources of differential refugee intensity in otherwise similar Tanzanian villages from the Kagera region to shed some light on the subject. In this sense, this study adds to the literature in three ways. First, I measure the impact of the arrival of refugees on a set of outcomes that not only are unexplored but also offer a more plausible approximation of the net effects of the shock. It can be argued that both beneficial and unfavorable impacts developed due to the sharp increase in population. For instance, the inflows of food aid into affected areas increased the supply of basic staples and in principle are not expected to hurt pre-shock levels of consumption and nutritional condition of domestic children. However, this may not be true if this shift in the supply of food did not offset the huge increase from the demand side that would have resulted in deep food scarcity, as some newspapers and anecdotal records of the time seems to testify. In addition, the presence of food aid could introduce incentives for local households to decrease their own production of food. Along the same lines, while more economic activity can raise the welfare among locals, it is also known that larger price variability can diminish it as well. Overall, one can identify others numerous mechanisms working in opposite directions and the use of variables that summarize potentially negative and positive effects such as the nutritional and health status of children may be more informative.

Second, given the timing of the surveys and research design employed in this study, I am able to examine the dynamics of the effects by measuring both the short and long term impacts of the shock. Empirical evidence that identifies the persistence of detrimental effects in early childhood

development hints at the troubling long-run consequences of flows of war refugees. More specifically, there is an emerging literature which leaves no doubt about the strong association between malnutrition and poor health during the first years of life with the level of cognitive, motor and social-emotional development of children and their subsequent school progress and labor market performance through to adulthood (e.g. Carneiro and Heckman, 2003; Heckman and Masterov 2004; Almond and Mazumder, 2005; Behrman and Rosenzweig, 2005; Heckman, 2006). Furthermore, the lastingness of childhood developmental deficits as a result of civil wars can be argued to have impacts on later economic growth.

Finally, the results of this paper are valuable evidence to support and mobilize global action in the scale required to prevent childhood deficiencies. Most programs of humanitarian assistance are devoted today to handling the extremely critical needs of refugees in several parts of the developing world. But, as the findings seem to suggest, a more comprehensive response from international donors and aid agencies is needed such that the human welfare of host communities is guarded against current and future civil wars as well.

2.2 Genocides, Flows of Refugees and the Kagera Region

The world refugee population increased continuously from the 1970's to the early 1990's, a trend that totaled a six fold increase during the period and reached a peak of nearly 18 million refugees in 1992. Since then, the number of forced migrants has fallen to levels of around 11 million in 2006. An overwhelming majority of such human displacement has resulted from civil wars in developing countries and Africa has a remarkable share: by the mid 2000's it had nearly one third of the worldwide number of refugees and internally displaced persons and approximately 60% of them were children under the age of 18 (UNHCR 2007; U.S. Committee for Refugees and Immigrants, 2007).

Among the several civil wars that broke out in the 1990's in Africa, the conflicts of Burundi in 1993 and Rwanda in 1994 largely attracted the attention of the international community due to the high level of atrocity. Although the first of these conflicts had roots in ethnic rivalries from the early 1970's, the conflict actually escalated into a bloody struggle at the end of 1993, after the assassination of the first democratically elected president. Over 100,000 Burundians perished and more than

800,000 fled into Rwanda, Tanzania, and Democratic Republic of Congo (RDC, before know as Zaire). The second of the genocides erupted a few months later, on April 6, 1994. The deadly attack on the presidential aircraft that carried the presidents of Rwanda and Burundi triggered a three-month killing campaign that claimed the lives of approximately one million people. In the aftermath of the genocide, more than one million refugees escaped to find haven in neighboring countries, particularly in RDC and Tanzania.

As the rest of the world witnessed the horrors of these wars, Kigoma and Kagera –two regions of Tanzania– were flooded with thousands of refugees. The Kagera region is located in the northwestern corner of the country and is the remotest region from Dar es Salaam, the main administrative center of the country. Kagera has borders with Rwanda and Burundi to the west and is divided into five districts (Bukoba, Muleba, Biharamulo, Karagwe and Ngara). The region is mostly rural and its major economic activity is agriculture, which engages more than 80% of the working population.[2] At the time of the two genocides, Kagera was the poorest among all the 20 regions in Tanzania, with an income per capita of roughly US$280 and more than 50% of the population living below the US$1 poverty line (National Bureau of Statistics of Tanzania, 2001). Kagera's level of poverty is further compounded by unpaved and badly maintained roads, poor water infrastructure (less than 10% of the population has access to safe water), limited health care and educational facilities and very low level of coverage of electricity and sanitation.

Tanzania has hosted refugees for more than three decades. Yet, given the sharp increase of population inflows in the 1990's the country changed its "open door policy" for a more restrictive program, which focused in promoting temporary protection of forced migrants and subsequent repatriation. In the case of the Burundian and Rwandan refugees, efforts were made to settle them into rudimentary camps along the country's western border. However, the heavy burden of hosting such massive influx of refugees made it impossible to meet all their critical needs. The conditions of Benaco, the largest shelter opened for the crisis and located in the Ngara district of the Kagera region, gives some sense of the scale of this tragedy. At one point, the camp hosted up to 400,000 refugees and experienced several problems such as water and food scarcity, lack of waste collection,

---

[2] Coffee, cotton and tea are the most important cash crops while bananas, beans and maize are the main food crops.

epidemics (bloody dysentery, malaria, pneumonia, meningitis and measles), malnutrition and an increase in criminal activities.

According to the UNHCR (2005), Kagera alone received between 500,000 and 600,000 refugees from the two genocides, which was equivalent to nearly 38% of its local population in 1993 (1,580,000 inhabitants). There is extensive evidence from the historical literature suggesting that many refugees blended with local populations. In fact, anecdotal stories from government officials and reports from agencies and non-governmental organizations seem to indicate that the efforts to separate migrants from Tanzanians were undermined for a variety of reasons, and I focus on three in particular.[3] First, a large number of refugees (mainly Burundians) ignored the restrictions and informally settled in areas outside the camps such as forests and river banks. Second, many refugees had to leave the delimited zones during the day for neighboring villages to supplement the shortage of the camps. Often they engaged in agricultural activities in areas typically cultivated by host communities, and also cultivated their own crops and collected firewood in areas restricted to locals. Third, refugees from Rwanda and Burundi have a similar ethnic ancestry (Hutu and Tutsi) to the populations in western Tanzania and thus, migrants were often hosted in the towns by relatives and other groups of people ethnically linked to them. Overall, it appears that villages close to refugee camps were not isolated from the massive arrival of migrants and host communities were frequently exposed to the increased presence of refugee settlements. Whether they benefited or not from such interactions is a question I address in the next sections.

3. Data

The database employed in this paper comes basically from two main sources. First, I use a pool of cross-sectional data at the household level from the 1992 and 1996 Tanzania Demographic and Health Surveys (TDHS) to estimate the short-term effects of the shock. The surveys –representative at the national and regional level– collected information from women aged 15-49, men aged 15-60 and their children in randomly selected households on a wide range of topics, including fertility, infant mortality education, health, nutrition, family planning and other basic characteristics of the

---

[3] I was very fortunate to hold a personal talk with Frederick Sumaye, a former Prime Minister of Tanzania (from 1995 to 2005) on April 12, 2007 in which he described in detail many of the issues associated with the crisis of refugees in the country, particularly in the Kagera region.

households (e.g. employment, housing amenities, assets).[4] The 1992 TDHS used a random cluster process to draw a total sample of 8,327 households. The 1996 TDHS, the survey followed the same sample frame and clusters covered in the 1992 TDHS and reached a total of 7,969 households. The samples used in the empirical models are restricted in basically three ways. First, the analysis is limited to households located in the Kagera, Mwanza and Mara regions, both the latter providing additional comparison groups, as will be explained in more detail below. Second, in order to match the timing of the events, only households from the 1996 TDHS that resided in the same cluster at least since 1993 were added to the final pooled sample. Third, the observational units in all the estimations are restricted to children under five. When pooled, the final sub-sample contains between 500 and 2,200 children, depending on the model, dependent variables and strategy implemented. For this part the outcomes of interest are under five infant mortality and child's anthropometrics and morbidity.

Since identification is derived from the geographic coordinates of the village, I merged the resulting dataset with the GIS module (restricted access) from the 1996 TDHS, which includes spatially geo-referenced information from all the clusters in both surveys. The GIS data is used to locate the relative position of the clusters with respect to topographic characteristics and bodies of water in the region of study, as well as to calculate the distance from each cluster to the border of the Kagera region with Rwanda.

For the long run impacts, I use the Kagera Health and Development Survey (KHDS), a five-round longitudinal household survey conducted in all the districts of the Kagera region. The first four waves were collected almost yearly between 1991 and 1994, while the latest round was carried out in 2004. The KHDS is a very rich dataset that contains modules on household demographics, education, health, anthropometrics, household activities, household and individual expenditures, local markets, among others, as well as questionnaires at the community, school and health facility-level. The survey interviewed a stratified random sample of more than 800 households from 51 clusters in 49 villages, for a total of nearly 5,500 individuals. Efforts were made to re-interview all

---

[4] One disadvantage of these datasets is that they do not ask questions on household incomes and expenditures. I try to overcome this by including in the estimation proxies of family expenditures such as the educational attainments of the parents and indicator variables of asset information on household ownership of several durables (e.g. radio, television, refrigerator, motorcycle, car, bicycles and appliances) as well as information on dwelling characteristics.

respondents in subsequent rounds and track those individuals who moved to a different village, region or country. Here I focus on the cohort of children between 10 and 15 years old by 2004 (who were born during 1989-1994) that lived in a certain village of the Kagera region at least since 1993. Finally, because this section of the study partially follows an identification design similar to that of the short term analysis, I also use the GIS portion of the KHDS (restricted access) that contains geo-reference information of all villages in the surveys. In addition, these datasets are also matched with information on altitude and historical climate data (monthly precipitation and number of days it rained in the month) collected by 21 weather stations in the Kagera region between 1980 and 2004.[5] I present and discuss sample means and standard deviations of the pooled samples for the key variables of the analysis as well as the identification strategies, empirical results and robustness checks in the next section.

4. Empirical Strategy and Findings

4.1 Identification

Appropriate data must be found in order to examine the effect of an immigration event on the native population. However, a lack of cross-sectional variation in the number of refugees at the village level has made this task hard. An additional and even more challenging issue is the difficulty of finding appropriate control groups for host communities. For example, poorer regions that are highly disadvantaged in some unobservable domains may be relatively more likely to have undernourished children and host more refugees (e.g. limited institutions and systems to control their arrival or assist them). If that is the case, any association between the number of refugees and welfare measures of domestic inhabitants will likely be driven by the underlying correlation between the level of development of host regions and these two outcomes rather than by the causal effect of the exposure to refugees. I attempt to overcome both the data limitations and potentially non-random nature of refugee intensity by exploiting some particular geographic features of the Kagera region and –for the short term impacts– of two neighbor regions. In addition to this design, I also exploit intra- and inter-cohort variation in childhood exposure/responsiveness to the shock to shed light on its long run effects. All of these strategies are implemented by performing double and

---

[5] I matched villages in the KHDS sample to the nearest weather station in order to proxy for their rainfall records.

triple difference analysis (D-D and D-D-D) on pre- and post-shock data, namely a pool of cross-sectional surveys and longitudinal data for the short and long term analysis, respectively.

With respect to the geographic conditions, I first take advantage of the existence and location of a chain of mountainous terrains, forest, game reserves, lakes and swamps that create a division between western and eastern Kagera. Evidence presented below indicates that their geographic placement produces a plausible natural barrier for population displacements at some latitudes (see Figure 2, Map (a) Topographic characteristics, Map (b) Land Use). Following this approach jointly with GIS data, I define treated districts as those located in the western Kagera (Karagwe and Ngara), which have borders with Rwanda and Burundi, while control villages are therefore the towns in the eastern districts of the Kagera region (Bukoba, Muleba and Biharamulo, see Figure 3, Map (d) Administrative Divisions). A couple of reasons underpin this strategy. First, a particular feature of civil conflicts in Africa and Asia is that the refugees that succeed in fleeing war into neighboring countries often settle very close to the borders of their own countries, i.e. the districts of Karagwe and Ngara. Second, the fact that the Kagera region is very small[6] and its districts were largely rural and similar in several baseline key factors (e.g. demographics, economic development, economic activities, levels of schooling, weather, morbidity, etc.) supports the use of the eastern districts as proper counterfactuals of the western districts. However, despite the exogenous placement of these topographic barriers, it is important to note that the eastern districts of Kagera were not completely isolated from the flows of refugees and still received a relatively smaller number of migrants. Hence, this approach is in reality capturing the variation in the treatment intensity of both zones, i.e. the difference between "high refugee intensity" and "low refugee intensity" areas.

Notwithstanding the lack of data to measure the exact number of refugees hosted at the village level, I use qualitative data from the KHDS to provide evidence in favor of the variation arising from the natural experiment described above. The fifth round of the KHDS (2004) asked community leaders to assess the intensity of the influx of refugees in their villages after 1994 (i.e. closeness to refugee resettlements) in a scale from one to four. Based on this information, Table 1 presents the proportion of respondents that reported having refugee resettlements in the proximity of their towns. Consistent with the eastern-western story, 83.3% and 60% of the village leaders from clusters

---

[6] With a width of nearly 120 km and length of 260 km that represent a total land area of approximately 29,000 square km, similar to the area of the states of Massachusetts or Vermont in the U.S.

in the west districts of Ngara and Karagwe claimed having refugees settled nearby. In contrast, no more than 20% of their counterparts from the "low intensity" districts indicated having displaced people settled in their neighboring areas.

My second strategy is also constructed from the geographic position of the villages and uses the distance from the clusters to the borders with Rwanda as a proxy for cluster-level refugee intensity.[7] The rationale behind this source of variation is that the distance to the borders in these largely rural areas with poor roads and transport systems reflect, to a large extent, geographical and economic constraints for migrants rather than socioeconomic predictors of children's outcomes within Kagera. Moreover, this strategy is suitable to estimate the impact of a continuous differential treatment as compared to the average treatment intensity effect among treated districts obtained from the "eastern-western" design. Figure 3 shows the relationship between the responses of community leaders stated previously and the distance of the cluster to the border. As expected, the scatter plot suggests the existence of a negative correlation between the two. In other words, those villages farther away from the conflict zones appear to have hosted fewer refugees (a correlation coefficient of -0.703).

I also employ these geographic identification strategies with two other Tanzanian regions that are thought to provide an additional control group. As shown in Figure 4, the large water area of Lake Victoria defines the border of the whole Kagera region to the east with Mara and Mwanza, two largely rural regions. With a surface of 68,800 square kms and a width of nearly 240 kms in the east-west direction, the lake produces a natural barrier for forced migrants trying to reach eastern regions away from Kagera. The lake is not heavily used to transport people within Tanzania. Ferries crossing the lake to travel between Tanzania, Uganda and Kenya are its most important means of transport. Although small ferries from the port in Bukoba also connect the Kagera region with Mwanza, there is not evidence that they were used to carry refugees to that part of the country, in part because of the high costs and the long distance between the main ports. To exploit this setting, I perform regressions that use the villages of Kagera as the treated units (i.e. "western") and the

---

[7] I used the village's distance to the border with Rwanda because the Rwandan genocide seems to have produced the largest temporary refugee migration shock to the Kagera region. In addition, given the geographic position of the Kagera region with respect to Rwanda and Burundi (both countries border the western band of Kagera), the alternative empirical design based on the village's distance to the latter yields basically identical quantitative effects (not shown).

villages of Mara and Mwanza as the comparison towns (i.e. "eastern"), and also regression with the treatment definition derived from the cluster-border distance design.

Finally, the empirical analysis uses data from the Kagera region together with intra- and inter-cohort variation in exposure/responsiveness to the shock to investigate the persistence of the effects in the long run. First, I compare treated and untreated children (0 and 5 years old in 1994) before and after (10 years later) the arrival of migrants. Second, I contrast this cohort with older cohorts within western and eastern districts and also between western and eastern districts, under the assumption that the latter cohorts were less sensitive to the shock by the timing of the events. Therefore, the second strategy (D-D-D) is expected to difference out changes across cohorts that affected the two regions similarly (by comparing across eastern and western regions), cohort invariant effects between districts (by comparing across cohorts) and individual fixed effects (by exploiting the longitudinal nature of the data).

4.2 Short Run Impacts

I assess the short-run effects of the influx of refugees on the status of local children by looking at the differential change of several indicators associated with three outcomes: nutrition, morbidity and infant mortality. Due to the nature of these outcomes and data restrictions,[8] I focus on children aged 0-4 for the first two outcomes and children aged 0-5 for infant mortality. To estimate the impacts, I pool TDHS cross-sectional data from the 1992 (pre-shock) and 1996 (after-shock) surveys.

The left panel of Table 2 presents summary statistics of baseline key variables comparing households from high refugee intensity (western=1) and low refugee intensity (western=0) treatment areas. Basically, there are not significant pre-shock differences between high and low treatment areas, except for the fact that the latter have household heads that were roughly 3.5 years older. Before the two genocides, families in Kagera overall had on average almost seven members, 1.4 children under five, around 81% of the parents were married and women were 18.6 years old when they experienced their first birth. The socioeconomic indicators of the sample are just a reflection of

---

[8] Malnourishment and anthropometric measures are more developmental in early childhood, particularly among children under five. Furthermore, DHS datasets often collect anthropometric measures only for children 0-48 months old at the time the survey is conducted.

the high levels of poverty prevalent in this region and in most of the rural areas of sub-Saharan Africa. Approximately 25% and 1% of the households have access to clean water and electricity, respectively; only 7.6% reported living in a house with hard floor and 28% said they owned a radio. Fathers have, on average, 5.4 years of schooling, while the school attainment of mothers is just 3.8 years. The right panel presents the analogous comparison for the bigger sample, that is between households in treatment regions (Kagera=1) and households in the control regions of Mara and Mwanza (Kagera=0). Some relevant differences seem to exist between the two groups. On the one hand, treated households are smaller (almost one fewer member), have younger household heads (nearly 3.3 years) and have better access to prenatal care. On the other hand, 56% of the households in control regions have access to safe water, which represents a large and significant difference of 30 percentage points with respect to the coverage of households in Kagera. Considering the importance that in theory these variables may have in determining the dependent variables of this study, the empirical analysis below includes pre-shock measures of these and other variables to control for any differential baseline socioeconomic variation between the two groups.

The general specification of the base regression equations is as follows:

$$Y_{itd} = \alpha + \beta_1 X_{idt=0} + \beta_2 \delta_i + \beta_3 \tau_t + \beta_4 R_d + \beta_5(\tau_t * R_d) + \beta_6(\tau_t * R_d * Z_{itd}) + \varepsilon_{itd} \qquad (1)$$

I regress the left-hand-side variables (e.g. nutrition indicators) of children i, in region d, at a time t on a vector of household demographics and baseline socioeconomic characteristics at the household and village level (X), fixed village and year effects (δ and τ), a dummy variable for treatment status (R, 1 if household in "west" village and 0 otherwise), the interaction term between τ and R and the interaction term between these two variables and Z, a sub-set of the vector X with characteristics of the children (e.g. age, sex) and the structure of the household (e.g. single-headed households). The main parameter of interest is given by β_5, which measures the treated-specific variation between 1992 and 1996, while β_6 captures differential effects for some sub-groups of children or families. Likewise, this general specification can be easily modified to accommodate the cluster-border distance identification strategy adopted in some of the models. All standard errors are clustered at the village-year level.

Table 3 presents reduced form estimates from the first set of models that exploits variation in the treatment dose within Kagera broken down by treatment indicator. I initially focus on the findings obtained from the binary definition of treatment intensity. If the arrival of refugees had an impact and their location in western Tanzania was exogenous, the identification strategies employed here should be able to capture it even in the simple D-D analysis. As shown in the table, there is a relative worsening in all the outcomes of high-intensity treated areas, although they are imprecisely estimated, except for the Z-score of weight-for-height (*whz*). However, despite the small size of the sample, the estimated negative effects are more precise once covariates that are thought to explain some of the remaining variance are included. The findings show a large and negative effect on the nutritional status of children from host communities that were more heavily flooded by refugees: a reduction of nearly 0.6 standard deviations in both the *whz* and the Z-score of height-for-age (*haz*). Figure 5 shows *whz* and *haz* kernel densities to depict these changes in more detail. While baseline nutritional distributions of children aged 0-4 were very similar between western and eastern villages, there is a relative worsening (shift to the left) for children in highly-treated towns 1.5 years after the refugees settled. With regard to the indicators of morbidity and infant mortality, they exhibit a parallel change during this period. Local children in regions with high refugee intensity were more likely to have diarrhea (18 percentage points) and fever (27 percentage points), and under-five mortality rate increased by about 10 percentage points.

Using the same sample, I turn to estimates from the preferred specifications that rely on the cluster-border distance as a source of exogenous variation in refugee intensity. Obviously, the sign of estimates using this continuous definition of the treatment have the opposite interpretation as that of the models that employ a binary treatment indicator.[9] Overall, the coefficients from the distance estimators are consistent with the direction of the impacts described above but show a higher level of precision –perhaps due to the fact the continuous indicator is better able to detect the cross-sectional variation in refugee intensity if the shock was relatively more severe in a few towns closer to the border. To get an idea of the magnitude of the estimates on the anthropometric outcomes, consider for instance the effect of moving a family 10 km father away from the border of regions in conflict (i.e. a reduction in refugee intensity). As a result of this geographic repositioning, domestic

---

[9] For instance, when looking at the impact of the shock on anthropometric measures, a positive coefficient on the village-border distance variable means that the nutritional status of children in towns farther away from the boundaries with Rwanda appear to be less affected by the influx of refugees.

children would have experienced an improvement of 0.072 to 0.094 standard deviations in their *whz* and *haz* indicators. An identical increase in the cluster-border distance would reduce the incidence rate of infectious diseases (by nearly 2.4 and 3.7 percentage points for diarrhea and fever, respectively) and under-five infant mortality by 1.6 percentage points.

I run the same models on data that contains households in villages from control areas to investigate the average treatment effect of refugee arrival in the Kagera region. The estimates of these models are summarized in Table 4 and on the whole confirm the negative impacts on children's well-being discussed in the empirical analysis above. Both simple and extended D-D regressions using the binary definition of the experimental groups show a worsening in the nutritional distribution of local children measured by the *whz* and *haz* (0.3 standard deviations), an increase in the presence of diarrhea (around 15 percentage points) and fever (close to 18 percentage points) as well as a higher rate of mortality among children aged 0-5 (between 5 and 7 percentage points). As shown in the same table, these findings are also consistent with those obtained from the cluster-border distance strategy, namely that the health status of domestic children in villages closer to highly densed refugee areas was heavily affected by the crisis. Overall, gender and other socio-demographic variables does not appear to be correlated with the magnitude of the impacts within the affected villages.

4.3 Long Run Impacts

The short run effects documented above could be argued to have impacts on later health and human capital accumulation as well. In order to trace out these potential effects, I exploit the same refugee episode together with rich data spanning a period of 10 years that include pre- and post-shock information. For this, the geographic position of the village of residency and the year of birth determine individual's exposure to refugee crisis. Therefore, this strategy exploits cross-sectional intra-cohort variation in exposure to the shock to estimate the causal effect of exposure on the outcomes of individuals who were age 0 to 5 in 1994 (10 to 15 in 2004).[10] This is identical to the D-D implemented above with the exception that now location and year of birth determine treatment

[10] The KHDS 2004 (fifth round) was designed to reinterview all individuals who were members of households interviewed in any of the preceding rounds collected between 1991 and 1994. However, due to movements in and out of the household and the split of some original households, a number of people surveyed in 2004 were not panel respondents. All the models in this section are run on two subsamples: panel respondents and a pool of time-series- cross-sectional respondents.

status. By definition, this D-D strategy will be robust to the occurrence of changes during the period 1994-2004 that may influence the outcomes of interest as long as these changes over time are not correlated with the geographic placement of the towns (i.e. refugee intensity).[11]

In an alternative design I attempt to control for potential intra-regionally common unmeasured sources of relevant variation over the period 1994-2004 by combining the intra-cohort variation discussed above with inter-cohort variation in the vulnerability/responsiveness of the stock of human capital to the shock. The older an individual was at the onset of the refugee crisis, the less sensitive her/his schooling, anthropometrics and other proxies of health were to the arrival of forced migrants.[12] Hence, the group of people 15-20 years old at the pre-shock time (1994) for whom school attainment and nutritional status were well defined provides additional control groups for treated and control regions. The general specification of the base regression used for this D-D-D strategy is given by:

$$Y_{idjt} = \alpha + \beta_1 X_{idjt} + \beta_2 \delta_k + \beta_3 \lambda_i + \beta_4 \tau_t + \beta_5 R_d + \beta_6 (\lambda_i * \tau_t) + \beta_7 (\tau_t * R_d)$$

$$+ \beta_8 (\lambda_i * R_d) + \beta_9 (\tau_t * \lambda_i * R_d) + \varepsilon_{itd} \quad (2)$$

In this equation I run the outcome of child i, in region d, of birth year j, at a time t (1 if after the shock, 0 if before) on a vector X of observable household, child level characteristics and yearly dummies to indentify idiosyncratic shocks (listed in the notes to the tables), fixed village, cohort and year effects (δ, λ, τ), and δ is a dummy for treatment group (1 if "Western", 0 if "Eastern"). The fixed effects parameters in the specification control for time-invariant characteristics of villages (β_2), most responsive cohorts (β_3), treatment villages (β_5), and the time-series changes in health and human capital outcomes (β_4). The second-level interactions control for changes over time in the responsive cohorts (β_6), changes over time for the whole treated region (β_7) and the time-invariant characteristics of treatment group in the experimental cohorts (β_8). Finally, the parameter (β_9) captures the variation in outcome Y that is specific to the responsive cohorts (relative to the less

---

[11] To the best of my knowledge, there is no strong evidence to suspect that during this period positive or negative shocks were systematically located in this way.

[12] In addition, some empirical evidence has shown that child development is less sensitive to negative health shocks for children above 5 years old and the effects, if any, may be reversible. See Behrman, Alderman and Hoddinott (2004)

sensitive cohorts) in treatment villages (relative to control villages) and in 2004 (relative to 1994). This framework is also adjusted to implement the specification that includes the distance to the border as the continuous definition of treatment. Again, disturbances for all models are allowed to be correlated for households within the same village and year.

Table 5 presents reduced-form estimates of the interaction of interest from the first set of regressions (D-D models) on stock variables that are used as proxies of health and human capital accumulation: height,[13] chronic morbidity (health problems for more than six months), school attainment and literacy. Analogous to most results discussed before, econometric models using the continuous definition of refugee intensity yield less variable point estimates as compared to those obtained from the binary treatment indicator. Overall, the coefficients seem to indicate that the negative impacts persisted over time for the affected cohorts in high intensity treated villages. For instance, for children age 0 to 5 at the arrival of refugees, exposure to the shock reduced height in early adulthood by 1.8 cm (1.2%), schooling by 0.2 years (7.1%) and literacy by 7 percentage points (8.6%). However, I do not find differential patterns of chronic illness prevalence between the two groups. The results of the D-D-D design (using the cohorts born long before the genocides as additional control groups) are shown in Table 6. While the coefficients are less precisely estimated (perhaps due to the inclusion of an additional difference), most parameters replicate the direction of the effects reported for the D-D strategy –and some even replicate the magnitude. Overall, the results of both D-D and D-D-D appear robust to the inclusion of individual-fixed effects and are empirical evidence that the massive arrival of refugees had negative long run effects on the process of human capital accumulation of affected cohorts of young children.

4.4 Robustness Analysis

The strategy adopted in this paper relies on the identifying assumption that, in the absence of the shock, the outcomes of treatment and control towns would have followed similar trends. In that case, the findings of this paper would be biased estimates of the true effect if the indicators of

---

[13] Other studies in the literature have also used height as a proxy for the stock of family investments in children in early childhood such as nutrition and health care (Steckel, 1986; Fogel, Engerman and Trussell, 1982; Fogel, 1994, Fogel and Costa, 1997; Smith, 1999).

nutrition, health and infant mortality of children were diverging in treated villages (or in towns closer to the border) with respect to control areas even before the genocides began. I investigate the possibility of this source of bias directly by using pre-shock information from the experimental regions collected by the KHDS project between 1991 and late 1993. The fact that treated regions were immune to the effects of the shock before Rwandans and Burundian fled to neighboring states implies that there should be no "treatment effect" during this period, i.e. the treatment indicator should not be statistically different from zero. Since the short run analysis used 1992 data to construct the baseline (i.e. 2 years before the large influx of refugees), I conduct a placebo test for the period 1992-1994 to rule out any differential changes along this interval of time. Although the KHDS is not the dataset used to estimate the short run impacts of the shock, I construct the same outcome variables. The results of these models are summarized in Table 7 and strengthen the consistency of the findings by revealing no evidence of confounding trends before the violent ethnic clashes occurred.

In order to test for any misspecification bias arising from the linearity assumed in the econometric models, I reestimate all the short term parameters using non-parametric methods. More specifically, I employ D-D propensity score using several kernel matching methods under the same assumption of the linear D-D, namely that the placement of the refugees among villages was orthogonal to any time-varying unmeasured or unobservable characteristics of treated and control areas. The estimates discussed on the previous sections have been obtained from small samples and, thus, non-parametric methods may lack power as compared with the OLS approach. To avoid the lost of statistical power, I use a larger sample from the DHS-1992 and DHS-1996 that includes as potential controls all rural children in Tanzania except for those in the Kigoma region which was also affected by the flood of refugees. Children in affected areas are matched to these children based on different specifications of a treatment status logistic regression that includes several pre-shock village and household covariates. The results for these exercises are presented in Table 8. Overall, the findings hold as all the quantitative effects are replicated with very high level of statistical significance. In addition, the magnitude of most impacts for "high refugee intensity" areas as well as for the average level effect remains very similar.

Perhaps the largest identification concern for this study stems from the fact that population displacement and household dissolution could be a response to the arrival of refugees, particularly among households of treated villages. And even more worrisome, if migratory behavior and household dissolution induced by the shock were not randomly determined within experimental groups leading to compositional changes that may bias the results. For instance, as migration is often costly, it may be the case that relatively richer households in areas more heavily hit by the shock were more likely to migrate to regions less or not affected by the inflows of refugees. On the other hand, poorer families with less physical assets might also be more mobile.

There is no evidence in any of the samples to confirm that such source of bias exists. With the data available, I conduct a series of simple quantitative exercises to shed some light on these issues. First, I examine the time of residency of households in the Kagera region with children under five in the 1996 Tanzania DHS survey –the after-shock sample used for the short-term analysis. Migration appears to be very low in this sample, with 806 out of 837 (96.3%) children under five belonging to households that had lived in the same village for more than three years, i.e. before the genocides. Even though temporary migration of some adult household members is a very common practice in rural Tanzania and other rural areas of East Africa, the small number of migrants in the sample is consistent with the idea that rural families have strong ties to their lands and are not very mobile as a whole. Furthermore, there are not important statistical differences between migrant households of treated and control villages, although the small number of observations (31 in total) does not allow reaching any meaningful conclusions in this respect. Finally, since migration in Kagera is mostly intra-regional, I run regression models of the migrant status of children's households in 1996 based on some time-invariant household level socio-demographic variables (i.e. proxies of pre-shock conditions) and the distance to the border of the location in 1996. If families in treated towns were pushed out of their villages by the massive number of immigrants and moved farther away, there should be a positive relationship between the longitudinal east-west distance of the current site of residency in the Kagera region to the border with Rwanda and the status of migration of the household. However, for several specifications I do not find any supportive evidence for this pattern of migration (see Table 9).[14]

---

[14] In terms of identification, the very low variation in the binary dependent variable for migration status (96.3% of the observations correspond to children within non-migrant households) restricts the inference of these models.

Similarly, the attrition of households is surprisingly low for the samples of the KHDS-91-04 panel employed in the long term analysis. In fact, 96.5% of the 788 households (760) with at least one child under five that were interviewed in the fourth round (late 1993-early 1994) could be recontacted ten years later in the final survey.[15] As expected for any longitudinal dataset, recontact rates were much lower at the individual level. In this case, 529 out of 774 (68.4%) children 0-5 years old in the pre-refugee-crisis sample were re-interviewed in the last round of the panel in 2004. I compare the sample of untracked children by treatment status on the wide range of pre-shock covariates that were used in the long-term section. The results indicate that there are some statistical differences: a lower proportion of untracked children in treated villages had families that were headed by a women and a larger fraction lived in households that had less access to safe water or electricity.[16] However, these differences are the result of some pre-existing characteristics of the regions as they mirror to a large extent some baseline differences (which I controlled for) between the panel respondents of each experimental group as well. Moreover, regression models on this sample (presented in Table 9) based on the distance to the border together with a set of pre-shock controls do not offer any evidence to support an association between refugee intensity and after-shock attrition. Finally, and consistent with the low levels of migration observed in the DHS sample, 84% of the panel children surveyed in the fourth and fifth rounds resided in households located in the same town throughout the entire period 1994-2004. Among these households, those living in the western districts of Bukoba and Muleba (control areas) were more likely to migrate, probably to Bukoba Town (the main urban center in Kagera), Dar es Salaam (the most important city in Tanzania) and other regions of the country (see Table 9). Again, there are not large differences between migrant and long-term resident household within and across regions.

There are also other sorts of issues that may possibly contaminate the internal validity of this paper. Selection bias for survival could be one of them if children who survived and grew up in refugee-stricken areas after the shock were healthier than their counterparts in control villages.[17] On one hand, if innate healthiness is modeled as a time-invariant input of the health production

---

[15] A recontacted household means that at least one member of the original household surveyed in 1994was re-interviewed in the KHDS-2004.

[16] In interest of space, these results are not included here but are available from the author upon request.

[17] In fact, Meng and Qian (2006) have found that selection bias for survival leads to an underestimation of the adverse long run impact of China's great famine on adult survivors.

function (e.g. parental genetic endowments transmitted to children), the fixed-effects models used to estimate the long run impact would get rid of this component. However, to further check this potential disparity in natural healthiness, I compare parental height distributions (a proxy of the genetic information to be inherited by their children) separately by gender between treatment and control villages and do not find evidence of this selection. That is, parents of those children observed in 2004 were not, on average, taller in treated regions (results not shown). On the other hand, if healthiness is assumed to follow a progressive accumulating process during childhood and this process affected outcomes later in life differentially between experimental groups, the estimates presented here would underestimate the real effect of such large influx of temporary migrants.[18]

Finally, two other sources of bias that may arise, especially for the long term analysis, are measurement error and the endogenous responses of national and international organizations to assist high-treatment areas. In regard to the first issue, the intra- and inter-cohort design makes it difficult to determine the real level of exposure (i.e. responsiveness) to the shock for individuals in each cohort. Besides, the distance from the town to the border with Rwanda may be a noisy measure of refugee intensity. Either one or the other, the intensity of treatment may be measured with error and the parameters would be biased towards zero (Meng and Qian, 2006). As for the second concern, if post-genocides programs of international humanitarian assistance and national public investments were targeted towards areas in western Kagera after 1994, the association between treatment status and the stock of human capital in early adulthood may be confounded by the effects of the subsequent interventions. In either case (measurement error and/or endogenous assistance), the results presented here would be still informative in a conservative scenario as they can be seen as lower bound estimates of the true impacts.

5. Interpretation: Why Negative Effects?

The influx of forced migrants into the Kagera region accounted for as much as one-third of the local inhabitants and nearly 80% of the total population in the two main recipient provinces within

---

[18] This type of selection could also be partially offset by changes in the allocation of resources within household. For instance, parents may invest relatively more resources in less healthy children in order to balance the likelihood of survival of all their children.

the Kagera region. Despite any potential benefits this enormous population movement may have brought to these regions, the main finding of this paper suggests that, at least in the case of young children, they were largely outweighed by its negative impacts. Yet, one caveat of the identification strategy adopted in this study is that it cannot disentangle the precise mechanisms underlying the observed deterioration in the outcomes of children in treated villages with respect to children in other villages of neighbor regions. Notwithstanding this shortcoming, in this section I briefly postulate some routes through which these events may have influenced negatively the well-being of young children.

The major channel seems to be the threats posed to the overall health of the most affected communities. The main prominent health problems in the Kagera region before the shock were malaria, AIDS/HIV, diarrhea, intestinal parasites and dysentery. The mass migration of refugees could have eased the emergence and spread of existing infectious diseases by food, water and environmental contaminating factors. These effects perhaps were further reinforced through the lack of proper sanitation and the collapse of already scarce health facilities. In addition, the arrival of a large group of people can easily multiply the transmission of malaria in villages with already high rates of infection and poor health prevention. In fact, malaria is caused by a parasite carried by the *Anopheles* mosquito, which feeds on human blood. The parasite is spread to people by the bite of an infected mosquito and travels through a person's bloodstream to the liver, where it grows for around two weeks until it starts attacking red blood cells and producing the associated life-threatening symptoms. The key feature of this process is that when a non-infected mosquito sucks up parasites from an infected person the mosquito becomes infected itself. Hence, increasing the amount of vulnerable people in the dimension of the shock analyzed here could have not only extended the cycle of infection but also increased the speed of transmission and the likelihood of infection. And among all individuals, children have the greater chance of being infected.

Exposure to the compounded effects of this unfavorable health environment could adversely affect the health status of parents as well as the amount and quality of other forms of investment in children such as health care, calorie consumption and parents' time spent with children. Overall, these mechanisms taken together are broadly in line with an intensification of the risks of developing diseases and harming early childhood development, even for those in utero if the health

condition of mothers was severely hit by the shock. Poor health at young ages can in turn disrupt human capital accumulation and future labor outcomes (e.g. wages, labor supply) by decreasing the returns to schooling. Even in rural communities with perhaps low returns to education, poor childhood health can reduce middle age and adult labor capacity for intensive livestock and agricultural occupations, by far the most prominent activities in the Kagera region.

An alternative hypothesis is overpopulation (seen as the ratio of the number of people to available resources rather than density) and its implicit increase in the competition for resources such as labor, food, land and wood. Unfortunately, there is not a formal accounting of the amount of food received by the villages in the area under study, but some anecdotal evidence seems to indicate that in general an important fraction of children (both migrants and locals) in refugee-hosting zones often experience severe starvation and malnutrition. This has been the case for refugee camps and nearby villages in Guinea which hosted people who escaped the genocide in Sierra Leone or in some of the refugee-recipient regions in Uganda, Chad, Kenya and Ethiopia that received migrants from the late Sudanese civil war in Darfur. More recently also, the World Food Program (WFP, 2006) have reported that the strain that Iraqi refugees is putting on Syrian neighboring regions have increased malnutrition rates for children under 5 among domestic and refugee families.

There are also other ways (perhaps less obvious) in which overpopulation could have had negative impacts on children's well-being. One case is the over-explotation of natural resources. In particular, the increased demand for wood gives a good illustration of this competition. It was reported by the UNHCR (2005) that refugees in western Kagera chopped thousands of threes for personal and commercial purposes. Consequently, children in areas around the camps were seen traveling longer distances in search of wood for fuel and shelter. Thus, not surprisingly, they appear to have carried a heavy burden of the refugee crisis in other domains as well as the findings of this paper seem to suggest.

6. Conclusions

This paper joins other efforts from an already growing literature that has been examining the contemporaneous and long run impacts of health shocks suffered in early stages of life. The

evidence offered by this study documents an indirect and long-lasting effect of civil wars on the health status and human capital of young children in refugee-hosting communities, and very likely their future economic growth. I have presented a set of results that appear to be robust to different subsamples and measures of treatment as well as to alternative specifications and estimation methods. But beyond any issues of internal validity, the difficulty for most papers lies in the generalization of their results. Despite the uniqueness of its identification strategy (i.e. timing of events, geographic characteristics, location of refugee settlements), the findings of this paper can be extrapolated to a large number of similar refugee movements in the developing world to shed some light on the possible impacts on the populations that receive them. In the last four decades, Sub-Saharan Africa has registered no less than 15 civil conflicts in countries like Angola, Burundi, Congo, Guinea-Bissau, Ivory Coast, Liberia, Mozambique, Nigeria, Rwanda, Sierra Leone, Somalia, Sudan and Uganda. Such high prevalence of war in Africa translated into an eightfold increase in the number of refugees between late 1960's and mid 1990's according to the UNHCR. These figures do not even include the amount of internally displaced persons or the thousands of migrants fleeing political oppression, as seems to be the case of more than 3 million poor Zimbabweans recently reported in the news to have left their country mostly for rural villages in Botswana, Mozambique and South Africa. Overall, the socioeconomic characteristics of all these refugees and their hosting communities resemble the conditions of vulnerability observed for the refugee crisis in the Kagera region: high levels of poverty, overpopulation, with most people being subsistence farmers and pastoralists that deal with erratic rains and face high levels of disease, malnutrition and death.

The empirical findings also show that the impacts of childhood exposure at the onset of the refugee migration persisted over time, at least until adolescence. There is considerable evidence from different developing countries that supports this observation by indicating that health shocks and growth lost in early childhood are not only partially offset later in life but correlated with cognitive development deficits and attained body size in adulthood (Martorell, et al. 1994; Simondon, 1998; Grantham-McGregor, et al, 1999a, 1999b; Hoddinott and Kinsey, 2001; Alderman, et al., 2003). Lastly, the detrimental effects of the shock among affected children in northwestern Tanzania may have other long-term consequences on future productivity and labor market outcomes as well. A growing body of evidence in the literature documents the positive association between height in adulthood and earnings (Behrman and Deolalikar, 1989; Behrman, 1993; Foster

and Rosenzweig, 1994; Strauss and Thomas, 1997 & 1998; Bleakely, 2002). For instance, using the point estimates of Strauss and Thomas (1997) on the earnings-height elasticity for Brazil, I predict that the reduced height observed among the Tanzanian children could account for as much as a 2.8% fall in their earnings in adulthood. This is, however, a very conservative figure as it does not include the additional negative impact on earnings arising from the possible drop in school attainment documented above.

Finally, from a policy perspective, the results of this paper together with those of previous research offer evidence to promote and mobilize global assistance in the scale required. Unfortunately –but understandably to some extent– some low-income countries overburdened by hosting a large number of displaced people have already began imposing restrictions on the acceptance of more refugees. Closing the borders, however, is far from the solution. Instead, as the findings seem to suggest, what is needed is a more comprehensive response from international donors and aid agencies –assuming obviously that this assistance in fact works— to protect the welfare of both refugees and host communities against current and future civil wars and the proximity of civilians to these events.

References

Alderman, H., J. Hoddinott and B. Kinsey, "Long Term Consequences of Early Childhood Malnutrition", World Bank: Washington, DC, 2003, Processed.

Alix-Garcia, Jennifer, "The Effect of Refugee Inflows on Host Country Populations: Evidence from Tanzania", working paper, Department of Economics, University of Montana, April, 2007.

Almond, Douglas and Mazumder Bhashkar, "The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data", *American Economic Review*, 2005, 95 (2), pp. 258-62.

Behrman, Jere, "The Economic Rationale for Investing in Nutrition in Developing Countries", *World Development*, 1993, 21 (11), pp. 1749-72.

Behrman, Jere and A. Deolalikar, Wages and Labor Supply in Rural India: The Role of Health, Nutrition and Seasonality. In D. Sahn, ed., *Causes and Implications of Seasonal Variability in Household Food Security*, 1989, The Johns Hopkins University Press, Baltimore, MD, pp. 107-118.

Behrman, Jere and Mark Rosenzweig, "The Returns to Birth Weight", *Review of Economics and Statistics*, 2005, 86, pp. 586-601.

Berlagea, Lode and Marijke Verpoortena, "Economic Mobility in Rural Rwanda: A Study of the Effects of War and Genocide at the Household Level", *Journal of African Economies*, Jan. 2007, 16(3), pp. 349-392.

Bevelander, Pieter and Christer Lundh, "Employment Integration of Refugees: The Influence of Local Factors on Refugee Job Opportunities in Sweden", IZA Discussion Paper No. 2551, January, 2007.

Bleakley, Hoyt, "Disease and Development: Evidence from Hookworm Eradication in the American South", *Quarterly Journal of Economics*, Feb. 2007, 122 (1), pp. 73-117.

Bolesta, Andrezej, "Socio-Economic Conditions of Palestinian Refugees in Jordan, Syria, and Lebanon", TIGER Working Paper Series No. 35, December, 2002.

Borjas, George, "Immigrants, Minorities, and Labor Market Competition", *Industrial and Labor Relations Review*, Apr. 1987, 40 (3), pp. 382-392.

Borjas, G. Freeman R. and L. Katz, "How Much Do Immigration and Trade Affect Labor Market Outcomes?", Brookings Papers on Economic Activity, 1997, Vol. 1997 (1), pp. 1-90.

Card, David, "The Impact of the Mariel Boatlift on the Miami Labor Market", *Industrial and Labor Relations Review*, Vol. 43 (2), Jan. 1990, pp. 245-257.

Card, David and Joseph Altonji, "The Effects of Immigration on the Labor Market Outcomes of Less-Skilled Natives," in John Abowd and Richard B. Freeman, editors., *Immigration, Trade and Labor*. Chicago: University of Chicago Press, 1991.

Chiswick, Carmel, "The Impact of Immigration on the Human Capital of Natives", *Journal of Labor Economics*, Oct. 1989, 7 (4), pp. 464-486.

Collier, Paul and Anke Hoeffler, "On Economic Causes of Civil War", *Economic Oxford Papers*, Oct. 1998, 50 (4), pp. 563-573.

Engel, Stefanie and Ana Ibanez, "Displacement Due to Violence in Colombia: A Household-Level Analysis", *Economic Development and Cultural Change*, 55 (2), Jan. 2007, pp. 335-65.

Gleditsch, Kristian and Idean Salehyan, "Refugees and the Spread of Civil War", *International Organization*, 60 (2), Spring 2006, pp. 335-66

Grantham-McGregor, S, L. Fernald and K. Sethuraman, "Effects of Health and Nutrition on Cognitive and Behavioral Development in Children in the First Three Years of Life", Low Birth weight, Breastfeeding and Protein-energy Malnutrition, Food and Nutrition Bulletin, 1999a, 20(1), pp.53-75.

Grantham-McGregor, S, L. Fernald and K. Sethuraman, "Effects of Health and Nutrition on Cognitive and Behavioral Development in Children in the First Three Years of Life", Infections and Micronutrient Deficiencies: Iodine, Iron and Zinc. Food and Nutrition Bulletin, 1999b, 20 (1), pp. 76-99

Fairlie, Robert and Bruce Meyer, "The Effect of Immigration on Native Self-Employment", *Journal of Labor Economics*, Jul. 2003, 21 (3), pp. 619-650.

Fogel, Robert., Engerman, Stanley and James Trussell, "Exploring the Uses of Data on Height", *Social Science History*, Trends in Nutrition, Labor, Welfare, and Labor Productivity, 1982, 6 (4), pp. 401-421.

Fogel, Robert, "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy", *American Economic Review*, Jun. 1994, 84 (3), pp. 369-395.

Fogel, R. and Dora Costa, "A Theory of Technophysio Evolution, With Some Implications for Forecasting Population, Health Care Costs, and Pension Costs", *Demography*, Feb. 1997, 34 (1), pp. 49-66.

Foster, A. and M. Rosenzweig, "A test for moral hazard in the labor market: Effort, health and calorie consumption", *Review of Economics and Statistics*, 1994, 76(2), pp. 213-227.

Friedberg, Rachel, "The Impact of Mass Migration on the Israeli Labor Market", *Quarterly Journal of Economics*, Nov. 2001, 116 (4), pp. 1373-1408.

Hansen, Jorgen and Magnus Lofstrom, "Immigrant Assimilation and Welfare Participation: Do Immigrants Assimilate into or out of Welfare?", *Journal of Human Resources*, 2003, 38 (1), pp. 74-98.

Hatton, Timothy and Jeffrey G. Williamson, "Refugees, Asylum Seekers and Policy Europe", NBER working paper 10680, August, 2004.

Hatton, Timothy and Jeffrey G. Williamson, "What Fundamentals Drive World Migration?, NBER working paper 9159, September 2002.

Hoddinott, J., and B. Kinsey, "Child Growth in the Time of Drought", *Oxford Bulletin of Economics and Statistics*, 2001, 63(4), pp. 409-436.

LaLonde, R.J. & Topel, R.H, "Immigrants in the American Labor Market: Quality, Assimilation, and Distributional Effects", *American Economic Review*, 81, pp. 297–301.

Martin, Adrian, "Environmental Conflict between Refugee and Host Communities", *Journal of Peace Research*, 42(3), May 2005, pp. 329-46.

Martorell, R., K.L. Khan and D.G. Schroeder, "Reversibility of Stunting: Epidemiological Findings in Children from Developing Countries", *European Journal of Clinical Nutrition*, 1994, 48(suppl), S45-S57.

Meng, Xin and Nancy Qian, "The Long Run Health and Economic Consequences of Famine on Survivors: Evidence from China's Great Famine", mimeo, November, 2006.

Miguel, Edward and Gérard Roland, "The Long Run Impact of Bombing Vietnam", working paper, Nov. 2006.

Montalvo, J. and Marta Reynal-Querol, "Fighting against Malaria: Prevent Wars while Waiting for the "Miraculous" Vaccine", *Review of Economics and Statistics*, Feb. 2007, 89 (1), pp. 165-177.

National Bureau of Statistics of Tanzania, "National Accounts of Tanzania, 1998-2001", 2001, Dar es Salaam, Tanzania.

Pischke, J.-S. and J. Velling , "Wage and Employment Effects of Immigration to Germany: An Anlysis Based on Local Labor Markets," CEPR-Discussion-Paper Nr. 935, 1994.

Simondon, K., F. Simondon, I. Simon, A. Diallo, E. Benefice, P. Traissac, and B. Maire, B., "Preschool stunting, age at menarche and adolescent height: A longitudinal study in rural Senegal", *European Journal of Clinical Nutrition*, 1998, 52, pp. 412-418.

Smith, James, "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status", *Journal of Economic Perspectives*, 1999, 13 (2), pp. 145-166.

Steckel, Richard, "Peculiar Population: The Nutrition, Health, and Mortality of American Slaves from Childhood to Maturity", *Journal of Economic History*, Sep. 1986, 46 (3), pp. 721-741.

Strauss, J. and Duncan Thomas, "Health and Wages: Evidence on Men and Women in Urban Brazil", *Journal of Econometrics*, 1997, 77 (1), pp. 159-187.

Strauss, J. and Duncan Thomas, "Health, Nutrition, and Economic Development", *Journal of Economic Literature*, 1998, 36 (2), pp. 766-817.

United Nations High Commission for Refugees, "Impacts of Refugees and Internally Displaced Persons on the Environment in Tanzania", Assessment Report, Disaster Management Branch and UNEP Division of Environmental Policy Implementation, 2005, Kenya.

United Nations High Commission for Refugees, "Global Report 2006", June, 2007, Geneva, Switzerland.

U.S. Committee for Refugees and Immigrants, "World Refugee Survey 2006", 2007, Washington DC.

Figure 1. Total Number of Refugees Hosted by Tanzania, 1991-2004
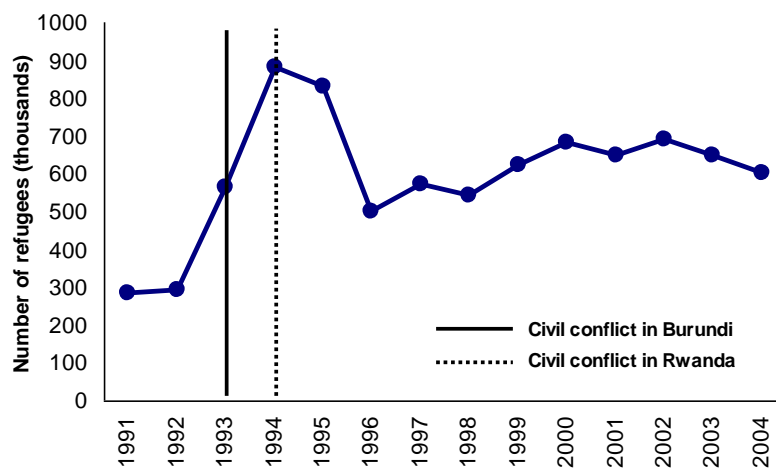
Source: United Nations High Commissioner for Refugees (UNHCR)

Table 1. Districts of Kagera Region by the Intensity of the Influx of Refugees

Indicator	High Treatment Regions		Low Treatment Regions			
	Ngara	Karagwe	Biharamulo	Muleba	Bukoba Rural	Bukoba Urban
Refugee influx intensity	83.3%	60.0%	20.0%	14.3%	5.9%	0.0%

Notes: The indicator of intensity is calculated using retrospective information from the fifth round of the Kagera Health and Development Survey (2004) and shows the percentage of communities that reported having refugees' resettlements close by in the village or the ward after 1993.

Figure 2. Maps of Kagera Region and Location of Experimental Villages

(a) Topographic Characteristics

(b) Land Use

(c) Rainfall

(d) Administrative Divisions
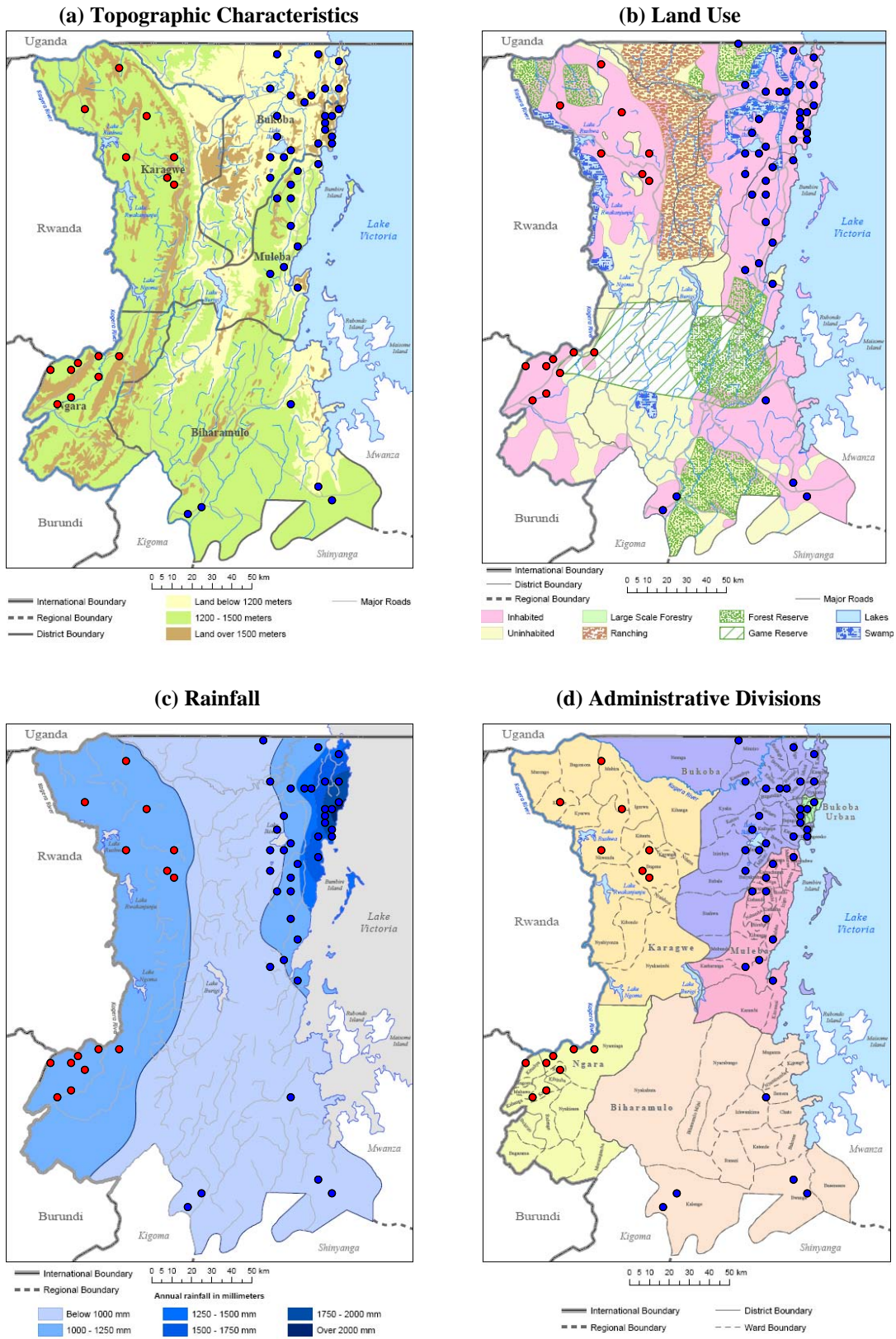
Source: Digital Cartography and GIS from Kagera Food and Security Project (KAFOSEC)

Figure 3. Relationship between the Intensity of the Influx of Refugees and the Distance to the Village to the Border with Rwanda

Note: The indicator of intensity is calculated using retrospective information from the fifth round of the Kagera Health and Development Survey (KHDS, 2004) that ranks (from 1 to 4) the distance from the village to refugees' resettlements after 1993. Distances from the villages to the Rwandan border were calculated using GIS and restricted-access geo-referenced data from the KHDS, 2004.

Figure 4. Maps of Kagera, Mwanza and Mara Regions and Location of Experimental Villages

Table 2. Pre-Shock Summary Statistics by Treatment Region

Variable	Total	High treatment areas (West=1)	Low treatment areas (East=1)	Difference	Total	Treatment areas (Kagera=1)	Control areas (Kagera=0)	Difference
Demographics								
Number of household members	6.97	6.71	7.14	-0.426	7.68	6.97	7.94	-0.972 ***
	[0.238]	[0.238]	[0.365]	[0.436]	[0.156]	[0.238]	[0.193]	[0.307]
Number of children	1.41	1.29	1.49	-0.205	1.51	1.41	1.54	-0.129
(five years and under)	[0.066]	[0.099]	[0.088]	[0.132]	[0.045]	[0.066]	[0.056]	[0.087]
Age of household head (years)	42.0	40.0	43.4	3.44 **	44.4	42.0	45.4	-3.338 ***
	[0.868]	[1.137]	[1.224]	[1.759]	[0.490]	[0.868]	[0.585]	[1.100]
Age of mother (years)	31.8	32.6	31.3	1.363	32.2	31.8	32.3	-0.447
	[0.529]	[0.863]	[0.666]	[1.090]	[0.285]	[0.529]	0.339	[0.628]
Age of children (years)	7.7	7.9	7.5	0.364	8.1	7.7	8.2	-0.525
	[0.372]	[0.591]	[0.481]	[0.762]	[0.193]	[0.372]	0.225	[0.435]
Proportion of parents married	0.809	0.857	0.780	0.077	0.807	0.809	0.806	0.003
	[0.026]	[0.036]	[0.035]	[0.051]	[0.013]	[0.026]	[0.015]	[0.030]
Household head is female	0.146	0.121	0.162	-0.041	0.146	0.146	0.147	-0.001
	[0.023]	[0.034]	[0.031]	[0.046]	[0.012]	[0.023]	[0.014]	[0.027]
Age of mother at first birth	18.6	19.0	18.4	0.638	18.3	18.6	18.2	0.419
	[0.227]	[0.391]	[0.273]	[0.477]	[0.105]	[0.227]	[0.118]	[0.256]
Socioeconomic indicators								
Access to clean water	0.257	0.220	0.278	-0.058	0.481	0.257	0.560	-0.303 ***
	[0.029]	[0.043]	[0.039]	[0.058]	[0.017]	[0.029]	[0.020]	[0.036]
Acess to prenatal care	0.787	0.770	0.800	-0.030	0.705	0.787	0.675	0.112 ***
	[0.027]	[0.044]	0.034	[0.056]	[0.015]	[0.027]	[0.018]	[0.033]
Household has electricity	0.013	0.022	0.007	0.015	0.021	0.013	0.024	-0.011
	[0.007]	[0.015]	[0.007]	[0.017]	[0.005]	[0.007]	[0.006]	[0.010]
Household has radio	0.283	0.285	0.282	0.003	0.280	0.283	0.279	0.004
	[0.030]	[0.047]	[0.039]	[0.061]	[0.015]	[0.030]	[0.018]	[0.035]
Household with hard floor	0.076	0.111	0.053	0.058	0.118	0.076	0.133	-0.057
	[0.017]	[0.033]	[0.019]	[0.038]	[0.038]	[0.017]	[0.051]	[0.054]
Family has bicycle	0.311	0.307	0.313	-0.006	0.352	0.311	0.368	-0.057
	[0.030]	[0.048]	[0.040]	[0.005]	[0.016]	[0.030]	[0.019]	[0.036]
Schooling								
Father's school attainment	5.38	5.25	5.48	-0.23	5.18	5.38	5.04	0.343
(in years)	[0.279]	[0.221]	0.171	[0.279]	[0.229]	[0.279]	[0.250]	[0.540]
Mother's school attainment	3.79	3.75	3.81	-0.07	3.78	3.79	3.78	0.009
(in years)	[0.210]	[0.352]	[0.261]	[0.438]	[0.108]	[0.210]	[0.127]	[0.245]
Labor force participation								
Mother working	0.831	0.857	0.814	0.043	0.857	0.831	0.867	-0.036
	[0.024]	[0.036]	[0.033]	[0.049]	[0.012]	[0.024]	[0.013]	[0.028]
Number of households	226	91	135		845	226	619	

Notes. Summary statistics based on the Demographic and Health Survey - Tanzania, 1992. Standard errors in square brackets. The symbols ***, (**) and [*] stand for significance at the 1%, (5%) and [10%] levels, respectively. See text for definitions of households in high and low treatment and control regions.

Table 3. Differential Short Run Local Impacts of the Influx of Refugees on Nutrition, Morbidity and Infant Mortality

	Sample: Villages in Kagera Region						
			Reduced form regressions				
	Pre-shock mean		D-D: West=1, East=0		D-D: Distance to the border		
Outcome	High treatment areas	Low treatment areas	(i)	(ii)	(i)	(ii)	N
Nutritional Indicators							
Weight-for-height z-score	-0.446	-0.490	-0.542 **	-0.596 **	0.0072 **	0.0077 **	457
	[1.107]	[1.203]	[0.222]	[0.247]	[0.0025]	[0.0034]	
Height-for-age z-score	-1.666	-1.746	-0.522	-0.660 **	0.0075 *	0.0094 **	457
	[1.356]	[1.391]	[0.307]	[0.302]	[0.0043]	[0.0041]	
Morbidity							
Children with diarrhea	0.036	0.046	0.135	0.183 *	-0.0024 *	-0.0028 **	504
	[0.186]	[0.211]	[0.090]	[0.093]	[0.0014]	[0.0012]	
Children with fever	0.250	0.267	0.217	0.270 **	-0.0037 *	-0.0040 *	504
	[0.434]	[0.443]	[0.125]	[0.100]	[0.0017]	[0.0021]	
Infant Mortality							
Death in the first five years of life	0.084	0.163	0.083	0.108 *	-0.0016 **	-0.0018 **	647
	[0.279]	[0.370]	[0.061]	[0.054]	[0.0007]	[0.0007]	
Controls?			No	Yes	No	Yes	

Notes: Robust standard errors clustered at the village-year level in square brackets. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children four years old and below from villages in Kagera region, except for infant mortality that was calculated on children who were five years old or younger. All multivariate difference-in-difference regressions contain pre-shock (1992) village, household and individual demographic controls for gender (child and head of the household), age (child and parents), parental schooling, maternal anthropometric measures, number of members within the household, prenatal care, binary variables describing the access to clean water and other physical characteristics of the household, and village and year fixed effects. See text for definitions of outcomes, experimental and non experimental individuals and regions, before and after years and design based on the distance of the village to the border with Rwanda.

Table 4. Aggregated Short Run Impacts of the Influx of Refugees on Nutrition, Morbidity and Infant Mortality

	Sample: Villages in Kagera, Mara and Mwanza regions						
			Reduced form regressions				
	Pre-shock mean		D-D: Kagera=1, Kagera=0		D-D: Distance to the border		
Outcome	Treatment areas	Control areas	(i)	(ii)	(i)	(ii)	N
Nutritional Indicators							
Weight-for-height z-score	-0.474	-0.252	-0.291 *	-0.333 *	0.0009 *	0.0011 **	1,619
	[1.165]	[1.188]	[0.170]	[0.175]	[0.0005]	[0.0005]	
Height-for-age z-score	-1.715	-1.695	-0.304 *	-0.303 *	0.0011 **	0.0011 **	1,619
	[1.376]	[1.258]	[0.183]	[0.184]	[0.0005]	[0.0005]	
Morbidity							
Children with diarrhea	0.042	0.100	0.151 ***	0.171 ***	-0.0004 ***	-0.0005 ***	1,796
	[0.202]	[0.300]	[0.044]	[0.047]	[0.0001]	[0.0001]	
Children with fever	0.260	0.373	0.179 **	0.201 ***	-0.0005 ***	-0.0006 ***	1,784
	[0.439]	[0.483]	[0.074]	[0.079]	[0.0002]	[0.0002]	
Infant Mortality							
Death in the first five years of life	0.136	0.128	0.047	0.076 *	-0.0002 ***	-0.0003 ***	2,286
	[0.343]	[0.334]	[0.039]	[0.041]	[0.0001]	[0.0001]	
Controls?			No	Yes	No	Yes	

Notes: Robust standard errors clustered at the village-year level in square brackets. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children four years old and below from villages in Kagera, Mara and Mwanza regions, except for infant mortality that was calculated on children who were five years old or younger. All multivariate difference-in-difference regressions contain pre-shock (1992) village, household and individual demographic controls for gender (child and head of the household), age (child and parents), parental schooling, maternal anthropometric measures, number of members within the household, prenatal care, binary variables describing the access to clean water and other physical characteristics of the household, and village and year fixed effects. See text for definitions of outcomes, experimental and non experimental individuals and regions, before and after years and design based on the distance of the village to the border with Rwanda.
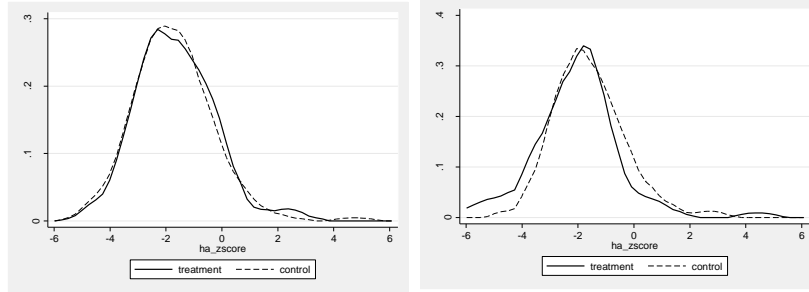
Figure 5. Short Run Distributional Changes in Measures of Nutritional Status

High and low treatment groups at t=0 and t=1 (H/A z-score Kernel Densities)

Treatment and control groups at t=0 and t=1 (H/A z-score Kernel Densities)
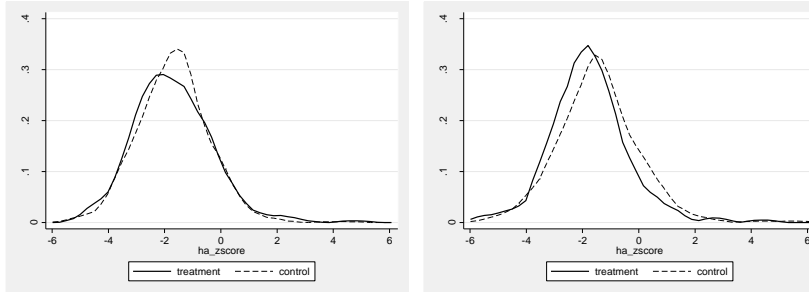
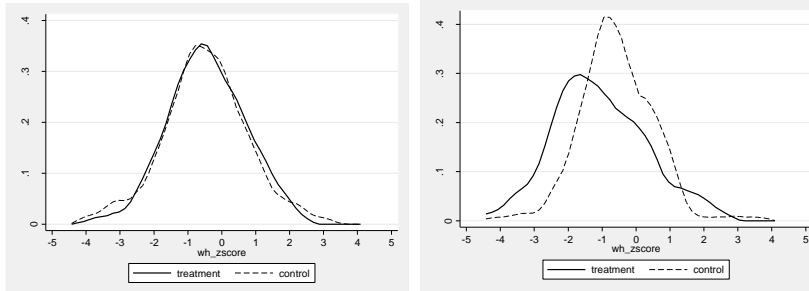**High and low treatment groups at t=0 and t=1 (W/H z-score Kernel Densities)**

Treatment and control groups at t=0 and t=1 (W/H z-score Kernel Densities)
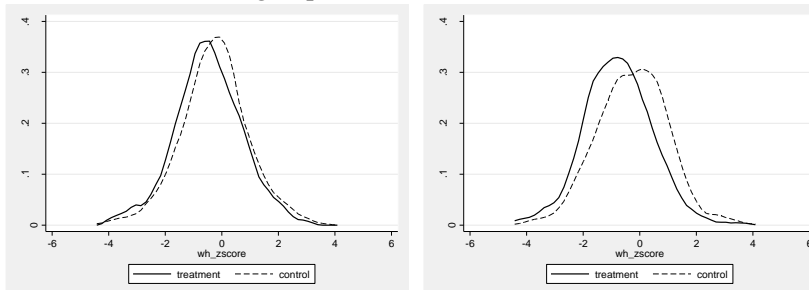
Table 5. Long Run Impacts on Health and Human Capital Accumulation Exploiting Intra-cohort Variation in Exposure to the Shock

	Pre-shock mean		Reduced form regressions							
			D-D: West=1, East=0				D-D: Distance to the border			
			Pooled		Panel		Pooled		Panel	
Outcome	West=1	East=0	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
Anthropometrics										
Height (centimeters)	84.90	86.31	-1.943	-1.719 **	-1.777 *	-2.184	0.0208	0.0196 **	0.0280 **	0.0289 **
	[1.030]	[0.639]	[2.068]	[0.680]	[1.051]	[1.135]	[0.0229]	[0.0094]	[0.0131]	[0.0142]
Observations			1,861	1,785	1,159	1,078	1,861	1,785	1159	1078
Morbidity										
Chronic disease	0.088	0.110	-0.022	-0.005	-0.066	-0.069	0.0003	0.0001	0.0006	0.0007
	[0.034]	[0.042]	[0.041]	[0.037]	[0.130]	[0.147]	[0.0005]	[0.0006]	[0.0017]	[0.0019]
Observations			1,548	1,340	855	722	1,548	1,340	855	722
Education										
School attainment (years)	---	---	-0.179	-0.172	-0.272	-0.200	0.0040 *	0.0045 ***	0.0064 ***	0.0051 **
			[0.163]	[0.120]	[0.193]	[0.204]	[0.0022]	[0.0014]	[0.0023]	[0.0025]
Observations			2,125	2,026	1,281	1,190	2,125	2,026	1281	1190
Does read/write?	---	---	-0.057	-0.068	-0.075	-0.065	0.0011	0.0013 **	0.0013 **	0.0012 **
			[0.067]	[0.048]	[0.0467]	0.051	[0.0007]	[0.0005]	[0.0005]	[0.0006]
Observations			2,127	2,028	1,284	1,193	2,127	2,028	1284	1193
Controls?			No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village-year level in square brackets. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children under five in 1994 (10-15 years old in 2004) who lived in the village since 1994 (for the pooled sample) or were tracked in the 2004 (for the panel). Chronic disease measures the presence of health problems for more than six months. Pre-shock means presented in the first two columns correspond to cross-sectional pooled samples. Depending on the outcomes, multivariate difference-in-difference regressions include controls for individual characteristics (age, sex), household socio-demographic structure (number of members, number of children, age of household head, single-headed households), parental education, dwelling conditions (type of floor and toilet), access to safe water, distance to closest school and health facility, household assets and expenditures, proportion of coffee grown in lands cultivated by the household, yearly dummies to capture household level shocks between 1994-2004, rainfall variability and time-invariant village effects (for the pooled sample). See text for definitions of outcomes, experimental and non experimental cohorts and regions, before and after years and design based on the distance of the village to the border with Rwanda.

Table 6. Long Run Impacts on Health and Human Capital Accumulation Exploiting Intra- and Inter-cohort Variation in Exposure to the Shock

				Reduced form regressions						
			D-D-D: West=1, East=0				D-D-D: Distance to the border			
	Pre-shock mean		Pooled		Panel		Pooled		Panel	
Outcome	West=1	East=0	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
Anthropometrics										
Height (centimeters)	84.90	86.31	-2.266	-1.435	-1.429	-1.878	0.0316	0.0176	0.0334	0.0337 *
	[1.030]	[0.639]	[1.939]	[1.533]	[1.493]	[1.630]	[0.0212]	[0.0166]	[0.0187]	[0.0202]
Observations			3,362	3,179	2,325	2,070	3,362	3,179	2,325	2,070
Morbidity										
Chronic disease	0.088	0.110	0.052	0.060	0.010	0.084	-0.0008	-0.0007	-0.0010	-0.0014
	[0.034]	[0.042]	[0.066]	[0.070]	[0.169]	[0.198]	[0.0008]	[0.0009]	[0.0021]	[0.0025]
Observations			2,963	2,544	1,956	1,561	2,963	2,544	1,956	1,561
Education										
School attainment (years)	---	---	-0.277	-0.083	-0.437	-0.347	0.0054	0.0039	0.0076 *	0.0074 *
			[0.454]	[0.429]	[0.379]	[0.372]	[0.0052]	[0.0050]	[0.0046]	[0.0041]
Observations			3,824	3,743	2,584	2,523	3,824	3,743	2,584	2,523
Does read/write?	---	---	-0.119	-0.111 *	-0.084	-0.070	0.0018 **	0.00195 **	0.0016 **	0.0014 *
			0.073	[0.066]	[0.060]	[0.060]	[0.0008]	[0.0008]	[0.0007]	[0.0007]
Observations			3,914	3,688	2,639	2,564	3,914	3,688	2,639	2,564
Controls?			No	Yes	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors clustered at the village-year level in square brackets. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children under five in 1994 (10-15 years old in 2004) who lived in the village since 1994 (for the pooled sample) or were also tracked in the 2004 round (for the panel). Chronic disease measures the presence of. Pre-shock means presented in the first two columns correspond to cross-sectional pooled samples. Depending on the outcomes, multivariate difference-in-difference regressions include controls for individual characteristics (age, sex), household socio-demographic structure (number of members, number of children, age of household head, single-headed households), parental education, dwelling conditions (type of floor and toilet), access to safe water, distance to closest school and health facility, household assets and expenditures, proportion of coffee grown in lands cultivated by the household, yearly dummies to capture household level shocks between 1994-2004, rainfall variability and time-invariant village effects (for the pooled sample). See text for definitions of outcomes, experimental and non experimental cohorts and regions, before and after years and design based on the distance of the village to the border with Rwanda.

Table 7. Pre-shock Difference-in-Difference Models on Nutrition, Morbidity and Infant Mortality, 1992-94

	Sample: Villages in Kagera Region						
			Reduced form regressions				
	Pre-shock mean		D-D: West=1, East=0		D-D: Distance to the border		
Outcome	High treatment areas	Low treatment areas	(i)	(ii)	(i)	(ii)	N
Nutritional Indicators							
Weight-for-height z-score	-0.109	-0.270	-0.121	-0.165	0.0005	0.0019	1,067
	[0.174]	[0.069]	[0.263]	[0.173]	[0.0031]	[0.0020]	
			{0.644}	{0.342}	{0.850}	{0.329}	
Height-for-age z-score	-2.201	-1.540	0.075	0.134	0.0000	-0.0013	1,069
	[0.143]	[0.090]	[0.273]	[0.130]	[0.0033]	0.0015	
			{0.783}	{0.307}	{0.997}	{0.393}	
Morbidity							
Child sick?	0.232	0.248	0.022	0.058	0.0000	-0.0004	1,115
	[0.036]	[0.020]	[0.061]	[0.043]	[0.0007]	[0.0005]	
			{0.717}	{0.185}	{0.960}	{0.478}	
Infant Mortality							
Death in the first five years of life	0.124	0.101	0.009	0.020	0.0001	0.0000	1,740
	[0.0242]	[0.011]	[0.030]	[0.022]	[0.0003]	[0.0002]	
			{0.778}	{0.365}	{0.637}	{0.896}	
Controls?			No	Yes	No	Yes	

Notes: Robust standard errors clustered at the village-year level in square brackets. Significant at 90(*), 95(**), 99(***) percent confidence. Two-tailed p-values reported in { }. The units of observation for nutritional and morbidity outcomes are children four years old and below from villages in the Kagera region, while units of analysis for infant mortality are households with at least one child under five which reported having/not having deaths of children under five in the last 24 months. Depending on the outcomes, multivariate difference-in-difference regressions include controls for individual characteristics (age, sex), household socio-demographic structure (number of members, number of children, age of household head, single-headed households), parental education, dwelling conditions (type of floor and toilet), access to safe water, distance to closest school and health facility, household assets and expenditures, proportion of coffee grown in lands cultivated by the household, rainfall variability and time-invariant village effects. See text for definitions of outcomes, experimental and non experimental regions, before and after years and design based on the distance of the village to the border with Rwanda.

Table 8. Difference-in-Differences Matching Estimates of Short Run Impacts on Nutrition, Morbidity and Infant Mortality
(Reduced Form Estimates)

Outcome	D-D: West of Kagera=1, Rest of Tanzania=0					D-D: Kagera=1, Rest of Tanzania=0			
	Type of Kernel [a]					Type of Kernel [a]			
	NN(10) bw = 0.01	E bw = 0.01	LL bw = 0.01	N	CS	NN(10) bw = 0.01	E bw = 0.01	LL bw = 0.01	N
Nutritional Indicators									
Weight-for-height z-score	-0.704 *** [0.227]	-0.729 *** [0.220]	-0.735 *** [0.255]	4,463	99.1%	-0.332 *** [0.086]	-0.318 *** [0.083]	-0.349 *** [0.010]	4,463
Height-for-age z-score	-0.453 ** [0.250]	-0.458 ** [0.242]	-0.455 ** [0.247]	4,277	99.1%	-0.165 [0.134]	-0.060 [0.097]	-0.072 [0.075]	4,277
Morbidity									
Children with diarrhea	0.283 ** [0.130]	0.282 ** [0.132]	0.288 ** [0.142]	4,904	100%	0.178 ** [0.081]	0.173 ** [0.079]	0.171 ** [0.082]	4,904
Children with fever	0.334 *** [0.120]	0.340 *** [0.118]	0.340 *** [0.121]	4,900	100%	0.177 *** [0.054]	0.169 *** [0.053]	0.165 *** [0.053]	4,900
Infant Mortality									
Death in the first five years of life	0.076 * [0.050]	0.076 ** [0.046]	0.075 ** [0.048]	5,139	100%	0.042 ** [0.028]	0.049 ** [0.029]	0.050 ** [0.028]	5,139

Notes: Bootstrapped standard errors presented in square brackets were obtained from 500 replications. Significant at 90(*), 95(**), 99(***) percent confidence. The units of observation are children born in the period 0-48 months preceding the survey in rural villages in Tanzania, except for infant mortality that was calculated on children who were five years old or younger preceding the survey. Units matched on the propensity score from a logistic regression on presence in the treatment region. Pre-shock covariates in the logistic regression include child's age and sex, mother's age (mother's body mass index for nutritional indicators), parental education, a binary for female-headed households, number of children under five, dummies for dwelling characteristics (earth floor, electricity), access to clean water, dummies for household assets (radio, vehicle) and prenatal care usage. Results presented in this table are very similar to those obtained with bandwidths ranging from 0.005 to 0.025. See text for definitions of outcomes, experimental and non experimental individuals and regions, before and after years

[a] NN(10) = 10 nearest neighbors; E = Epanechnikov; LL = local linear; bw = bandwith; CS = common support.

Table 9. Effects of Refugee Intensity on Household Migration and Attrition

Dependent variable	Sample: Tanzania DHS 1996		
	LPM	Probit	N
Migratory status after the shock	0.0002	0.0001	837
(=1 if child in household that moved, 0 otherwise)	[0.0005]	[0.0001]	
	{ 0.624}	{ 0.570}	

Dependent variable	Sample: Tanzania KHDS 1994		
	LPM	Probit	N
Attrition status after the shock	0.0005	0.0004	774
(=1 if child in untracked household, 0 otherwise)	[0.0006]	[0.0007]	
	{0.466}	{0.620}	
Migratory status after the shock	0.0001 *	0.0012 **	529
(=1 if child in household that moved, 0 otherwise)	[0.0006]	[0.0006]	
	{0.096}	[0.037]	
Controls?	Yes	Yes	

Notes: Robust standard errors clustered at the village-year level in square brackets. Probit coefficients reported correspond to marginal effects. Significant at 90(*), 95(**), 99(***) percent confidence. Two-tailed p-values reported in { }. Independent variable of interest is the distance of the town of residency to the border with Rwanda. Model using the DHS-1996 sample contain controls for gender (child and head of the household), age (child and parents), number of members within the household and parental schooling. Models using the KHDS-1994. include controls for individual characteristics (age, sex), household socio-demographic structure (number of members, number of children, age of household head, single-headed households), parental education, dwelling conditions (type of floor and toilet), access to safe water, distance to closest school and health facility, household assets and expenditures, proportion of coffee grown in lands cultivated by the household, rainfall variability and time-invariant village effects.