WINNERS AND LOSERS AMONG A REFUGEE-HOSTING POPULATION:
Consumption, economic activities and agglomeration

Jean-François Maystadt* Philip Verwimp†

Draft and incomplete: Please do not cite without permission
(10 February 2009)

Abstract

Every year, thousands of refugees are forced to leave their countries of origin and are hosted by their neighboring countries. However, very few is known about the impact of these refugees on the local economy and its inhabitants. Based on hypothesis formulated during a two-month iterative field research, a theoretical framework is used to understand how the refugee inflow would affect the good and labour markets of the local economy. We then test the theoretical predictions regarding the potential winners and losers among the refugee-hosting population, using household panel data collected in the region of Kagera in Tanzania. Preliminary results suggest that the local hosts do not necessarily suffer from the refugee presence. Net economic benefits could even emerge provided a sufficient mass of refugees is gathered. Furthermore, the economic benefits appear to be unevenly distributed among the refugee-hosting population. The agricultural worker is likely to suffer the most from an increase in competition on the labor markets and the surging prices of several goods. On the contrary, the non-agricultural worker and the self-employed farmer are in a better position to benefit from such a refugee inflow.

*Université catholique de Louvain, CORE and FNRS, B-1348 Louvain-la-Neuve, Belgium - *E-mail:* Jean-Francois.Maystadt@uclouvain.be. The first author is grateful to the Center for the Study of Forced Migration (University of Dar es Salaam), Joachim De Weerdt (EDI) and Jonathan Wosley, UNHCR (in particular Yacoub El-Hillo and Eveline Wolfcarius), WFP (in particular Taban Lokonga), several NGOs and many other respondents for the support received during fieldwork activities. We are indebted to all interview respondents for the time they accepted to dedicate to our research. The first author also ackowledges financial support from the Fonds National de la Recherche Scientifique (FNRS) and from the Communauté Française (Bourse de voyage 2008).

†Fund for Scientific Research (FWO), University of Antwerp and ECARES (ULB)

Keywords: Refugees, Tanzania, agglomeration
JEL Classification: O12, O18, R12, R23
We are grateful to Joachim De Weerdt, Fernanda Estevan, Miren Lafourcade, Luca Marchiori, Walter Steingress, Jacques Thisse as well as participants to PAI meeting (13 Octobre 2008, Bruxelles), ECRU seminar (21 November 2008, Louvain-la-Neuve), the fourth annual HiCN conference (5 December 2008, Yale University, New haven), the International Conference on New Economic Geography (13 December 2008, Passau University), the UCL DW (22 January 2009, Louvain-la-Neuve) and the second CEPR conference on "Transnationality of Migrants" (23 January 2009, LLN) for useful comments. We are indebted to the World Bank and Kathleen Beegle for granting access to confidential geographic data of the Kagera Health and Demographic Survey (2004). Any remaining errors are naturally ours.

1 Introduction

Every year, thousands of people are forced to leave their countries in a desperate attempt to find a safe heaven, their life being threatened by civil conflicts. End of 2007, UNHCR (2007) reported 11.4 million refugees in the world, whose 2.2 million originated from Sub-Saharan Africa. Contrary to some popular ideas, these refugees do not invade the industrialized world but are in widespread majority, hosted by their neighboring countries. Media and host governments have often pressured the international community to help them in supporting the burden induced by this massive inflow of refugees. However, very few is known about the exact nature of the impact of the refugees on the local hosts. From a policy point of view, the United Nations High Commission for Refugees (UNHCR) seems increasingly aware that the issue of the impact of the refugees on their local hosts has been overlooked for too long. For example, in Tanzania, the UN agency for refugees is implementing a new program easing the transition phase following the closure of all camps in the region of Kagera. [1] A better understanding of the main channels through which the local population is affected by massive refugee inflows would help in improving the efficiency of such programs.

Our collective memory tends to perceive refugee camps as an unorganized mass of temporary tents and occupied by passive refugees under assistance. Nevertheless, this common view is far from many experiences. First, some temporary situations can end up by lasting sometimes very long. This has resulted in what UNHCR has called the problem of protracted refugee situations (Crisp, 2003; Slaughter and Crisp, 2009). City-sized refugee camps have mushroomed in often very poor areas, where inhabitants themselves struggle to make a living. Despite their traumatic experiences and their poor health conditions when they arrived, still these refugees came in with productive capacities, even with some assets (human capital, livestock, etc) or at least, keep networks to get access to these assets (de Montclos and Kagwanja, 2000; Werker, 2007). Therefore, refugees are likely to endorse some important economic functions and have significant impacts on their hosts' livelihoods. The purpose of this paper is to contribute to a better understanding of the long-term impact of refugees on the local communities.

The issue of the impact of a refugee settlement on local economies does not seem to have attracted very much research interest. As far as we know, Chambers (1986) in an article entitled "Hidden Losers? The impact of rural refugees and refugee programs on poorer hosts" is the first to have paid attention to this issue. Based on scattered evidence and rural experiences, this author argues that the presence of a refugee camp has mixed consequences for the host population through an increase in price, wage competition and competition for natural resources. The better-off and more visible hosts would be more likely to gain from the presence of refugees and refugee programs while by contrast, poorest hosts could loose from competition for food, work, wages, services and common property resources. He points to these vulnerable hosts as the hidden losers. Since then, authors such as Kuhlman (2002), Whitaker (1999), Landau (2004) and several papers from the UNHCR´

[1] In June 2008, the Lukole camp (Ngara district) was the last camp of the region of Kagera to be closed making the Kagera region free of refugees for the first time since about 15 years.

s Evaluation and Policy Analysis Unit have provided field-work support of this contrasted impact. Unfortunately, we are not much more advanced than the "scattered evidence" on which Chambers (1986) bases his analysis. As summarized by Whitaker (1999, 2), "refugees are assumed to have a different impact on diverses classes, genders, sectors and regions within the host country, but little empirical evidence has been done on this issue."

Few economists seem to have pay attention to this subject or at least, related ones. Some authors such as Hatton and Williamson (2004) focus on the causes that lead refugees to seek asylum in developed countries and in particular, in the European Union. They also study its impact on native-born workers. In a cross-country setting, Azam and Hoeffler (2002) also test empirically the factors explaining the number of refugees per country of origin. Stark (2004) formalises the dynamic process that distinguishes a refugee flow from an immigration move while Bubb et al. (2007) theoretically study the system of refugee protection based on the 1951 Convention Relating to the Status of Refugees as a Pareto-improving contract that bound states to provide a more efficient level of the global good of protection. The issue of efficiency of food aid which is also related to the subject has also been much studied. Its impact on prices and trade have been explored by notably Barret (2002), Dercon and Krishnan (2004), and Donovan et al. (2006).[2] Nevertheless, those subjects are much more general than the presence of refugees, per se. As far as we know, two papers are much closer to ours, at least in studying related issues in the Tanzanian context. First, AlixGarcia and Saah (2008) assess the impact of a refugee camp presence through changes in agricultural prices between 1995 and 1998. They also test the impact of massive refugee inflows on the holding of assets. Two important results might be emphasised. On the one hand, they found a significant increase in the prices of some agricultural goods (bananas, beans and milk) and a decrease in the price of the aid-delivered good (maize). On the other hand, they found some indications of coping strategies from households reducing their expenditure of more expensive consumption goods and acquiring more valuable durable assets. Another related paper is the one by Baez (2007) who assesses the short and long run causal effects of hosting refugees on health conditions of children. The author offers evidence of adverse impacts, almost 1.5 years after the shock, on children's health.

This paper is organized as follows. In the next section, we will describe our case study and list a series of hypothesized effects that have been collected through our fieldwork. [3] Section 3 introduces a theoretical framework aiming at identifying who are likely to benefit or suffer the most from the changes induced by the refugee inflows. Section 4 presents the empirical results of our study and discuss the robustness of these results. Finally, we will

[2]Our paper is also related to studies assessing the impact of migrants on local labour markets (card, 1990; Hunt, 1992). However, our approach differs from these wage-type equation analysis in the sense that we seek to undertake a general equilibrium assessment of the impact of refugee inflows on the welfare of the local population, not restricted to labour employment.

[3]The quantitative analysis presented in this paper has been fed by a two-month iterative field research (Udry, 2003). In order to refine some of our hypothesis, about 30 interviews were conducted; data (notably refugee camp location and population) were completed; and some reports were collected to better understand the economic environment of the region and the issues (management, interaction between refugees and local people) related to the refugee presence.

conclude in section 5.

2 Refugee inflows in Kagera

2.1 Context

The Kagera region is a very remote region in North-Western Tanzania. The region is impacted between the Victoria Lake, Uganda, Rwanda and Burundi. Kagera is one of the poorest regions of the country in terms of income per capita with an average of 149,828 Tanzanian Shillings (Tzs, i.e. 166$ a year) per habitant (NBS, 2003). In 2002, about 2 million people lived in a region of 29,241 squared kilometers and mainly rely on subsistence agriculture. One particularity of this region results from its recent history of refugee hosting that makes it an extraordinary "laboratory" to study the impact of refugee inflows on the local population. On the one hand, the magnitude of the phenomenon is a case in point. From October 21, 1993, between 250,000 and 300,000 Burundians fled into Tanzania following the assassination of the President of Burundi. As reported by Rutinwa (2002, 28), a new influx of refugees of 250,000 came then from Rwanda from April 28, 1994, within only 24 hours. This influx generated from the crash of the plane carrying the Presidents of Rwanda and Burundi is known as the triggering factor of the Rwandan genocide. This movement, described by the United Nations High Commission for Refugees (UNHCR) as the largest and fastest exodus it had ever witnessed, was followed in the next two months by a number of nearly one million refugees fleeing Rwanda to Tanzania. To put the reader into the picture, in 1995, there remained about 700,000 refugees in the region of Kagera, whose population accounted for about 1.5 million people at that time. [4] Such a human inflow, representing more than one third of the regional population (even more than one half in Karagwe and Ngara districts), had a financial counter-part. Interviewees stressed the massive flow of money that entered into the local economy through the humanitarian pipeline. On the other hand, the unanticipated and localized nature of the events provides a tool to isolate the impact of the refugee inflows on the local population from other factors. As witnessed by a local aid worker, "they came very unexpectedly. The local population was never expecting such a thing. Just overnight, so many people were around ... They came like a swarm of loco bees". The unanticipated nature of the events linked to political assassinations is also underlined by AlixGarcia and Saah (2008). Furthermore, a change in the refugee policy implemented by the Tanzanian Government restricts the movement of the refugees to 4 kilometers around city-sized refugee camps. [5] These movement restrictions, coupled with geographical features limiting the spatial spread of the impact (Baez, 2007),

[4] Other refugees also came from the Republic Democratic of Congo from 1997. However, our region of interest, Kagera, did not host these refugees. Congolese camps are mainly located in the neighboring region of Kigoma.

[5] Tanzania has a long history of hosting refugees. In 1972, refugees already fled from Burundi and were directed towards, the so-called old settlements. However, these refugees are very distinct from the 1993-1994 waves. These "old" refugees are not settled into closed camps; were targeted by an "open-door" policy whose aim was to integrate these newcomers (even through naturalization)

provides an exceptional framework to distinguish refugee-hosting areas from others.

2.2 Observations

Given the magnitude of the phenomenon, the massive inflow of refugees is likely to have affected the local economy. Based on fieldwork and secondary sources, the following main impacts have been identified:

- The price of some goods seems to have sharply increased, threatening the food security of some households. The increase of prices should have resulted from an increasing demand from aid workers but also from the refugees themselves. The arrival of international organizations (UNHCR, NGOs) and their staff (local and international) induced a significant increase in demand from people with much higher purshasing power. Food aid could also have had a surging effect on prices. The World Food Program (WFP), the UN agency in charge of providing food to refugees, could purchase on the local markets. But the main effect came from the refugees themselves. In order to diversify their food diet, refugees exchange the received ratio of maize, maize flour or cooking oil against other products such as bananas, cassava, palm oil, beans, rice, meat, fish, alcohol, etc. Whitaker (1999) reports that about 75% of the food distributed to refugees were traded. More conservative assessments by WFP gives an estimation of about 20-30% (WFP and UNHCR, 1998). Refugees also sell non-food items such as blankets and plastic sheets. Such trading activities is easy to observe and even institutionalized by the aid community and the Ministry of Home Affairs (MHA), through the creation of a so-called common market around each refugee camp, where refugees and local people were invited to trade. For example, the common market of Lukole opened between 1994 and 2003 and closed to one of the largest Rwandan refugee camp, is estimated to have been the biggest market of the Kagera region, after the one of the capital town (Bukoba).[6] Despite these observations, this perceived effect remains an empirical question. As reported by one of our respondents, while the increase in price has been imputed to the refugees arrival, it is difficult to distinguish this increase from the whole issue of inflation in the country.
- One of the most often reported effects is the use of refugees as cheap labor by the local population. Using refugees to till and harvest land, local agricultural production is

and ensure self-sustainability of their livelihoods on the Tanzanian territory and were located far from the borders of their country of origin (Rutinwa, 2002). It has to be noted that despite the movement restrictions, 1993-1994 refugees could still receive permission to work outside the camp, provided they came back at night.

[6]The common market is one of the first institution to be closed by the Tanzanian authorities, when they are planning to close a refugee camp and try to give refugees some incentives to repatriate. It has also to be noted that in a 2001 article on refugees in western tanzania, *The Economist* (TheEconomist, 2001) also witnesses such an increase in demand: "Half a million extra months increase demand for almost everything. The more enterprising locals profit from this, ploughing up extra acres to take advantage of soaring food prices, selling cooking pots and clothes in the camps, and even supplying foreign aid workers with chocolate , cheese and satellite-television dishes".

reported to have in some cases doubled following the refugee arrival, land availability not being a major constraint in Kagera (WFP and UNHCR, 1998). Cheap labour was also used in sectors such as construction, housekeeping or catering. As confirmed by Pr. Rutinwa, refugees were paid below the minimum wage that was about 1000 Tzs a day around that time ... During a focus group he undertook for his study (CSFM, 2003), one woman declared " At least one thing we like about these refugees, they are cheaper. They are managable in terms of the amount you pay them compared to what you pay a Tanzanian to do the same amount of work". Such an increase in cheap labour should generate a downward pressure on the wages of agricultural workers. Whitaker (1999) reports that the wage paid to casual laborers dropped by about 50% in many areas. Such a depressing effect on wages for the less skilled is supported by historical evidence provided by Chambers (1986).

- The effect on the wages seems to have been the opposite for the most educated people. As illustrated by one interviewee who used to be chief accountant in a cooperative union when the refugees came in, his salary was below 100,000 Tzs, while the drivers of the international organisations could earn more than 200,000 Tzs. This confirmed the observation by Waters (1999) that salaries in the relief operations were about two to three times the level of salaries for similar positions elsewhere in Tanzania. The resulting increase in the wage of more skilled people is also reported to have attracted employees from other areas. Landau (2004, 45) reports that "A district that was once designated as a labour reserve has now become a major destination for Tanzanians from all over the country seeking waged employment with international and non-governmental organizations."

- Business also seemed to have flourished in the refugee-hosting areas. Whitaker (1999, 7) underlines that "With the increased local market, there was an upsurge in business and trade conducted by both local hosts and refugees. Tanzanian entrepreneurs from around the country also flocked to the area. Commercial centers developed in the refugee camps with daily markets and countless shops and restaurants". The economic landscape has completely changed in this respect. Some sleepy (Ngara market, Omukalinsi) or even desert places (Lukole, Benaco, etc) became very active market places. Business seemed to go on even when the refugees repatriate. An aid worker involved in repatriation, underlined how refugees coming originally far away from the Burundian border, repatriate just behind the border to continue their trading activities. The variety of non-agricultural products also seemed to have expanded as "several enterprising Tanzanians even opened shops with catered to expatriate aid workers' tastes for chocolate, cheese, European wines and satellite televisions" (Landau, 2004, 47). This is also the case for local people as some utilities were sold by refugees and new products (cans, etc) entered this very remote area. Despite this boom, the impact on the existing businesses was rather ambiguous. The renewed attractiveness of refugee-hosting areas seems to have gone along with fiercer competition following the entry of other entrepreneurs. Such reported increase in competition could have driven some existing petty businesses out of the market. First, a lot of entrepreneurs are reported to have come from Mwanza, Shinyanga and other Tanza-

nian regions to open a business. People of Haya-origin from Kilimanjaro were often reported as very business-minded and keen to jump on these new business opportunities. As underlined by one active district refugee coordinator in one refugee camp, refugees were also allowed to conduct businesses and were very active in this respect. One of our respondents also pointed to the 'unfair" nature of competition, because of the absence of taxes for commodities traded by refugees and the free services provided to them: "it was unfair competition because you cannot compete somebody who has no overhead costs." At the end, the effect on the existing businesses will be an empirical issue.

- Infrastructure seemed to have largely improved following investment made by international organisations in terms of road accessibility. Whitaker (1999, 12) noted that "In Kagera region, more than 15 million dollars went towards the rehabilitation of main and feeder roads, airstrips, and telecommunications infrastructure", making "internal transportation cheaper and easier for host communities". This might be more important that it seemed in a region where the remoteness of the village in which one his born is an important determinant of the likelihood of growing out of poverty (DeWeerdt, 2006).

- Another important reported effect is the improvement of health and sanitation services. If the services were not necessarily available to the local population at the origins of the crisis, the UNHCR and its implementing partners have progressively made these services available to the local hosts. Around 30% of the health services beneficiaries are reported to be local people. Well, there is no doubt that the standards of these services are very much higher compared to what was delivered to the local population in these very remote areas, before the refugees came in.

- We should also underline the possible destabilizing effects of such population movements. First, the sudden flow of so many people had devastating environmental effect (Berry, 2008). As witnessed by one of our respondents, "When refugees came in, they started to pull down the timber, the wooden materials from the buildings, ... the schools, even the coffee plantations to get firewood". This had dramatic effect, in particular for women, who had to allocate more time to collect firewood, in a region where it constitutes the main source of energy for cooking activities. [7] Although aid workers contest it given very tight controls they implement, the spread of desease has been attributed to the coming of refugees living in poor sanitary conditions. At least, regarding wood collection and desease spreading, one could imagine that these effects have been spatially more limited than the economic effects described above (a perimeter of 15-20 kilometers has been reported for the wood destruction around the camps, in the worst cases). Fieldwork also suggests that these effects have decreased overtime following NGO interventions in these respects. Finally, security problems have been reported by officials and police officers. The attribution to the refugees of

[7] Since then, one could observe an increased awarness among aid agencies for this issue. Notably, devastated areas around old camps (for example in Karagwe) contrast with more recent Congolese camps in the region of Kigoma where wood collection has been strongly regulated.

the security problems has been debated (CSFM, 2003). However, one should, at least, recognize that the inflow of refugees has gone along with a cross-border diffusion of weapons and an increase in criminality (who ever is perpetrating these crimes).

These reported main effects constitute an interesting starting point to understand the complexity of the issue. Nevertheless, we may wonder what are the magnitude, the persistency and the spatial distribution of these effects. First, such narrative evidence does not tell us whether the negative effects are likely to be compensated by some economic benefits. Given the nature of the economic effects, the benefits are also likely to be unevenly redistributed. [8] As a first approximation, looking at the change of consumption per adult equivalent (PAE, see data description) for all surveyed individuals between 1991 and 2004, the last sticks of the figure 1 suggest that households living closed to refugee camps have increased their consumption PAE but less proportionally than in other areas. But the picture is qualified when looking at poverty groups. [9] Initially richer people seem to have better seized the opportunities generated by the proximity to refugee camps. They indeed increased their consumption PAE much more (about 38%) than their counterparts in other areas (about 16%). On the contrary, the poorest have seen their situation improved less (10%) than their counterparts in other areas (16%). Such chart suggests that the Chambers (1986)'s hidden losers can well hide some winners. Second, we do not have any idea about the persistency of the above reported effects. For example, the changes in factor prices (wages, prices of goods, ...) should have been followed by modifications in the consumption and production behaviors. As suggested above, some negative effects (wood devastation, disease risk) may have been reduced overtime. Therefore, section 3 will study in a general equilibrium setting how changes in the good and labour markets will affect differently several agents of the local economy. Finally, we do not only conjecture economic gains to be distributed differently among the refugee-hosting population. The figure 2 also indicates that contrary to the common wisdom considering the presence of refugees uniquely as a burden, the establishment of refugee camp does not either generate more people to get out of the sample (due to untraced migration or death) between 1991 and 2004 or give more incentives to migrate (traced) outside the initial village.[10] Clearly, the establishment of a refugee camp does not only generate negative externalities. Given the localized nature of the

[8] Many respondents were quite aware of this inequality effect : " Those who were creative, who were business-minded, they made a lot of money. Even in the villages, those who are farmers, everything produced on the farm, had a market there. For all over the district, it was going there on bycicle, going to get money from the camps. The situation was worst for the most vulnerable groups like the eldery who do not know what to eat next days, those with inability, those with chronic illness but those who were energetic, creative, they made a good money out of it."

[9] Following Collier et al. (1986), we transformed the consumption data of the KHDS data (see data description) by adult equivalent and taking into account price differences between locations. Based on 1991 data, we create four groups, which should reflect the level of poverty in the sample. The identification of people in a village closed to a refugee camp is based on the reply from the community leader to the KHDS question "were there any refugee settlement closed by?". Other identification criteria are described in section 4.1.

[10] Among the 3794 individuals that were interviewed in 1991 and 2004 in our database (therefore, traced when moving from their original village to another village, another region or another country),

negative effects (environmental degradation, disease, etc), we might then wonder whether the balance between economic benefits and costs differs over space. The existence of a non-monotonic relationship between the proximity to refugee camps and welfare is an empirical question. On the one hand, some respondents were convinced that villages very closed to refugee camps, all in all, suffer from the negative effects, while habitants at intermediate distance could maximize the economic benefits generated by the establishment of refugee camps and minimize the costs. [11] On the other hand, other respondents clearly reject this non-monotonic hypothesis. For them, the economic benefits such as health accessibility, outweigh the negative externalities, even for the closest to refugee camps. Given our questions related to the persistency and the distribution of the above effects, the paper will seek to assess, both theoretically and empirically, how the establishment of refugee camps in the region of Kagera has affected the local population in the long run and through which channels. A particular attention will be given in section 4 to allow for a differentiated effect of such establishment over space.

3 Theoretical framework

The theoretical framework should help us to identify how agents may adapt to the refugee arrival and the related changes in factor prices. The purpose is to describe how some agents may be able to benefit from the inflow of refugees, while others seem to suffer from it. The below theoretical model will mainly focus on the effects on the labour and good markets. [12]

2,830 individuals are still living in the same location (75%) in 2004. The emigration rate computed by village is quite unevenly distributed (from 8% to 43%). What appears rather clear is that the establishment of refugee camps does not seem to have been a big dispersion force. On the contrary, a higher share of the population followed overtime tend to move away from unaffected villages. If the same exercise was performed with the active population, i.e. 1,330 individuals followed overtime, 1011 lived in 2004 in the same village than 10 years ago (76%). The contrast between refugee-hosting communities (17%) and others (26%) is even stronger, confirming the slight indication that contrary to common wisdom, the establishment of a refugee camp may well be an attraction force.

[11] In a report drafted in September 1995 (about one year and a half after the arrival of Rwandan refugees) for a local NGO, Adams et al. (1995, 31) support such a non-monotonic relationship by stating that "While the whole district has been affected by the arrival of the refugees (through improved transport links, rising prices, loss of livestock, damage of schools and hospitals, the re-direction of already over-streched district services to the refugees, an expanding economy, increased opportunities for trade, etc), the negative consequences have tended been very localized. Those villages in the immediate vicinity of the camp have borne the brunt of theft and robbery, the loss of water, fuel and construction materials, threats to household food security, health problems, and social change."

[12] AlixGarcia and Saah (2008) use another framework to understand how refugee inflows affect local prices, through food aid provision. Although interesting, we cannot use a similar framework given our focus on long-run General Equilibrium effects and the distribution of these effects. Our theoretical framework could also be compatible with a New Economic Geography (NEG) extension that would be able to incorporate in a GE setting the change in transportation costs and the mobility of some agents, possibly induced by the refugee inflows.

Controlling for other factors or explanations will be an empirical issue.

We represent a local economy whose population (normalised to one) is composed of H landed people and L landless people. The landed people are endowned with one unit of land, that can be invested in one of the two sectors of the economy, the agricultural sector or the manufacturing sector (non-agricultural). The landless people are composed of L_a agricultural workers and L_m manufacturing workers.

3.1 Preferences

The preferences of all individuals are represented by Cobb-Douglas preferences, with CES sub-utility for the differentiated non-agricultural goods. [13] The utility of the agent i is composed of its consumption of homogeneous agricultural good A and differentiated non-agricultural goods M.

$$U_i = M^{\mu} A^{1-\mu} \tag{1}$$

$$\text{with} \quad M = \left[\int_0^N q_m(s)^{\frac{\sigma}{\sigma-1}} ds \right]^{\frac{\sigma-1}{\sigma}} \tag{2}$$

$$\text{with} \quad \sigma > 1$$

$$\text{s.t.} \quad GM + p_a A \leq Y$$

σ represents the elasticity of substitution between two varieties of the non-agricultural good, while G is the price index of the non-agricultural goods.

From the identical preferences of the individuals, one can obtain the demand functions and we can re-write the price index:

$$q_m(s) = p_m^{-\sigma} G^{\sigma-1} \mu Y \tag{3}$$

$$q_A = \frac{(1-\mu)Y}{p_a}$$

$$G = p H_m^{\frac{-1}{\sigma-1}} \tag{4}$$

$$\tag{5}$$

The indirect utility function will be given by introducing the demand functions into the utility function:

[13] Therefore, I will assume that the share of expenditure on the agricultural good and the manufacturing good remains constant. Non-homothetic preferences might be more suitable in our context, like Murata (2008) or Murphy et al. (1989).

$$V_i = \Theta \frac{y_i}{G^{\mu} p_a} \quad (6)$$

$$\text{with} \quad \Theta = \mu^{\mu}(1-\mu)^{1-\mu} \quad (7)$$

3.2 Technologies

The economy is represented with two sectors of production. The agricultural sector performs under constant returns to scale, using one unit of land and L_a units of agricultural labor. The non-agricultural sector exhibits increasing returns to scale, with non-agricultural labor as an input. We assume that the price of the agricultural good is equal to 1.

$$\pi_a = q_a - w_a L_a \quad (8)$$

$$\text{with} \quad q_a = L_a^{\alpha}$$

The first-order condition gives $w_a = \alpha L_a^{\alpha-1}$ and $\pi_a = (1-\alpha)L_a^{\alpha}$. The non-agricultural sector needs β variable requirement of skilled labor L_m and one landed person, who decides to become entrepreneur. The structure of this sector is the monopolistic competition.

$$\pi_m = p_m q_m - \beta w_m q_m - w_m \quad (9)$$

Deriving π_m with respect to p_m gives $p^* = \frac{\sigma}{\sigma-1} w_m \beta$, i.e. the monopolistic competition result following which the mark-up of price over marginal cost is constant. By the zero profit condition, the equilibrium production is independent of the number of entrepreneurs: $q^* = \frac{\sigma-1}{\beta}$. Equalizing this equilibrium production with the demand equation, we obtain the following wage equation, i.e. the wage given to the skilled workers and the rent kept by the entrepreneur:

$$w_m = \frac{\mu L_a^{\alpha}}{\sigma H_m - \mu(L_m + H_m)} \left[\alpha + (1-\alpha)H_a\right] \quad (10)$$

We can show that the non-agricultural wage will be positive provided the number of non-agricultural workers is sufficiently large. Otherwise, the non-agricultural sector does not exist.

$$w_m > 0 \quad \text{if} \quad L_m > \frac{\sigma - \mu}{1-\alpha} \frac{\alpha + (1-\alpha)2H}{3\mu} = \widetilde{L_m} \quad (11)$$

This condition will be most easily met if:

- σ is low, i.e. the non-agricultural goods are sufficiently differentiated.
- μ is large, i.e. the constant share of expenditure on non-agricultural goods should be sufficiently large.
- H is large, i.e. the number of landed people should be large enough.
- α is small, i.e. the labor productivity in the agricultural sector should be small enough.

Equalizing w_m and r_a determines the distribution of non-agricultural entrepreneurs:

$$H_m = \frac{\mu[\alpha + (1-\alpha)(H + L_m)]}{(1-\alpha)\sigma} \tag{12}$$

3.3 Theoretical predictions

Given the fact that H_m does not depend on L_a, a change in landless agricultural workers, which represent the refugee inflow in our present case, can be shown to affect the different agents through two channels: the changes in income and in prices. V_a, V_{π_a} and V_m denote respectively the indirect utility of the agricultural worker, the landed investing in agricultural activities, the landed becoming a non-agricultural entrepreneur. We assume that given the existence of the non-agricultural sector, $L_m > \widetilde{L_m}$.

- The agricultural worker suffers from an increase both in competition on the labour market ($\frac{\delta w_a}{\delta L_a} < 0$) and in prices following the demand pressure on the good market ($\frac{\delta G}{\delta L_a} > 0$ if $L_m > \widetilde{L_m}$).

$$\frac{\delta V_a}{\delta L_a} = \Theta \frac{\frac{\delta w_a}{\delta L_a} G^\mu - \mu w_a \frac{\delta G}{\delta L_a}}{G^{2\mu}} < 0 \tag{13}$$

- The landed investing in agricultural activities benefits from an increase in its agricultural rent ($\frac{\delta \pi_a}{\delta L_a} > 0$) but suffers from the increase in the price index ($\frac{\delta G}{\delta L_a} > 0$). Given the fact that H_m does not depend on L_a, one can show that the indirect utility has increased following the refugee inflow.

$$\frac{\delta V_{\pi_a}}{\delta L_a} = \Theta \frac{\frac{\delta \pi_a}{\delta L_a} G^\mu - \mu \pi_a \frac{\delta G}{\delta L_a}}{G^{2\mu}} \tag{14}$$

$$\text{Given} \quad V_{\pi_a} = L_a^{\alpha(1-\mu)} \frac{(1-\mu)^{1-\mu} H_m^{\frac{\mu}{\sigma-1}} (\sigma-1)^\mu (1-\alpha)}{[\alpha + (1-\alpha)H_a]^\mu (\sigma\beta)^\mu}$$

$$\frac{\delta V_{\pi_a}}{\delta L_a} > 0 \quad \text{if} \quad L_m > \widetilde{L_m} \tag{15}$$

- The non-agricultural entrepreneur (also true for the non-agricultural worker) also benefits from an increase of its income but has to support an increase in the price index. However, given the fact that H_m does not depend on L_a and L_a increases w_m,

one can show that the indirect utility of the non-agricultural agent will increase with the number of landless agricultural workers.

$$\frac{\delta V_m}{\delta L_a} = \Theta \frac{\frac{\delta w_m}{\delta L_a} G^{\mu} - \mu w_m \frac{\delta G}{\delta L_a}}{G^{2\mu}} \tag{16}$$

$$\text{Given} \quad H_m \quad \text{and} \quad L_m > \widetilde{L_m} \quad , \quad \frac{\delta V_m}{\delta L_a} > 0$$

4 Empirical analysis

Based on the theoretical predictions, this section seeks to assess the relationship between the refugee inflow and the welfare of the hosting population. Furthermore, we would like to see whether the net costs and/or the net benefits are likely to be unevenly distributed among the refugee-hosting population.

4.1 Data description

We use the Kagera Health and Development Survey (KHDS) dataset collected by Economic Development Initiatives (EDI) and the World Bank (Beegle et al., 2006). Based on the World Bank LSMS (Living Standards Measurement Study) standards, the KHDS data provide a very comprehensive survey on several dimensions of the individual and household well-being such as the levels of consumption, income, assets, the occurrence of shocks, but also some community and facilities characteristics such as the availability of public services, the participation to collective groups, etc. In 1991-1994 (4 waves), the KHDS interviewed up to four times 915 households and their members from fall 1991 to January 1994. The households were selected from 51 communities, from the 6 districts of the region of Kagera. In addition to the representativeness of this survey, one interesting feature of this survey is the outstanding exercise of tracing most individuals from the original 915 households, about ten years later, in 2004. Because people had move out from their original households, the KHDS 2004 interviewed about 2,700 households and their members, including those having moved outside their village of origin, the Kagera region and even Tanzania. As indicated by Beegle et al. (2006), the field team achieved an excellent rate of recontact of 93%.

These data are particularly adequate for assessing the impact of the refugee inflows of 1993-1994 on the local population. First, we can be certain that the first wave of the KHDS surveys has been undertaken before October 21, 1993, date of the assassination of the President of Burundi and signaling the start of the refugee crisis in the Kagera region. [14] Therefore, the data should allow us to distinguish the effect of the refugee inflows from

[14] Given the procedure followed to interview a household (i.e. when one household drops it is replaced by a new one starting again wave 1), the facts that the first passage of fieldwork took place between September 1991 and May 1992, that each passage lasts between 6 and 7 months, and that no household was interviewed for the first time in the fourth passage, a household would be interviewed in the first wave on July 1993 at the latest (the majority between September 1991 and may 1992).

some initial differences between villages, households or individuals. Second, the location of the different villages throughout all the region allows us to introduce a key heterogeneity in our sample, depending on whether the individuals are living in a village closed to a refugee camp or not. We use a sample of 3,510 individuals and 812 households (that duplicate into 2126 households by 2004), that have been interviewed in 1991 (before October 21, 1993) and in 2004. [15]

- The *dependent variable* is computed as the consumption (transformed into log) per adult equivalent, in real 2004 prices. The adult equivalent transformation is applied using the method proposed by Collier et al. (1986) for Tanzania, while the Lapeyres and Fisher indexes are used for price correction. [16]
- The *treatment variable* should capture the impact of the refugee camp per se. One possibility is to use a dummy variable, denoted RC, indicating whether the community leader has replied positively to the KHDS question "Were there any refugee settlements closed by?". [17] Less vulnerable to reporting errors, our fieldwork allows us to proxy the effect of the establishment of a refugee camp by the distance between any village and the refugee camps as well as the estimated number of refugees by camp (using the 1995 estimates of the refugee population collected through fieldwork and considered as the peak time of the refugee presence). We compute an average of the refugee population weighted by an exponential distance function, called the refugee impact : $RI_{h,v} = \sum_{c=1}^{13} exp^{-\alpha d_{v,c}} pop_c$, where c, from 1 to 13 refugee camp and v is the village where the household h is living. In the baseline results, we will use this measurement of refugee impact with $\alpha = 1$ and keep other measurements for further robustness checks.
- The *activity variables* define the main occupation of each household. The main occupations introduced are subsistence agriculture, agricultural self-employed , wage employment (distinction between unskilled -farming and fisheries - and skilled labour is made), and non-agricultural self-employment (business and livestock). To determine what is the main occupation of a household, we first compute on the basis of the time spent in a particular activity by each household member and the value of profits or wage realized through this activity, the total income generated by each household

[15] To be more precise, from 3796 individuals followed overtime, 254 individuals have been dropped due to missing income data and 32 due to duplication for reasons given in Beegle et al. (2006, 39).

[16] The aggregated consumption data provided by EDI (http://www.edi-africa.com) have been used for comparability reasons (recall periods, common definition of components).

[17] It has to be noted that Baez (2007) uses the fact to belong to the districts of Ngara and Karagwe as an identification. We have some concerns about using this indicator as it might rather capture a district-specific unobserved effect rather than the effects of the refugees. First, all villages from Ngara and Karagwe districts do not all report a closed-by refugee settlement. Only eight out of eleven do so. Furthermore, four villages in three neighboring districts also report closed-by refugee settlements. Provided we use the community questionnaire as identification criteria, this would represent 12 villages, i.e. approximately 192 households out of about 816 households or 24%. It represents between 31 and 29 percent of the population in Kagera region.

and the share of this income over the total income generated by this household. We apply the Lapeyres and Fisher index and the adult equivalent method of Collier et al. (1986). [18] We define a main activity as the one for which the percentage is above the average of the rest of the population. This makes the definition less dependent to an arbitrary threshold (like e.g. larger than one half) but disregard the marginal income sources. [19] For example, an individual generating 99% of his income from subsistence income and 1% from business is better to be considered as mainly involved in subsistence agriculture. On the contrary, someone doing more than 15% of his income (sample average) will be considered as doing business. It also means that one household can have more than one main activity, in case it would specialize into two activities. The activity variables combined with the refugee one will be key to see whether different payoffs are associated with some activities in refugee-hosting areas, compared to other (control) villages.

- *Household characteristics* are introduced, such as the size of the household, the proportion of literate members, the number of children, the value of land as the main asset of the household (transformed in PAE and into logarithm).
- As explained below, other variables will be introduced and explain progressively.

[18] We use the aggregated data given in the 1991-1994 KHDS database and apply the same methods described in KHDS (2004) to compute the 2004 data. For subsistence income, we convert the value of each consumed item according to its amount, its season and its imputed value. The income from self-employed agricultural activities is computed on the basis of the gross revenues minus the costs of household-level activities in farming. Employment income is recognized as the income received as an employee of a private individual or of an institution other than the household, for remuneration in cash or in kind (KHDS, 2004). The household-level value is computed by summing the employment income earned by each member, multiplied by the time this member reports to have worked at each job in the last 12 months. The Laspeyres and PAE transformation have been applied. We replicate the KHDS (2004) method for the year 2004 to compute the non-agricultural self-employed activities.We compute the livestock income by computing the net revenues from livestock. The second method based on reported net revenues and proposed by KHDS (2004) is used. deMel et al. (2009) indeed show that using reported profits provides more accurate measure of profits than the method using (underreported) revenues and expenses. Fishing activities are also added to this category to make 1991 data comparable with 2004 data. One technical problem for the non-agricultural activities appears regarding the way we should treat losses. Indeed as some losses are encountered, some share of income are above 1 or below 0. For 1991, we have about 4 households (or 16 individuals) who have a share of income generated from business above 1, mainly due to losses in the agricultural sector, and about 4 households (or 16 individuals) who have a negative share of income due to business losses. For 2004, the former increases to 92 individuals and the later reduces to one individual. We restrict the negative values to zero.

[19] For another issue (work satisfaction) but a similar choice of threshold, Winkelmann and Winkelmann (1998) also take the sample average as a threshold.

4.2 Baseline results

We start by estimating the following regression at the household level, h:

$$\begin{aligned} log(V_{htv}) &= \beta_0 + \beta_1(RI_{htv}) + \beta_2 Activity_{ht} + \beta_3 Activity_{ht} * RI_{htv} \\ &+ \beta_4 Z_{ht,1991} + \beta_5 \alpha_t + \beta_6 \alpha_h + \epsilon \end{aligned} \tag{17}$$

- V_{htv} represents the indirect utility of the household h living in village v at time t, measured by the consumption per adult equivalent, correcting for price differences, as explained above.
- RI is the treatment variable, defining the extent to which a household is affected by the refugee inflow. As explained above, we use an inverse distance function (exponential) weighted by the 1995 refugee population.
- $Activity$ defines the main occupation(s) of the household h.
- Z represents some household characteristics such as the size of the household, the number of children, the proportion of literate members and the stock of land assets.
- α_h and α_t are respectively a household fixed effect and a time dummy

Like most household panel analysis, we start by restricting our sample of households to the members that have not been traced (excluding new households and/or those that have moved to another location). A household fixed effect is used to reduce the endogeneity risk by controlling for any unobserved time-constant household characteristics, while a time dummy captures any change overtime which is common across all villages. This seems to be particularly important when introducing the activity variables, as unobserved household characteristics might determine the choice of activities and affect the ability of the household to increase his consumption overtime. Like for all our specifications, to correct for heteroskedasticity and correlation of errors within the same district, we use robust standard errors in the baseline regressions, by clustering at the district level (Bertrand et al., 2003). In column (1) of table 3, we found a non-significant coefficient for the effect of the refugee impact on welfare. In the rest of the table, we found a non-monotonic relationship between the establishment of refugee camps and the dependent variable. We also introduce the activity variables. With the exception of the non-agricultural worker who seems to benefit from the refugee presence, none of our theoretical predictions are supported. [20] However, these results should be considered with suspicion as they may be threatened by important estimation biases. As in many panel data analysis that do not interview those creating a new household (e.g. children in 1991 that compose a family before 2004), table 3 does not

[20] In regression (4), like in all the remaining specifications, we test the robustness of our results to the inclusion of (time-varying) village dummies. There is an obvious risk that such dummy captures part of the effect of the refugee impact but robustness to its inclusion , would strengthen our results

include the split-off households. As explained by Dercon and Shapiro (2007) and Verwimp and Bundervoet (2008), not including the split-off households may create some bias in the estimations. Assume for example, that those who create a new household are younger and more able to adapt to economic changes (eventually induced by the arrival of refugees), you would underestimate the welfare of the original household by only comparing those who still belong to the original household in 1991 and 2004. In other words, a selection bias can be introduced if the split-off households have different characteristics compared with the members of their initial households (Witoelar, 2005). The sample used to compute table 4 includes the split-off households in a household fixed effect framework. To focus on the impact on the local hosts, we exclude, in a first analysis, the households that have migrated outside the region of Kagera between 1991 and 2004. Nevertheless, although they are excluded from the sample, we still need to take into account that these households have changed location. [21] Otherwise, we could introduce an important omitted variable bias. Migration is indeed known as an important household strategy to smooth consumption overtime (Rosenzweig and Stark, 1989; Stark, 1991) but could also affect other explanatory variables such as the choice of activity. Migration could affect the familly labour cost (e.g. in agricultural activities) or help to enter into some activities thanks to the role of financial intermediary that could be played by the migrant (Stark, 1980; Taylor and Wyatt, 1996). Table 4 introduces the proportion of the initial household members who have moved outside the region of Kagera (either in Tanzania or in Uganda) as well as the household decision to move outside the village of origin and its neighborhood within the Kagera region. Working with panel data, we complete these variables by the reported history of past migration of the household (proportion of migrants reported in 1991). Nevertheless, the migration variables could be endogenous due to simultaneity or omitted variable problem. As usual in migration studies, you need to clarify the causal link between the change in welfare and the migration decision. In other words, you do not know whether the decision to migrate is explained by the change in consumption or some unobserved characteristics (skills, motivations, etc) of the household or the opposite. Therefore, we use instrumental variables to deal with the endogeneity of the migration decisions. Similar to Scott Rozelle and deBrauw (1999), we construct an instrumental variable, capturing the spillover effect at the village level, i.e. the number of migrants by village but excluding the ones of the concerned household. We also compute the past value of this instrument, using 1991 data. Regarding the decision to migrate within the region of Kagera, we also use the proportion of the initial household migrating, excluding the concerned household to keep this instrument exogenous to the decision to move itself. Table 5 shows the first-stage fixed effect regressions. As expected, the measurements of migration spillovers positively affect our migration variables, with a highly significant F-test. The Anderson Correlation LR statistics rejects the null hypothesis of underidentification,

[21] Migration is likely to be an important phenomenon to control for. Collier et al. (1986) already stated that "In Tanzania migration is known to be a phenomenon with a long history which affects a very substantial part of the population." In our sample, we have about 608 individuals out of 3510 individuals who have moved outside their village (or a neighboring village). Interestingly, it is only slightly lower than the more restrictive (outside original village) ratio (27 %) obtained by Collier et al. (1986) with another sample and at another time.

while given the Stock and Yogo (2005) statistics, we can strongly reject the null hypothesis of weak instruments. Given the Hansen overidentification test, we cannot reject the null hypothesis that our instrumental variables are valid. [22]

Regression (1)-(4) of table 4 use the sample that includes the split-offs and the migration variables, without instrumental variables. We turn to regressions (5)-(8) that deal with the potential endogeneity of the migration variables. Regression (5) of table 4 at least indicates that, when split-off households are included and contrary to common wisdom, the establishment of refugee camps does not only release negative externalities on the local population. The impact is rather small but still is on average significant and positive. An increase by a standard deviation of the RI index would on average increase the consumption per adult equivalent by about only 1%. That is being said, our results also shed light on the two hypothesis that had been formulated in the previous sections. On the one hand, our theoretical framework suggests that the effects on the good and labour markets are likely to be unevenly distributed among the local population. Therefore, we introduce the activity variables defined in section 4.1. Very consistently throughout our various specifications, those involved in agricultural activities as self-employed farmers appear to be better-off by the refugee proximities, compared to their counter-parts in other areas. Among the winners, are also to be found the worker in non-agricultural activities and those depending on subsistence agriculture. For the worker in non-agricultural activities, such result is consistent with the evidence that skilled workers have been relativelly better valued by NGOs and international organizations in refugee-hosting areas.[23] One of the most negatively affected by the the refugee camps seems to be the agricultural worker, likely to face fiercer competition on the labour markets and increased prices for his purchases. An increase by a standard deviation of the RI index (at average level) would deteriorate the welfare of the agricultural worker by about 21-25% (see columns (3) and (4) of table 13). Despite the reported boom in this sector, our results suggest that those involved in business activities would experience a welfare deterioration following a similar change of the RI index of about 22-25% (see table 13). One possible explanation is the reported increase in competition in the business sector, following the arrival of more productive entrepreneurs from other regions (given the tracing nature of our data, newcomers are not included in our sample) who would have driven the existing local businesses out of the market. Although not treated specifically in the theoretical model, those involved in livestock are better-off. [24] Such positive effect

[22] Like for most 2SLS estimation, we had to partial out some variables to test the validity of our instruments. Our procedure was first to check if our coefficients and standard errors remain unaltered when partialling out the household-specific variables (size, proportion of literate members, etc) and then, correct the covariance matrix to make such tests feasible.

[23] Although we have not hypothesized any effect for this agent, the positive effect for those into subsistence agriculture might be more difficult to explain. One possibility might be that they would be in a better position to substitute more expensive consumption goods by their own production or that in the survey, the reported value of own production is higher given the expected higher prices for some goods.

[24] Those activities were merged with the business activities in previous works. The same results than the ones provided for the business activities were found.

is not only explained by the increased demand both from refugees and non-governmental organizations but mainly by a disruption of the supply of livestock at the time of the refugee arrival. In a mission identification report, the World Food Program (Program, 1995, 4) describes the evolution of the livestock sector in the two districts that have massively hosted refugees : "Livestock is widespread everywhere but only in some areas were large cattle established after independence with the arrival of Tutsis refugees from Rwanda, who brought their herds with them. In 1984, in Ngara, around 18,000 head of cattle were recorded and 128,000 in Karagwe ... As a result of refugees influx, about 50% of the cattle population in some major livestock keeping villages of Ngara district were taken out of the country when the Tutsis decided to return to Rwanda after the RPF government took power there, leaving Ngara and Karagwe with only about 8,000 and 65,000 cattle respectively."

Among the household control variables, the positive and negative coefficients of respectivelly the proportion of literate members within the household and the size of the household need to be underlined. Consistently with Beegle et al. (2007), table 4 also shows that the correlation between welfare and migration is only attributed to unobserved household characteristics that determine the migration decisions. On the other hand, our fieldwork helps us to hypothesize a non-monotonic relationship between the refugee impact and welfare. We found a U-shaped relationship between RI and the dependent variable. Nevertheless, it appears to be the reverse than the one that was expected. Our results suggest that the closer you are or the biggest the refugee population is, the more you benefit from it. By contrast, at some distant point or below a certain level of refugee population, you might suffer from the refugee presence. In the next section, we will discuss further how we could understand the roots of this non-monotonic relationship.

4.3 Discussion and robustness

4.3.1 Alternative specifications

We have imposed a particular structure to our baseline specification by adopting a household fixed effect specification and by restricting our sample to those household members who have not migrated outside the region of Kagera. First, we change our baseline specification by using an initial household fixed effect. Similar to Altonji et al. (1992) and Witoelar (2005), such initial household fixed effect will control for the unobserved characteristics that have been inherited from one's initial household. This might be particularly relevant in an environment of imperfect markets (credit, insurance, etc) where family ties, networks and family learning process might be particularly important for activity choice. Table 6 shows that our results remain robust to this specification. The only change is related to the livestock activities that the refugee presence does not seem to affect anymore. This could suggest that those entering into this type of activities without the necessary family support or connections might actually be less successful in benefiting from the market gap that has resulted from the sudden increase in demand and fall in supply. Second, one concern might be related to the way migratory flows would affect our results. On the one hand, a catastrophic view on the refugee crisis might suspect that the coefficient of the RI index is overestimated

because of an attrition problem. In other words, if local people tend to die (e.g. due to disease) or escape the refugee-hosting areas (without being traced), our results might only reflect the change of welfare of those better equipped to face such a shock. However, given the lower attrition and emigration rates depicted in chart 5, we can rule out this possibility. On the other hand, the opposite might also be true. The magnitude of our coefficients might be underestimated, as the mobility of some agents would tend to reduce the magnitude of the impact, notably through economic activities. Although section 1 gives clear reasons for the localized nature of the refugee impact, our estimates can then be considered as a lower-bound of the real impact. We also remind that in order to reduce this bias, we control for migration variables. As a robustness check, table 7 considers a sample that would include those who have migrated. In this case, we change the migration variables by introducing dummies indicating directly whether the household head has decided to migrate outside the region of Kagera. Table 7 confirms our main results.

Third, we have another point of concern with regards to our specification. We made an implicit assumption that the split-off households become independent from its initial household. A different assumption would be that the split-off households are altruist vis-à-vis their family and therefore, pool all their income and their consumption with their initial household. Altonji et al. (1992) and Witoelar (2005) propose a method to actually test the altruist hypothesis. Using an initial household fixed effect, an altruist household would result in a distribution of consumption per adult equivalent within the extended family independent of the distribution of income (net of remittances). Table 8 strongly rejects the null hypothesis of altruism among households of our sample, as the income net of transfers positively affects the consumption per adult equivalent when controlling for an initial household fixed effect and other time-varying household-caracteristics. This result holds, using different samples and different instrumental variables to deal with the potential endogenity of the income variable. [25]. Finally, table 9 tests whether our results capture long-lasting effects. Indeed, we use the fact that refugees from Burundi and Rwanda differ in their duration of stay in Tanzania. Refugees from Rwanda have indeed been repatriated in 1996 while the last refugees from Burundi left the region of Kagera in July 2008. Although we should remain cautious about the interpretation of this table [26], our most robust results suggest that the

[25] More information is given in the explanatory note of table 8. For comparability reasons, we first use in regressions (1) and (2) the instruments proposed by Witoelar (2005), i.e. the value of land, of farm productive equipments and of non-farm productive assets. However, we have some doubts about the exogenous nature of these instrumental variables. Said differently, we might suspect these variables to affect the consumption variable, by another channel than the income effect. Therefore, we introduce the distance through road networks to the main trading partners as an instrumental variables. The first-stage regressions are also provided in table 8. Such instrumental strategy is applied to both samples, excluding or including the households migrating outside the region of Kagera.

[26] We keep this distinction as a robustness checks, because we are likely to capture unobserved differences between refugees from Rwanda and Burundi. For example, the environmental degradation is reported to have been much fiercer in Rwandan refugee camps, while such degradation has been much more under control in more recent Burundian camps. On the contrary, refugees from Rwanda were reported to be much more skilled or hard-working.

effects of the refugee presence are long-lasting, mainly for the non-agricultural activities. One possible explanation for the sectoral distinction could be related to the fact that the effect on the agricultural sector is likely to fade away as soon as the refugees are repatriating. On the contrary, the effect on the non-agricultural sector may persist due to the fact that refugees continue to trade with their hosts, when returning to their country of origin (see section 2.2 on the trade links between refugees and local hosts after repatriation).

4.3.2 Interpretation : Scale versus spatial effect?

Based on our baseline estimation (table 4), we have already pointed that the establishment of refugee camps (more exactly, a positive change of a standard deviation at an average RI) has on average a positive and significant impact of less than 1% on welfare. This is a very small impact but we found a very differentiated effects on the local population. The main winners seem to be the skilled workers that enjoy increased job opportunities and the self-employed farmers who could benefit from a very cheap labour force. However, using the most robust results (columns (7) and (8) of table 4), columns (3) and (4) of table 13 qualify this finding. In net terms, the non-agricultural worker will benefit from the refugee presence, only if he is very closed to a refugee camp and if the refugee population is very large. At an average value of the RI index, its welfare will decrease by 7-9% following an increase by a standard deviation of this index while at the maximal value of the index (i.e. the closest village or the most populated refugee camp), he will improve his situation by a similar percentage (6-7 %). A similar qualification applies to another winner, the self-employed farmer. On the contrary, the main losers, i.e. the agricultural worker and those involved in businesses will experience a deterioration of their welfare (in a range of about 21-25% at the average RI index) but this negative effect will also be softened, the closer they are located to a refugee camp or the more numerous the refugee population is. Such qualified interpretation comes from the non-monotonic relationship found between the RI index and the welfare of the local population. Table 13 indicates that the minimum of the U-shaped relationship found between the RI index and the dependent variable is low but stands within the range of possible values. Independently of the effects through economic activities, a variation of the standard deviation of the RI index (at the average level) would decrease consumption per adult equivalent by about 8-10%, while at the closest and/or the most refugee-populated village, it could increase consumption by about 5-6%.

Despite this preliminary interpretation, the non-monotonic relationship between the RI index and the dependent variable does not necessarily receive easy interpretation. Indeed, one does not know whether the relationship is driven by its spatial (proximity to refugee camps) component or its scale (refugee population) component. One possible approach is to introduce alternative treatment variables to be able to decompose a scale effect from a spatial one. Tables 10 and 11 introduce four alternative variables of interest. Columns (1) to (3) of table 10 use the simple dummy variable indicating whether the community leader replies positively to the question "Were there any refugee settlement closed by?". Such variable is not significant and is likely to be very noisy given possible reported and measurement errors. Regressions (4) to (6) restrict the computation of the RI index for

each village to the closest refugee camp. Our previous results are confirmed with similar coefficients. Such simplification allows us to decompose the RI index for the closest camp into a scale and a spatial dimension. Table 11 presents the results. The spatial dimension seems to keep a similar non-monotonic relationship. However, the minimum stands largely above the range of possible values. On the contrary, the effect through economic activities, seems to be well captured by the spatial decay function. The opposite is true for the scale dimension. The non-monotonic relationship is confirmed, while the effects through economic activities are less precisely captured. The next step is then to introduce simultaneously the two dimensions, like in table 12. We have a confirmation that it is the scale dimension which is driving the non-monotonic shape. Such a result might seem surprising but this suggests that a minimum mass of refugees is needed to generate positive externalities that would benefit the local population (independently of the main economic occupations). Looking at the minimum of the U-shaped relationship, table 13 indicates that this minimum mass could amount to about 53,000 to 57,000 refugees. Such a figure should be considered with a great degree of cautiousness. For us, it does not contradict the UNHCR recommendations to its own staff that in general, "large camps of over 20,000 people should generally be avoided" (UNHCR, 1982) because of possible problems of accommodation, security, etc. We indeed do not consider these internal constraints to the size of a refugee camp. However, our results could suggest that in this approximation of the optimal size, UNHCR should internalize the effect of the size on the local hosts, which is not necessarily as expected. A smaller size is not necessarily optimal. For the activity variables, the scale dimension is relevant for the agricultural worker who could face fiercer competition on the labour markets for larger refugee population and for the self-employed farmer who could benefit from a cheaper workforce with a large refugee population. On the contrary, the spatial dimension per se only captures negative externalities: the closer you are, the worst it should be. Contrary to what was hypothesized following our fieldwork, there is no non-monotonic relationship between this dimension and the dependent variable, at least in the long run. The spatial dimension seems however to explain very much the effects on the non-agricultural sector.

4.3.3 Can a refugee camp generate agglomeration economies?

The existence of a scale effect following the establishment of a refugee camp might seem puzzling. Given the fact that we control for the effects through household economic activities, the scale effect should capture externalities that materialize outside the individual economic units, i.e. what is called Marshallian agglomeration economies (Combes et al., 2008c). Since Marshall (1890), regional scientists and geographical economists know that the spatial concentration of economic activities could generate positive externalities through a denser and more efficient labour market (labour pooling), the share of specialized inputs through forward and backward linkages (Hirschman, 1958) and the spread of innovation through spillover effects. Our results suggest that the concentration of economic activities around the refugee camps may have generated such agglomeration economies, provided a sufficient mass of refugees was gathered. Such Marshallian economies constitute some kind of theoretical black boxes (Combes et al., 2008c). However, as shown by Combes et al. (2008a) and

Combes et al. (2008b), the existence of these agglomeration economies can be empirically investigated by computing the following variables at the village level:

- The population density should proxy for the size of the labour market and the intensity of local interactions. [27] Such variable could capture the increased interactions between households, resulting from the renewed dynamism of markets surrounding the establishment of a refugee camp. We expect such phenomenon to increase with the refugee population.
- The occupational diversity is computed as the inverse of the Herfindhal index (standard measure of concentration).

$$div_{v,t} = \frac{1}{H_{v,t}} = \frac{1}{\sum_o (spe_{v,o,t})^2} \tag{18}$$

  Such diversity is likely to ease innovation between sectors but also to allow the local population to become more resilient to negative shocks specific to one occupation (Jacobs, 1984; Glaeser et al., 1992). Given credit and insurance market imperfections, this last effect could be an important welfare-improving channel for households facing major constraints to diversify their risks and smooth consumption overtime (Morduch, 1995; Rosenzweig and Binswanger, 1993; Rosenzweig and Wolpin, 1993).

- The levels of "Specialisation" are included to control for localization economies. We compute it by main occupation (a in subscript) as the total income generated from one occupation over the total income.

$$spe_{a,v,h,t} = \frac{\sum_h inc_{a,v,h,t} - inc_{a,h,t}}{\sum_h \sum_a inc_{a,v,h,t} - inc_{a,h,t}} \tag{19}$$

  To keep it completely exogenous to the concerned household, we substract the concerned household's income from the computation (it becomes then a village-level variable with a household-specific correction). Such specialisation index should capture externalities within the occupations and has been called localization economies (Hoover, 1936, , cited in Combes et al 2008a) or Marshall-Arrow-Romer economies by Glaeser et al. (1992). The concentration of economic activities belonging to the same sector should ease imitation and diffusion of innovations and new production processes as well as the sharing of specific inputs. However, increasing the spatial concentration of economic activities within the same sector intensifies labour market and product market competition. Although the net effect on welfare is a priori ambiguous, we could expect this second effect to matter more in refugee-hosting areas for some sectors such as the business one.

[27] Such a variable is usually captured by the local employment density. Employment has a different meaning in an underdeveloped region such as the Kagera region, as it is mainly an activity of last resort (therefore, it represents a small share of occupations). Still, we would like to capture the increased interactions between households in a more active market that could be one welfare-improvement channel of refugees.

- Other factors need to be controlled for. This scale effect could also capture the effect of the access to other markets. We will introduce a market potential index based on the average income per household computed by village. We use a market potential similar to Harris (1954), i.e. $PM_{v=r} = \sum_{v=s=1}^{51} \frac{Y_s}{d_{r,s}}$. Traditionally, it is also needed to control for local amenities or local public goods. We introduce two variables, measuring the accessibility (distance) to the closest health center and to the closest secondary school. On the one hand, section 2.2 has underlined the increased accessibility to higher-quality health services. On the other hand, the reported increase in tax capacities of refugee-hosting areas due to the revived economic dynamism could affect the provision of public goods such as education.

First of all, we assess whether the inclusion of these variables radically alter our previous results. In other words, the question is to know whether the found scale effect persists when controlling for agglomeration economies. [28] Regression (1) of tables 14 and 15 introduces our main variable of interest, the population density in a Household Fixed Effect specification with standard errors clustered at the district level. Regressions (2) and (3) introduce then our main control variables. Despite the loss of efficiency following the inclusion of additional variables, the scale effect remains robust. With the exception of the least robust regression (1), the population density has a positive and significant coefficient. Among the

[28] Not finding similar results would jeopardize our identification strategy. It could indeed mean that we would infer a significant impact of the presence of refugees from possibly unrelated agglomeration economies. This would also raise some doubts about the exogenous nature of our treatment variable. It has to be noted that Baez (2007, 12) points to such an endogeneity problem. According to him, "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." Although we consider his concerns as legitimate, we do not share his view. First of all, controlling the border was not at all something in the hands of the local authorities. The refugee inflow was so massive that at that time, it was before all a security issue and borders were rather of the military resort. The choice of location was mainly taken by the Ministry of Home Affairs and UNHCR. Among the criteria reported to have been used, there was an important cost issue . The refugee inflows was the largest in UNHCR history and highly unanticipated. From October 1993 to April 1994, refugees crossed the border between Burundi and Tanzania and stay closed to local communities, without formal assistance. There were so numerous that it was reported to be too costly to move them far away from the border. Therefore, contrary to the UNHCR handbook of Emergencies and international law recommendations, refugee camps were located pretty closed to the border. So, if there was a choice of location to be made, this choice was restricted in the area closed to the border. In addition, as confirmed by officials, this was reinforced by the willingness of the Tanzanian government to ease the repatriation process and reduce as quick as possible the risk to create a small Rwandan or Burundian conflict within the Tanzanian border. Such geographic restrictions on the choice of location certainly reduce the endogeneity problem. Our identification strategy indeed rests on the fact that there was no choice of location between far away and closed villages.

control variables, only the distance to the closest health center has a negative and significant sign. More interestingly, we then seek to assess whether the interactions between these new variables and the refugee population could explain the previously found scale effect. Given their spatial nature, the market potential and the access to health services are interacted with the proximity to the closest refugee camp. Regressions (4) and (5) indicate that the scale effect fades away by the introduction of these interaction terms. In refugee-hosting areas, the denser labour market and the more intense local interactions have significantly improved the welfare of the local population. On the contrary, the diversity index does not bear any explanatory power, specific to refugee-hosting areas. [29] Access to health services has improved welfare, but even more in areas closed enough to refugee camps. A positive and significant variable is found for the distance to the closest secondary school. It cannot receive any meaningful interpretation, as the education systems provided to the refugees and the local population are completely segmented. [30] Still, we are well aware of the potential endogeneity of the population density. The inclusion of several variables such as the local amenities and the migration variables might not be enough to deal with the potential endogeneity issue (Combes et al., 2008a). An increase in welfare could actually attract people from other regions. Nevertheless, despite this simultaneity problem, it remains interesting to witness that agglomeration forces captured by the population density seem to have been more at work in refugee-hosting areas. On the other hand, we adopt an instrumental variable strategy in regressions (6) and (7), by instrumenting the population density by the 10-year lagged population density. Our previous results are confirmed. Finally, the purpose was to investigate further the found scale effect. Despite this welfare-improving channel through agglomeration economies, our results do not reject the existence of negative externalities (environmental degradation, security issues, etc). Although we have shown that the average effect is marginally positive, when we control for the effect on economic activities (that remain very similar), for the agglomeration forces, including the ones generated by the establishment of refugee camps, the scale variable becomes negative and remains significant.

5 Conclusions

Following an iterative fieldwork that helps us to clarify some hypothesis, the paper sheds some light on the impact of refugee inflows on their local hosts. Contrary to conventional wisdom, we found on average a slightly positive impact on local people's welfare. However, such an impact appears to be highly differentiated among the refugee-hosting population. The agricultural worker suffered from fiercer competition on the labour markets and increased prices on the good markets in refugee-hosting areas. On the policy side, programs (e.g. education, microfinance, etc) might target those agricultural workers to help them to cope with such a shock on the local economy. On the contrary, those involved in agricul-

[29] Note that the same is true when we interact this variable with the spatial dimension.

[30] The amenity variables are only introduced in this section on the discussion of the results, as they may be threatened by an important endogeneity bias. For example, school creation targeted to poorer villages could explain the unexpected sign for the school variable.

tural activities as self-employed farmers, have benefited from the supply of a cheap labour force while the worker in the non-agricultural sector tends to have improved their situation (as skills have been highly valued by NGOs and international organizations). Contrary to our theoretical prediction, the self-employed in business activities does not seem to have improved its welfare. Such results contrast with the business boom reported by interview respondents during our fieldwork. Possible reconciliations might call for integrating an increase in competition (with selection effect) into our analysis. Understanding the factors affecting the entry and exit into such activities is certainly a subject for further works and could path the way for more policy-oriented conclusions aiming at easing the adoption of coping strategies by the local population following such a refugee inflow.

Furthermore, our empirical analysis rejects the hypothesized non-monotonic relationship according to which negative externalities (environmental degradation, security issues, disease spread, etc) would overcome the economic benefits in villages much more affected by the refugee presence. On the contrary, we found the opposite relationship : the closer you are or the more populated the refugee camp is, the more likely you could benefit from positive external effects (even when controlling for the impact on economic activities). By investigating further the roots of this U-shaped relationship between the RI index and the consumption per adult equivalent, we found that a minimum mass of refugees could be needed for these positive externalities to materialize. Our discussion of the results suggest that such a scale effect could be explained by the presence of agglomeration economies (driven by a denser labour market and increase in local interactions) around the refugee camps. Such findings could call for a radical change of paradigm regarding the impact of refugees on the local population. By minimizing the negative externalities whose existence is also confirmed by our analysis, improving the capabilities of the local people to cope with such a structural change and eventually, by integrating progressively these refugees into the local economy, the presence of refugees could actually constitute a unique asset (rather than a burden) to break down some traps of underdevelopment.

References

Adams, M., Samuel, V., Sekiku, J., and Tibandebage, F. (1995). The impact of the rwandese influx on tanzanian villages. karagwe district, kagera region. tanzania. ACORD (Agency for Cooperation and Research in Development) Report of a Field Visit 20 August - 7 September 1995.

AlixGarcia, J. and Saah, D. (2008). The effect of refugee inflows on host communities: Evidence from tanzania. Working Paper.

Altonji, J. G., Hayashi, F., and Kotlikoff, L. J. (1992). Is the extended family altruistically linked? direct tests using micro data. *The American Economic Review*, 82(5):1177–1198.

Azam, J.-P. and Hoeffler, A. (2002). Violence against civilians in civil wars : Looting or terror. *Journal of Peace Research*, 39(4):461–485.

Baez, J. E. (2007). Civil wars beyond their borders: The human capital and health consequences of hosting refugees. Job market paper.

Barret, C. B. (2002). Food aid and commercial food trade. *OECD Background paper*.

Beegle, K., DeWeerdt, J., and Dercon, S. (2006). Kagera health and development survey 2004 basic information document. *mimeo. The World Bank*.

Beegle, K., DeWeerdt, J., and Dercon, S. (2007). Poverty and wealth dynamics in tanzania: Evidence from a tracking survey. In Narayan, D. and Petesch, P., editors, *Moving out of Poverty*, page forthcoming in volume 2. Palgrave MacMillan and The World Bank.

Berry, L. (2008). The impact of environmental degradation on refugee-host relations: a case study from tanzania. *UNHCR EPAU Research paper*, 151.

Bertrand, M., Duflo, E., and Mullainathan, S. (2003). How much should we trust differences-in-differences estimates. *mimeo*.

Bubb, R., Kremer, M., and Levine, D. (2007). The economics of international refugee law. *mimeo*.

card, D. (1990). The impact of the mariel boatlift on the miami labour markets. *Industrial and Labor Relations Review*, 43(2):245–257.

Chambers, R. (1986). Hidden losers? the impact of rural refugees and refugee programs on poorer hosts. *International Migration Review*, 20(2):245–263. Special issue: Refugees: Issues and directions.

Collier, P., Radwan, P., Wangwe, S., and Wangwe, A. (1986). *Labour and poverty in Rural Tanzania*. Oxford University Press and ILO, Oxford.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008a). Spatial wage disparities: sorting matters. *Journal of Urban Economics*, pages 723–742.

Combes, P.-P., Lafourcade, M., Thisse, J.-F., and Toutain, J.-C. (2008b). The rise and fall of spatial inequalities in france: A long-run perspective. CEPR Discussion paper DP7017.

Combes, P.-P., Mayer, T., and Thisse, J.-F. (2008c). *Economic Geography: The Integration of Regions and Nations*. Princeton University Press.

Crisp, J. (2003). No solution in sight : The problem of protracted refugee situations in africa. Center for Comparative Immigration Studies Working Paper 68.

CSFM (2003). The impact of the presence of refugees in northwestern tanzania. Center for the Study of Forced Migration. University of Dar es Salaam.

de Montclos, M.-A. P. and Kagwanja, P. M. (2000). Refugee camps or cities? the socio-economic dynamics of the dadaab and kakuma camps in northern kenya. *Journal of Refugee Studies*, 13(2):205–222.

deMel, S., McKenzie, D. J., and Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of Development Economics*, 88(1):19–31.

Dercon, S. and Krishnan, P. (2004). Food aid and informal insurance. *World Bank Policy Research Working Paper*, 3217.

Dercon, S. and Shapiro, J. S. (2007). Moving on, staying behind, getting lost: Lessons on poverty. mobility from longitudinal data. In Narayan, D. and Petesch, P., editors, *Moving out of Poverty*, volume 1, chapter 3, pages 77–126. Palgrave MacMillan and The World Bank.

DeWeerdt, J. (2006). Moving out of poverty in tanzania's kagera region. *EDI Working Paper*.

Donovan, C., McGlinchy, M., Staatz, J., and Tschirley, D. (2006). Emergency needs assessments and the impact of food aid on local markets. *MSU International Development Working Paper*, 87.

Glaeser, E. L., Kallal, E., Sheinkman, J., and Shleiffer, A. (1992). Growth in cities. *Journal of Political Economy*, 100:1126–1152.

Harris, C. (1954). The market as a factor in the localization of industry in the united states. *Annals of the Association of American Geographers*, 64:315–348.

Hatton, T. J. and Williamson, J. G. (2004). Refugees, asylum seekers and policy in europe. NBER Working Paper 10680.

Hirschman, A. (1958). *The strategy of economic development*. New Haven, CT: Yale University.

Hoover, E. (1936). *Location Theory and the Shoe and Leather Industries*. Cambridge, MA: Harvard University Press.

Hunt, J. (1992). The impact of the 1962 repatriates from algeria on the frensh labor market. *Industrial and Labor Relations Review*, 45(3):556–572.

Jacobs, J. (1984). *Cities and the Wealth of Nations*. New York, Random House.

KHDS (2004). User's guide to the kagera health and development survey datasets. Development Research Group. The World Bank.

Kuhlman, T. (2002). Responding to protracted refugee situations. *A case study of Liberian refugees in Côte d'Ivoire*. UNHCR Evaluation and Policy Analysis Unit.

Landau, L. B. (2004). Challenge without transformation: refugees, aid and trade in western tanzania. *Journal of Modern African Studies*, 42(1):31–59.

Marshall, A. (1890). *Principles of Economics*. London. Macmillan. Edition 8 Published in 1920.

Morduch, J. (1995). Income smoothing and consumption smoothing. *Journal of Economic Perspectives*, 9(3):103–114.

Murata, Y. (2008). Engel's law, petty's law and agglomeration. *Journal of Development economics*, 87:161–177.

Murphy, K. M., Shleifer, A., and Vishny, R. (1989). Income distribution, market size and industrialization. *Quarterly Journal of Economics*, 104:537–564.

NBS (2003). Kagera region socio-economic profile. National Bureau of Statistics of Tanzania and Kagera Regional Commissioner. dar es Salaam.

Program, W. F. (1995). Refugee affected areas development project. identification/formulation mission report. World Food Program Report No: 85/95 TCP-URT 58 10 August 1995.

Rosenzweig, M. and Binswanger, H. (1993). Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal*, 101(2):56–78.

Rosenzweig, M. and Wolpin, K. (1993). Credit market constraints, consumption smoothing and the accumulation of durable production assets in low-income countries: investments in bullocks in india. *Journal of Political Economy*, 101(2):223–244.

Rosenzweig, M. R. and Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural india. *Journal of Political Economy*, 97(41):905–926.

Rutinwa, B. (2002). The end of asylum? the changing nature of refugee policies in africa. *Refugee Survey Quarterly*, 21(1-2).

Scott Rozelle, J. E. T. and deBrauw, A. (1999). Migration, remittances, and agricultural productivity in china. *American Economic Review*, 89(2):287–291.

Slaughter, A. and Crisp, J. (2009). A surrogate state? the role of unhcr in protracted refugee situations. UNHCR EPAU Research Paper 108.

Stark, O. (1980). On the role of urban-to-rural remittances in rural development. *Journal of Development Studies*, 16:369–374.

Stark, O. (1991). *The migration of labor*. Cambridge (USA): Basil Blackwell.

Stark, O. (2004). On the economics of refugee flows. *Review of Development Economics*, 8(2):325–329.

Stock, J. and Yogo, J. H. (2005). Testing for weak instruments in iv regression. In Stock, J. H., editor, *Identification and Inference for Econometrics Models: A Festschrift in Honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press.

Taylor, J. and Wyatt, T. (1996). The shadow value of migrant remittance, income and inequality in a household farm economy. *Journal of Development Studies*, 32(6):899–912.

TheEconomist (2001). The penalty of kindness. The Economist, 23 August 2001.

Udry, C. (2003). Fieldwork, economic theory and research on institutions in developing countries. *unpublished*.

UNHCR (1982). Handbook for emergencies. UNHCR memo.

UNHCR (2007). 2007 global trends. refugees, asylum-seekers, returnees, internally displaced and stateless persons. United Nations High Commission for Refugees.

Verwimp, P. and Bundervoet, T. (2008). Consumption growth, household splits and civil war. HiCN-Households in Conflict Network Working Paper 48.

Waters, T. (1999). Assessing the impact of the rwanda refugee crisis on development planning in rural tanzania 1994-1996. *Human Organization*, 58(2):142–152.

Werker, E. (2007). Refugee camp economies. *Journal of Refugee Studies*, 20(3):461–480.

WFP and UNHCR (1998). Household food economy assessment greater lukole camp. ngara district tanzania. World Food Program/United Nations High Commission for Refugees Joint assessment 10 July- 12 August 1998.

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

Winkelmann, L. and Winkelmann, R. (1998). Why are the unemployed so happy? evidence from panel data. *Economica*, pages 1–15.

Witoelar, F. (2005). Inter-household allocations with extended families: Evidence from the indonesia family life survey. *Economic Growth Center Discussion Paper. Yale University*, 912.

Figure 1: Change in consumption per adult equivalent, by initial poverty groups

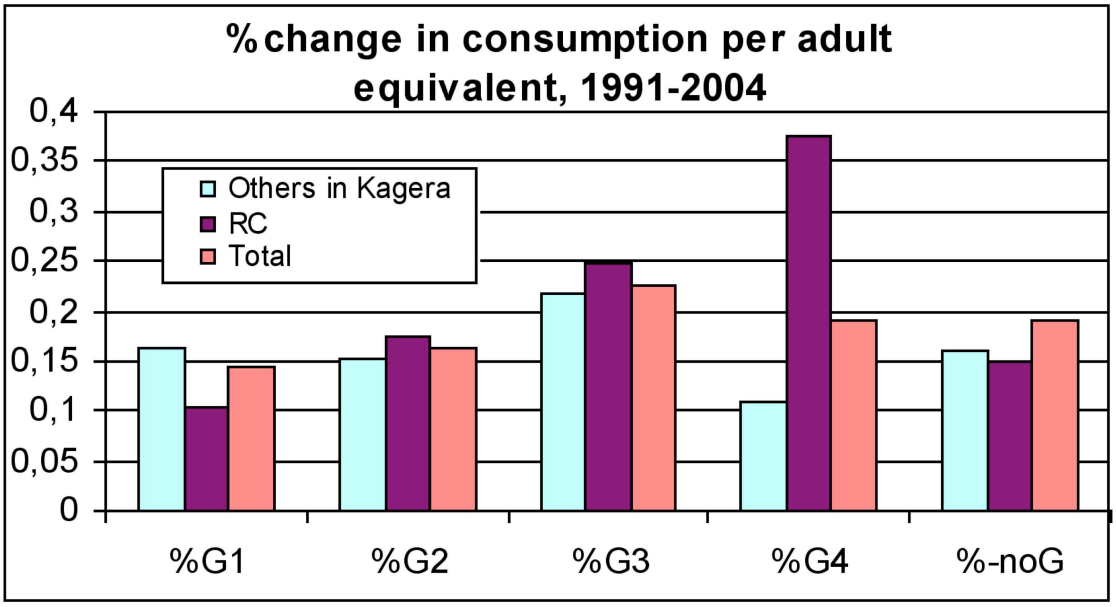

Figure 2: A refugee camp is not only a dispersion force

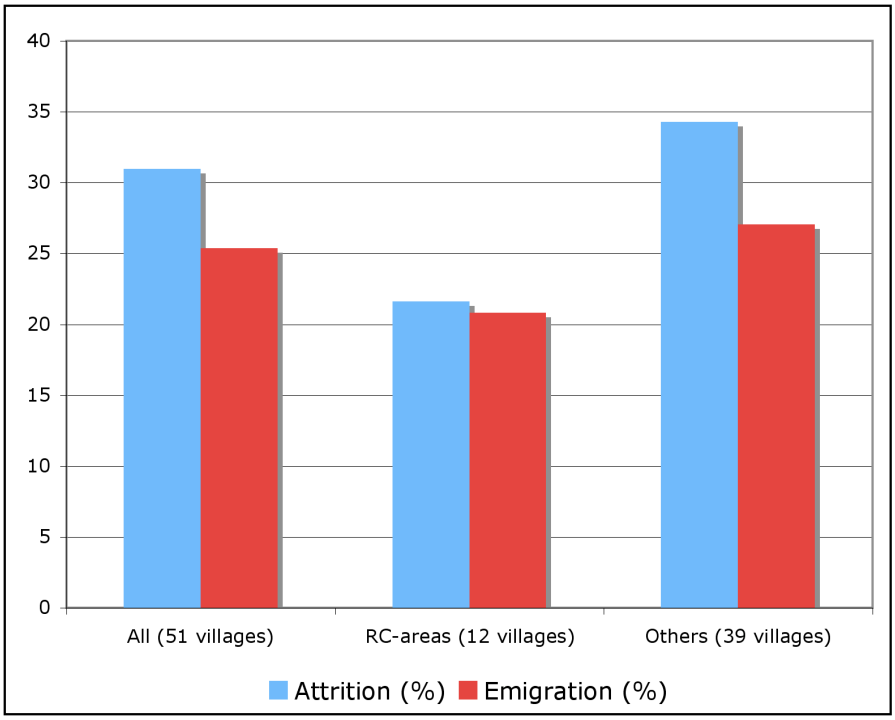

Figure 3: Own calculation based on KHDS

Table 1: Descriptive statistics for panel data (village-specific data)

RI	4252	2.970049	24.95085	0	261.4514
RI(Burundi)	4252	2.970049	24.95085	0	261.4514
RI(Rwanda)	4252	2.79e-07	3.11e-06	0	.0000349
RC	4252	.1298213	.3361462	0	1
RI(closest)	4252	2.897639	24.2017	0	253.3578
$exp(d^{-1}_{v,closest})$	4252	.0000304	.0002136	0	.0021113
Population(ref)	4252	7299.906	22951.92	0	120000
Log(density)	3832	4.577739	1.080623	3.135628	7.244486
Log(diversity)	3940	.859769	.3150989	.1850301	2.060453
Log(MP)	3940	12.98923	.7953359	11.53168	14.67735
Spe (Subs.)	3940	.548723	.1903971	.028286	.8671061
Spe (Sales)	3940	.1132917	.1618308	0	.9093747
Spe (Non Agri.L)	3940	.0651852	.104731	0	.6156561
Spe (Agri. L)	3940	.0395378	.0593182	0	.3065133
Spe (Bus)	3940	.1544627	.1513967	0	.7504783
Spe (LVST)	3940	.0079959	.0125649	0	.0803695
$\text{Log(density)}_{t-10}$	3832	4.261956	1.062095	2.90129	6.686915
Log(School)	3844	2.05474	1.081159	0	4.394449
Log(health)	3543	2.348483	1.179115	0	4.60517

RC is a dummy variable indicating whether the community leader has replied positively to the question "Were there any refugee settlements closed by?"; RI(closest) applied the same average of the refugee population weighted by an exponential distance function; $d^{-1}_{v,closest}$ is the inverse exponential distance function between each village and the closest refugee camp; Refugees(closest) is the number of refugees in the closest refugee camp.

Table 2: Descriptive statistics

Variables	Obs	Mean	Std. Dev.	Min	Max
Cons(PAE, 1991)	2046	217646.7	144837	20007.14	1491798
Cons (PAE, 2004)	2064	306802.9	243458.3	23775.18	2758511
Cons(per cap.,1991)	2040	188035.4	153851.9	14951.97	1256506
Cons(per cap.,2004)	2067	248641.5	247665.9	16969.46	4771619
Cons(PAE,t)	4110	262420	205393.8	20007.14	2758511
log(cons(PAE,t))	4110	12.27243	.6169704	9.903845	14.8302
Size(1991)	2120	6.898113	3.702479	1	30
Size(2004)	2064	4.768895	2.581935	1	22
Log(land,1991)	2126	11.66682	2.4293	0	17.92384
Log(land,2004)	2126	9.875798	4.652381	0	17.76653
Prop. Lit. (1991)	2126	.5248823	.2479797	0	1
Prop. Lit (2004)	2126	.5695399	.2942705	0	1
Subs.(1991)	2126	.5865475	.4925684	0	1
Subs.(2004)	2126	.7168391	.4506399	0	1
Crop sales(1991)	2126	.3222013	.4674296	0	1
Crop sales(2004)	2126	.0456256	.2087208	0	1
Agri. L(1991)	2126	.2121355	.4089164	0	1
Agri. L(2004)	2126	.0898401	.2860198	0	1
NonAgri.L(1991)	2126	.0277516	.1642991	0	1
NonAgri.L(2004)	2126	.1745061	.3796334	0	1
Business(1991)	2126	.1853246	.388652	0	1
Business(2004)	2126	.2196613	.4141146	0	1
LVST(1991)	2126	.2878645	.4528742	0	1
LVST(2004)	2126	.1444026	.3515802	0	1
Migr(kag)	2126	.173095	.3784184	0	1
Migr(out)	2126	.0733772	.2608161	0	1
Prop.migr.(Kag)	2126	.1398416	.2245621	0	1
Prop.migr.(out)	2126	.0585224	.1480509	0	1
Prop. Migr.(Kag, corrected)	2126	.1031855	.1940602	0	1
Prop. Migr.(out, corrected)	2126	.0435935	.1230459	0	1
Past prop.Migr.(Kag)	2126	.0332681	.0976359	0	.8181818
Past prop.Migr(out)	2126	.1262255	.1920051	0	.9166667
Spillover(out, 2004)	2126	3.786924	4.797376	0	23
Spillover(out, 1991)	2126	14.98542	13.73207	0	53
Spillover(kag, 2004)	2126	11.76999	5.782917	1	32
Spillover(kag, 1991)	2126	2.936971	3.144336	0	13

Activity dummies respectively indicate whether one household generates in a greater proportion than the sample average its income from subsistence agriculture (Subs.), self-employed agricultural activities (Crop Sales), employed agricultural worker (Agri. L), employed non-agricultural worker (Nonagri. L) and self-employed business (Business) or livestock (LVST) activities ; while other variables represent the size of the household (Size), the value of land transformed in real terms and in per adult equivalent (Land), the proportion of literate people in the household (Prop. Lit.). Prop. Migr(Kag) and Prop. Migr(out) respectively indicate the proportion of individuals within the initial household who have migrated outside their original village or its neighborhood in the region of Kagera (Kag) and outside the region of Kagera (out). Migr.(out) and Migr.(kag) respectively indicate whether the household head has migrated outside their original village or its neighborhood in the region of Kagera (Kag) and outside the region of Kagera (out). Spillovers represent the migration spillovers at the village level while HH Prop. Migr. (corrected), the number of household members migrating (excluding the one for which the index is computed).

Table 3: Household Fixed Effect (excluding split-off households)

	(1) HHFE	(2) HHFE	(3) HHFE	(4) HHFE
	split. excl.	split. excl.	split. excl.	split. excl.
RI	0.000208	-0.0453***	-0.0518***	-0.0607***
	(0.000264)	(0.00978)	(0.0105)	(0.0108)
RI^2			2.49e-05***	6.90e-05***
			(5.69e-06)	(1.07e-05)
Subsistence		0.00716	0.00217	-0.0157
		(0.0832)	(0.0821)	(0.0689)
Crop Sales		-0.0745	-0.0734	-0.0669
		(0.186)	(0.187)	(0.162)
Agri		0.141	0.134	0.174
L		(0.125)	(0.124)	(0.137)
Non Agri		0.198	0.195	0.110
L		(0.131)	(0.129)	(0.243)
Business		0.308**	0.305**	0.309*
		(0.113)	(0.111)	(0.145)
LVST		0.0793	0.0766	0.0797
		(0.216)	(0.217)	(0.203)
Subs.		0.0460***	0.0461***	0.0434***
*RI		(0.00938)	(0.00943)	(0.00792)
Sales		7.817	7.532	12.53
*RI		(25.13)	(25.28)	(31.33)
NonAgri.		0.0463***	0.0465***	0.0441***
L*RI		(0.00981)	(0.00986)	(0.00869)
Agri		-0.0337*	-0.0278	-0.0138
L*RI		(0.0148)	(0.0145)	(0.0151)
Business		-0.00707[0.156]	-0.00164	0.00515
*RI		(0.00423)	(0.00426)	(0.00459)
LVST		-0.00115	-0.00125	-0.000902
*RI		(0.00165)	(0.00165)	(0.00163)
Log(land)	-0.000438	0.00110	0.000654	0.00474
	(0.00729)	(0.00695)	(0.00698)	(0.0125)
Prop.Lit.	0.179	0.147	0.149	0.0959
	(0.178)	(0.200)	(0.200)	(0.191)
Size	-0.0717***	-0.0789***	-0.0790***	-0.0771**
	(0.0145)	(0.0173)	(0.0173)	(0.0231)
α_t	0.0494	0.0155	0.0200	-0.0320
	(0.0405)	(0.0851)	(0.0841)	(0.108)
Cluster D				sign.
Constant	12.52***	12.47***	12.48***	12.03***
	(0.189)	(0.122)	(0.115)	(0.228)
Obs.	979	979	979	979
R^2	0.131	0.187	0.189	0.343

*Note: *** p< 0.01, ** p< 0.05, * p< 0.1; Robust standard errors in parentheses and clustered at the district level (allowing for correlation of errors within districts). Between brackets, are included the p-value of coefficients closed to significance.*

Table 4: Household Fixed Effect (excl. migrants and incl. split-offs)

	(1) HHFE	(2) HHFE	(3) HHFE	(4) HHFE	(5) HHFE (2SLS)	(6) HHFE (2SLS)	(7) HHFE (2SLS)	(8) HHFE (2SLS)
	split. incl	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.
RI	0.0005	-0.004***	-0.007***	-0.008**	0.0005***	-0.0045***	-0.007***	-0.008***
	(0.0003)	(0.0009)	(0.0017)	(0.003)	(0.0002)	(0.0008)	(0.0014)	(0.002)
	[0.177]							
RI^2			1.11e-05*	1.35e-05			1.01e-05**	1.23e-05**
			(5.24e-06)	(8.19e-06)			(4.41e-06)	(6.07e-06)
				[0.160]				
Subs.		-0.128**	-0.131**	-0.122**		-0.126***	-0.128***	-0.120***
		(0.0495)	(0.0485)	(0.0459)		(0.0325)	(0.0325)	(0.0297)
Crop Sales		-0.0606	-0.0624	-0.0571		-0.0583	-0.0594	-0.0502
		(0.0566)	(0.0556)	(0.0580)		(0.0364)	(0.0361)	(0.0397)
Non Agri		0.135*	0.133*	0.0981		0.142***	0.141***	0.111***
L		(0.0527)	(0.0519)	(0.0607)		(0.0314)	(0.0315)	(0.0304)
Agri		0.107	0.105	0.100		0.109**	0.108**	0.105**
L		(0.0871)	(0.0863)	(0.0902)		(0.0485)	(0.0483)	(0.0506)
Business		0.195**	0.194**	0.187*		0.212***	0.212***	0.211***
		(0.0750)	(0.0751)	(0.0730)		(0.0731)	(0.0728)	(0.0675)
LVST		0.0102	0.00981	0.0139		0.0102	0.0102	0.0153
		(0.0935)	(0.0937)	(0.0940)		(0.0585)	(0.0585)	(0.0564)
Subs.		0.0049***	0.0045***	0.0051***		0.005***	0.0046***	0.0051***
*RI		(0.0009)	(0.0008)	(0.0008)		(0.0008)	(0.0008)	(0.0007)
Sales		0.0018**	0.0019**	0.002***		0.0018***	0.0019***	0.0019***
*RI		(0.0005)	(0.0005)	(0.0004)		(0.0006)	(0.0006)	(0.0006)
Non Agri		0.004***	0.0039***	0.0044***		0.0041***	0.0039***	0.0044***
L*RI		(0.0005)	(0.0005)	(0.0004)		(0.0005)	(0.0005)	(0.0005)
Agri		-0.0018**	-0.0017**	-0.0018**		-0.0017***	-0.0017***	-0.0018***
L*RI		(0.0005)	(0.0005)	(0.0005)		(0.0003)	(0.0003)	(0.0003)
Business		-0.0021***	-0.002***	-0.0019***		-0.0021***	-0.002***	-0.002***
*RI		(0.0003)	(0.0003)	(0.0004)		(0.0002)	(0.0003)	(0.0003)
LVST		0.0008**	0.0008**	0.0008*		0.0009***	0.0009***	0.0009**
*RI		(0.0003)	(0.0003)	(0.0004)		(0.0003)	(0.0003)	(0.0004)
Log(land)	-0.005	-0.0004	-0.0005	0.003	-0.004	0.0006	0.0005	0.003
	(0.007)	(0.008)	(0.008)	(0.0075)	(0.005)	(0.0056)	(0.0057)	(0.004)
Prop.Lit.	0.183	0.157	0.156	0.142	0.180***	0.153**	0.151**	0.137**
	(0.109)	(0.111)	(0.111)	(0.0893)	(0.0693)	(0.0690)	(0.0693)	(0.0556)
Size	-0.0386	-0.0402*	-0.0401*	-0.0412*	-0.0379***	-0.0398***	-0.0397***	-0.0409***
	(0.0206)	(0.0187)	(0.0187)	(0.0176)	(0.0124)	(0.0112)	(0.0112)	(0.0102)
Split-off	0.121	0.113	0.113	0.105	0.135***	0.119***	0.117***	0.109***
	(0.0826)	(0.0745)	(0.0739)	(0.0793)	(0.0485)	(0.0432)	(0.0423)	(0.0410)
Migr.(Kag)	0.151	0.226*	0.225*	0.0719	0.272	0.335	0.332	0.300
	(0.0886)	(0.0984)	(0.0989)	(0.0927)	(0.271)	(0.278)	(0.279)	(0.422)
Prop.	0.330	0.355	0.357*	0.357*	0.187	0.329	0.352	0.445
Migr.(out)	(0.186)	(0.177)	(0.176)	(0.167)	(0.392)	(0.383)	(0.385)	(0.417)
α_t	0.0712	0.0550	0.0590	0.0551	0.0298	0.0286	0.0353	0.0508
	(0.0479)	(0.0574)	(0.0543)	(0.0654)	(0.0681)	(0.0811)	(0.0829)	(0.0514)
Cluster D				sign.				sign.
Obs.	3803	3803	3803	3803	3674	3674	3674	3674
R^2	0.185	0.235	0.235	0.294	0.181	0.232	0.233	0.287
F-Test	201.78***	4.16***	4.03***	18.53***	53.65***	4.83***	4.81**	4.43*
Anderson					147.68***	143.48***	143.66***	90.82***
Stock-Yogo					51.09	49.18	49.21	29.72
Hansen[a]					0.124	0.124	9.298	
(p-value)					(0.68)	(0.72)	(0.0023)	

*Note: *** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance. Robust standard errors in parentheses and clustered at the district level. a The overidentification test was obtained by partialling out the household-specific variables. We first check that due to the Frisch-Wargh-Lovell theorem, the coefficients of the remaining variables are unaltered. We then correct the covariance matrix to make the overidentification test feasible.*

Table 5: Household Fixed Effect (excl. migrants): First stage regression

	HHFE	HHFE	HHFE	HHFE
	(1)	(2)	(3)	(4)
Dep Var	Prop. Migr.(out)	Prop. Migr.(out)	Migr.(Kag)	Migr.(Kag)
Spill	0.00642***	0.00631***		-0.00197
(Out)	(0.000749)	(0.000587)		(0.00144)
Spill		0.00138	-0.00440***	-0.00381
(Kag)		(0.000777)	(0.00149)	(0.00268)
Corr. Prop.		0.0358	0.458***	0.460***
Migr.(Kag)		(0.0455)	(0.0397)	(0.0946)
α_t	-0.0685**	-0.0835**	0.232***	0.207***
	(0.0185)	(0.0220)	(0.0219)	(0.0432)
RI	0.000130	7.25e-05	-0.00170	-0.00117
	(0.000386)	(0.000350)	(0.00220)	(0.00124)
RI^2	2.67e-06**	2.81e-06***	8.58e-06	6.76e-06
	(7.56e-07)	(6.91e-07)	(5.98e-06)	(4.07e-06)
Subsistence	0.00335	0.00311	-0.0203	-0.0229
	(0.00636)	(0.00663)	(0.0160)	(0.0286)
Crop Sales	-0.0201	-0.0207	-0.0291*	-0.0315
	(0.0236)	(0.0232)	(0.0176)	(0.0191)
Non Agri	0.00743	0.00652	-0.0631***	-0.0682
L	(0.0184)	(0.0187)	(0.0240)	(0.0403)
Agri	-0.0162	-0.0181	-0.0267	-0.0278
L	(0.0133)	(0.0139)	(0.0180)	(0.0254)
Business	-0.00154	-0.000123	-0.151***	-0.154***
	(0.0117)	(0.0114)	(0.0168)	(0.0239)
LVST	-0.0123	-0.0119	-0.00364	-0.00430
	(0.0236)	(0.0237)	(0.0149)	(0.0231)
Subs.	-0.000822***	-0.000754***	-0.000742	-0.000725***
*RI	(0.000199)	(0.000186)	(0.00176)	(0.000153)
Sales	0.00111***	0.00110***	-0.000219	-0.000254
*RI	(6.55e-05)	(6.50e-05)	(0.00142)	(0.000204)
Non Agri	-0.000916***	-0.000848***	0.000123	0.000156
L*RI	(0.000133)	(0.000124)	(0.00134)	(0.000191)
Agri	1.13e-05	2.01e-05	-2.18e-05	-8.05e-06
L*RI	(7.02e-05)	(6.75e-05)	(0.000746)	(0.000127)
Business	0.000141	0.000123	0.000156	0.000182
*RI	(9.21e-05)	(9.31e-05)	(0.000829)	(0.000193)
LVST	-0.000360***	-0.000350***	-0.000523	-0.000532***
*RI	(7.14e-05)	(7.02e-05)	(0.000511)	(9.60e-05)
Log(land)	0.00127	0.00123	-0.00878***	-0.00844
	(0.00108)	(0.00102)	(0.00200)	(0.00732)
Prop.Lit.	0.00969	0.00920	0.0393	0.0406
	(0.0262)	(0.0254)	(0.0265)	(0.0218)
Size	-0.000702	-0.000606	-0.00247	-0.00178
	(0.00200)	(0.00203)	(0.00210)	(0.00408)
Split-off	0.0574***	0.0559***	-0.0602***	-0.0603
	(0.0124)	(0.0122)	(0.0192)	(0.0311)
Constant	0.0234	0.0202	0.155***	0.175
	(0.0352)	(0.0310)	(0.0319)	(0.110)
Obs.	3879	3879	3879	3879
R^2	0.305	0.308	0.292	0.295
F-Test	16.89***	17.2***	37.07***	114***

**** p< 0.01, ** p< 0.05, * p< 0.1; Robust standard errors in parentheses and clustered at the district level. The variable "Corr. Prop. Migr(Kag)" is the proportion of the initial household which has migrated within the region of Kagera, corrected to keep it exogenous. (see table 2)*

Table 6: Initial Household Fixed Effect (excl. migrants and incl. split-offs)

	(1) IHFE	(2) IHFE	(3) IHFE	(4) IHFE	(5) IHFE (2SLS)	(6) IHFE (2SLS)	(7) IHFE (2SLS)	(8) IHFE (2SLS)
	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.
RI	0.00047*	-0.0034***	-0.006***	-0.006***	0.0007	-0.004*	-0.0057*	-0.0057*
	(0.0002)	(0.0006)	(0.0011)	(0.0011)	(0.0005)	(0.0017)	(0.0022)	(0.0022)
					[0.140]			
RI^2			1.13e-05**	1.13e-05**			7.35e-06	7.35e-06
			(3.75e-06)	(3.75e-06)			(1.32e-05)	(1.32e-05)
Subs.		-0.157**	-0.159**	-0.159**		-0.147*	-0.148*	-0.148*
		(0.0487)	(0.0484)	(0.0484)		(0.0666)	(0.0680)	(0.0680)
Crop Sales		-0.0643	-0.0660	-0.0660		-0.0373	-0.0384	-0.0384
		(0.0346)	(0.0338)	(0.0338)		(0.0942)	(0.0960)	(0.0960)
Non Agri		0.115**	0.113**	0.113**		0.149	0.148	0.148
L		(0.0407)	(0.0403)	(0.0403)		(0.0906)	(0.0931)	(0.0931)
Agri		0.0944	0.0932	0.0932		0.113	0.112	0.112
L		(0.0624)	(0.0619)	(0.0619)		(0.0644)	(0.0651)	(0.0651)
Business		0.186**	0.186**	0.186**		0.273	0.272	0.272
		(0.0564)	(0.0564)	(0.0564)		(0.261)	(0.263)	(0.263)
LVST		0.0291	0.0288	0.0288		0.0322	0.0322	0.0322
		(0.0633)	(0.0634)	(0.0634)		(0.0616)	(0.0617)	(0.0617)
Subs.		0.0042***	0.0039***	0.0039***		0.0048**	0.0046*	0.0046*
*RI		(0.0007)	(0.0007)	(0.0007)		(0.0018)	(0.0021)	(0.0021)
Sales		0.0013***	0.0014***	0.0014***		0.0013**	0.0014**	0.0014**
*RI		(0.0003)	(0.0003)	(0.0003)		(0.0004)	(0.0005)	(0.0005)
Non Agri		0.003***	0.003***	0.003***		0.0034**	0.0034**	0.0034**
L*RI		(0.0004)	(0.0003)	(0.0003)		(0.0009)	(0.001)	(0.001)
Agri		-0.0016***	-0.0016***	-0.0016***		-0.0017***	-0.0016***	-0.0016***
L*RI		(0.0003)	(0.0003)	(0.0003)		(0.0003)	(0.0003)	(0.0003)
Business		-0.0025***	-0.0024***	-0.0024***		-0.0025***	-0.0025***	-0.0025***
*RI		(0.0002)	(0.0002)	(0.0002)		(0.0004)	(0.0005)	(0.0005)
LVST		0.0002	0.0002	0.0002		0.0005	0.0005	0.0005
*RI		(0.0002)	(0.0002)	(0.0002)		(0.001)	(0.001)	(0.001)
Log(land)	-0.00758	-0.00185	-0.00191	-0.00191	-0.00156	0.00275	0.00262	0.00262
	(0.00431)	(0.00457)	(0.00458)	(0.00458)	(0.0143)	(0.0152)	(0.0154)	(0.0154)
Prop.Lit.	0.227**	0.189**	0.187**	0.187**	0.225***	0.181**	0.180**	0.180**
	(0.0696)	(0.0660)	(0.0664)	(0.0664)	(0.0582)	(0.0702)	(0.0698)	(0.0698)
Size	-0.0463**	-0.0477**	-0.0476**	-0.0476**	-0.0431***	-0.0460**	-0.0460**	-0.0460**
	(0.0148)	(0.0136)	(0.0136)	(0.0136)	(0.0129)	(0.0138)	(0.0138)	(0.0138)
Split-off	0.0998	0.0953*	0.0948*	0.0948*	0.132**	0.115	0.113	0.113
	(0.0519)	(0.0469)	(0.0466)	(0.0466)	(0.0636)	(0.0606)	(0.0615)	(0.0615)
	[0.113]					[0.117]	[0.125]	[0.125]
Prop.	0.341*	0.362**	0.364**	0.364**	0.332	0.488	0.504	0.504
Migr.(out)	(0.146)	(0.141)	(0.140)	(0.140)	(0.417)	(0.599)	(0.583)	(0.583)
Migr.(Kag)	0.145*	0.217**	0.216**	0.216**	0.718	0.677	0.670	0.670
	(0.0618)	(0.0681)	(0.0683)	(0.0683)	(1.105)	(1.232)	(1.244)	(1.244)
IHFE	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.
Cluster D				sign.				sign.
Constant	11.81***	11.91***	11.91***	11.91***	11.69***	11.79***	11.79***	11.79***
	(0.0819)	(0.0675)	(0.0675)	(0.0675)	(0.304)	(0.373)	(0.373)	(0.373)
α_t	0.0698*	0.0615	0.0652	0.0652	-0.0490	-0.0213	-0.0159	-0.0159
	(0.0319)	(0.0363)	(0.0339)	(0.0339)	(0.203)	(0.223)	(0.233)	(0.233)
Obs.	3803	3803	3803	3803	3803	3803	3803	3803
R^2	0.569	0.598	0.598	0.598	0.517	0.565	0.566	0.566

*Note: *** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance.Robust standard errors in parentheses and clustered at the district level.*

Table 7: Household Fixed Effect (incl. migrants and split-offs)

	(1) HHFE	(2) HHFE	(3) HHFE	(4) HHFE	(5) HHFE (2SLS)	(6) HHFE (2SLS)	(7) HHFE (2SLS)	(8) HHFE (2SLS)
	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.
RI	0.0006*	-0.0039***	-0.0059***	-0.007**	0.0007***	-0.0041***	-0.0057***	-0.0067***
	(0.0003)	(0.0008)	(0.0013)	(0.0023)	(0.0001)	(0.0006)	(0.0013)	(0.0016)
RI^2			9.04e-06	1.27e-05			7.46e-06*	1.08e-05*
			(4.67e-06)	(7.76e-06)			(4.10e-06)	(6.15e-06)
			[0.111]	[0.163]				
Subs.		-0.146**	-0.147**	-0.122**		-0.141**	-0.142**	-0.121***
		(0.0548)	(0.0536)	(0.0421)		(0.0577)	(0.0584)	(0.0292)
Sales		-0.0813	-0.0828	-0.0703		-0.0762*	-0.0776*	-0.0652
		(0.0538)	(0.0530)	(0.0622)		(0.0401)	(0.0404)	(0.0421)
Non-agri		0.144*	0.142*	0.134*		0.154***	0.152***	0.148***
L		(0.0606)	(0.0599)	(0.0629)		(0.0541)	(0.0538)	(0.0340)
Agri		0.0876	0.0864	0.0936		0.0931	0.0918	0.0977*
L		(0.0971)	(0.0964)	(0.0954)		(0.0941)	(0.0948)	(0.0591)
Business		0.173*	0.173*	0.169*		0.201**	0.200**	0.196***
		(0.0778)	(0.0779)	(0.0729)		(0.0807)	(0.0812)	(0.0707)
LVST		0.00676	0.00622	0.00456		0.00814	0.00759	0.00464
		(0.0862)	(0.0865)	(0.0815)		(0.0552)	(0.0555)	(0.0486)
Subs.		0.0047***	0.0044***	0.0043***		0.0049***	0.0046***	0.0044***
*RI		(0.0008)	(0.0009)	(0.0005)		(0.0006)	(0.0006)	(0.0004)
Sales		0.0021***	0.0022***	0.0023***		0.0021***	0.0022***	0.0023***
*RI		(0.0004)	(0.0004)	(0.0004)		(0.0002)	(0.0002)	(0.0002)
NonAgri.		0.0035***	0.0034***	0.0037***		0.0036***	0.0035***	0.0037***
L*RI		(0.0005)	(0.0005)	(0.0003)		(0.0004)	(0.0004)	(0.0002)
Agri		-0.0018**	-0.0018**	-0.0019**		-0.0018***	-0.0018***	-0.0018***
L*RI		(0.0005)	(0.0005)	(0.0005)		(0.0003)	(0.0003)	(0.0003)
Business		-0.002***	-0.0019***	-0.0017***		-0.0021***	-0.0019***	-0.0018***
*RI		(0.0003)	(0.0004)	(0.0004)		(0.0002)	(0.0002)	(0.0003)
LVST		0.0005	0.0005	0.0006*		0.0006***	0.0006***	0.0007***
*RI		(0.0003)	(0.0003)	(0.0003)		(0.0002)	(0.0002)	(0.0002)
Log(land)	-0.0111	-0.00515	-0.00518	-0.000901	-0.00928	-0.00333	-0.00348	-0.000815
	(0.00845)	(0.00846)	(0.00849)	(0.00911)	(0.0196)	(0.0133)	(0.0135)	(0.00543)
Prop.Lit.	0.231*	0.193	0.192	0.171*	0.222	0.184	0.183	0.166***
	(0.109)	(0.101)	(0.102)	(0.0786)	(0.155)	(0.115)	(0.115)	(0.0498)
		[0.116]	[0.119]		[0.151]	[0.112]	[0.112]	
Size	-0.0352	-0.0365	-0.0364	-0.0387*	-0.0343***	-0.0359***	-0.0358***	-0.0383***
	(0.0219)	(0.0197)	(0.0197)	(0.0178)	(0.0112)	(0.0112)	(0.0112)	(0.0106)
	[0.169]	[0.123]	[0.124]					
Split-off	0.147	0.137	0.136	0.123	0.155***	0.143***	0.143***	0.134***
	(0.0874)	(0.0789)	(0.0787)	(0.0782)	(0.0508)	(0.0439)	(0.0435)	(0.0479)
	(0.154)	(0.144)	(0.144)	(0.175)	(0.00224)	(0.00111)	(0.00105)	
Migr.(Kag)	0.147	0.218*	0.217*	0.0826	0.343	0.388	0.385	0.371
	(0.0769)	(0.0935)	(0.0940)	(0.0975)	(0.303)	(0.304)	(0.305)	(0.456)
	[0.114]							
Migr.(out)	0.362**	0.383**	0.381**	1.969***	0.432	0.462	0.453	1.911***
	(0.115)	(0.111)	(0.111)	(0.223)	(1.259)	(0.970)	(0.979)	(0.170)
Cluster D				sign.				sign.
α_t	0.0241	0.00520	0.00817	0.00518	-0.0165	-0.0320	-0.0287	-0.0140
	(0.0530)	(0.0597)	(0.0577)	(0.0572)	(0.0709)	(0.0638)	(0.0656)	(0.0506)
R^2	0.226	0.274	0.274	0.332	0.218	0.268	0.268	0.322
Obs.	4104	4104	4104	4104	3966	3966	3966	3966
Anderson					68.81***	67.93***	66.24***	95.77***
Stock-Yogo					17.42	16.93	16.66	16.85
Hansen (a)					3.58	5.063	5.07	
(p-value)					(0.167)	(0.079)	(0.079)	

*Note: *** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance.*

Robust standard errors in parentheses and clustered at the district level.

Table 8: Altruism Test based on Altonji et al (1992) and Witoelar (2005)

	(1) IHFE (2SLS)	(2) IHFE (2SLS)	(3) IHFE (2SLS)	(4) IHFE (2SLS)	(5) IHFE (2SLS)	(6) IHFE (2SLS)	(7) IHFE (2SLS)	(8) IHFE (2SLS)
IV	Witoelar	Witoelar	Dist (Road)	Dist (Road)	Dist (Road)	Dist (Road)	Dist (Road)	Dist (Road)
Split-off Migr.(Out)	incl. excl.	incl. excl.	incl. excl.	incl. excl.	incl. excl.	incl. incl.	incl. incl.	incl. incl.
Dep Var	log(cons)	log(cons)	log(cons)	log(cons)	log(cons)	log(cons)	log(cons)	log(cons)
Log(netinc)	0.125	0.142	0.517***	0.585***	0.610***	0.536***	0.583***	0.603***
	(0.0726)	(0.0828)	(0.0116)	(0.0145)	(0.0258)	(0.0174)	(0.0142)	(0.0272)
	[0.146]	[0.146]						
α_t	0.168***	0.0968***	0.251***	0.153**	0.185**	0.316***	0.166**	0.192*
	(0.00926)	(0.0159)	(0.0495)	(0.0414)	(0.0519)	(0.0756)	(0.0633)	(0.0797)
RI	0.0003*	0.0004*	-1.34e-05	0.0002	0.0004*	-1.46e-05	0.0004	0.0005**
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)
Size	-0.0521**	-0.0493**	-0.0534***	-0.0373**	-0.0354**	-0.0494**	-0.0297**	-0.0281*
	(0.0152)	(0.0143)	(0.0119)	(0.00953)	(0.00982)	(0.0144)	(0.0113)	(0.0124)
Prop.Lit.	0.223***	0.219**	0.193***	0.175**	0.177**	0.316***	0.207**	0.208**
	(0.0552)	(0.0550)	(0.0455)	(0.0524)	(0.0539)	(0.0445)	(0.0558)	(0.0567)
dumnewhh		0.0850		0.0558	0.0510		0.0423	0.0379
		(0.0492)		(0.0515)	(0.0514)		(0.0474)	(0.0471)
Prop. Migr. (out)^b		0.274		0.140	0.146		0.590***	0.585***
		(0.137)		(0.120)	(0.120)		(0.124)	(0.126)
Migr.(Kag)		0.218*		0.352***	0.363***		0.346***	0.354***
		(0.0915)		(0.0459)	(0.0424)		(0.0432)	(0.0400)
Log(land)				-0.0487***	-0.0477***		-0.0499***	-0.0488***
				(0.00820)	(0.00811)		(0.00600)	(0.00585)
Log(Bus)					0.000865			0.00125
					(0.00358)			(0.00530)
Log(eqip)					-0.00834			-0.00670
					(0.00481)			(0.00485)
Constant	10.37***	10.11***	5.926***	5.655***	5.368***	5.621***	5.668***	5.439***
	(0.847)	(0.948)	(0.178)	(0.148)	(0.294)	(0.262)	(0.202)	(0.352)
IHFE	sign.	sign.	sign.	sign.	sign.	sign.	sign.	sign.
Obs.	3781	3781	3781	3781	3781	4082	4082	4082
R-squared	0.609	0.621	0.333	0.290	0.248	0.247	0.275	0.241
First Stage								
Dep Var	log(netinc)	log(netinc)	log(netinc)	log(netinc)	log(netinc)	log(netinc)	log(netinc)	log(netinc)
Log(land)	0.0629***	0.0584***		0.0680***	0.0584***		0.0617***	0.0521***
	(0.009)^{iv}	(0.009)^{iv}		(0.00811)	(0.00860)		(0.00629)	(0.00603)
Log(Bus)	0.0393***	0.0395***			0.0395***			0.0370***
	(0.004)^{iv}	(0.004)^{iv}			(0.00402)			(0.00677)
Log(eqip)	0.0294**	0.0291**			0.0291**			0.0265*
	(0.008)^{iv}	(0.008)^{iv}			(0.00776)			(0.0104)
Log(road) (Uganda)			-0.266***	-0.135***	-0.619***	-0.679***	-0.651***	-0.663***
			(0.032)^{iv}	(0.026)^{iv}	(0.037)^{iv}	(0.023)^{iv}	(0.019)^{iv}	(0.015)^{iv}
Log(road) (Rwanda)			-0.627***	-0.444***	-0.794***	-0.969***	-0.975***	-0.834***
			(0.062)^{iv}	(0.051)^{iv}	(0.052)^{iv}	(0.05)^{iv}	(0.046)^{iv}	(0.043)^{iv}
Obs.	3857	3857	3857	3857	3857	4162	4162	4162
R-squared	0.379	0.388	0.328	0.368	0.388	0.316	0.359	0.376

*Note: *** p< 0.01, ** p< 0.05, * p< 0.1; Robust standard errors in parentheses and clustered at the district level. Columns (1) and (2) use the instruments proposed by Witoelar (2005), i.e. the value of land, farm productive equipments and non-farm productive assets. Regressions (3) to (8) use the distance to the main economic partners as IV, the distance by road to Uganda and Rwanda. To make this distinction clear in the first-stage estimation, iv indicates instrumental variables that are excluded from the second-stage regression presented in the above part of the table. Regressions (6) to (8) differ from (1) to (5) as the sample used includes those migrating outside the region of Kagera. Therefore, b indicates that the variable "Prop. Migr.(out)" is replaced by the direct measure of the decision to move outside the region of Kagera by the household in the columns (7) and (8).*

Table 9: Household Fixed Effect (excl. migrants): Do the effects last overtime?

	(1) HHFE split. incl.	(2) HHFE split. incl.	(3) HHFE split. incl.	(4) HHFE (2SLS) split. incl.	(5) HHFE (2SLS) split. incl.	(6) HHFE (2SLS) split. incl.
RI(BU)	0.000460 [0.181]	-0.00691***	-0.00803**	0.000520***	-0.00677***	-0.00785***
	(0.000296)	(0.00166)	(0.00271)	(0.000185)	(0.00137)	(0.00200)
$RI(BU)^2$		1.10e-05*	1.32e-05[0.167]		1.01e-05**	1.21e-05**
		(5.35e-06)	(8.19e-06)		(4.41e-06)	(6.08e-06)
RI(RW)	1616	3910***	4140	1279	3216[0.156]	5685[0.152]
	-1739	(750.0)	-2821	-1556	-2267	-3968
$RI(RW)^2$		dropped	dropped		dropped	dropped
Subs.		-0.133**	-0.124**		-0.131***	-0.121***
		(0.0486)	(0.0460)		(0.0328)	(0.0297)
Sales		-0.0616	-0.0551		-0.0588	-0.0484
		(0.0561)	(0.0587)		(0.0362)	(0.0402)
NonAgri.		0.128*	0.0959		0.135***	0.108***
L		(0.0500)	(0.0596)		(0.0309)	(0.0301)
Agri.		0.110	0.103		0.112**	0.108**
L		(0.0876)	(0.0913)		(0.0486)	(0.0507)
Business		0.193*	0.186*		0.208***	0.209***
		(0.0757)	(0.0738)		(0.0737)	(0.0685)
LVST		0.00961	0.0129		0.0102	0.0148
		(0.0947)	(0.0956)		(0.0592)	(0.0573)
Subs.		0.00456***	0.00507***		0.00467***	0.00516***
*RI(BU)		(0.000823)	(0.000746)		(0.000777)	(0.000741)
Sales		0.00193**	0.00197***		0.00193***	0.00188***
*RI(BU)		(0.000492)	(0.000460)		(0.000604)	(0.000619)
NonAgr.		0.00395***	0.00438***		0.00400***	0.00441***
L*RI(BU)		(0.000416)	(0.000424)		(0.000505)	(0.000462)
Agr.		-0.00176**	-0.00187**		-0.00175***	-0.00184***
L**RI(BU)		(0.000475)	(0.000464)		(0.000288)	(0.000300)
Business		-0.00195***	-0.00186***		-0.00198***	-0.00196***
*RI(BU)		(0.000323)	(0.000377)		(0.000259)	(0.000305)
LVST		0.000802**	0.000813*		0.000860***	0.000911**
*RI(BU)		(0.000292)	(0.000385)		(0.000284)	(0.000364)
Subs.		-1895	-1882		-1221	-2503
*RI(RW)		-1220	-2113		-1748	-2218
Sales		-8273	-8585		-7456	-8895**
*RI(RW)		-5922	-5325		-5477	-4428
NonAgr.		4886	5380		4963***	5003***
L*RI(RW)		-3933	-3771		-1033	-1309
Agr.		-16649957*	-10007517		-16863899***	-10254088***
L*RI(RW)		(7.92e+06)	(6.98e+06)		(3.75e+06)	(3.41e+06)
Business		-6714***	-7135***		-5956*	-7889***
*RI(RW)		(707.5)	-1217		-3193	-2465
LVST		8980**	9586***		8198***	8428***
*RI(RW)		-2266	-1768		-2559	-2279
α_t	0.0702	0.0612	0.0564	0.0296	0.0405	0.0525
	(0.0499)	(0.0527)	(0.0652)	(0.0676)	(0.0835)	(0.0515)
HHvar	incl.	incl.	incl.	incl.	incl.	incl.
Migr. var	incl.	incl.	incl.	incl.	incl.	incl.
Cluster D			sign.			sign.
Obs.	3803	3803	3803	3674	3674	3674
R^2	0.185	0.237	0.295	0.181	0.235	0.288

*RI(BU) represents the RI index, restricted to the refugees from Burundi while RI(RW) corresponds to the RI index restricted to the refugee camps, populated of refugees from Rwanda. *** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance. Robust standard errors in parentheses and clustered at the district level.*

Table 10: Household Fixed Effect (excl. migrants) with alternative treatment variables

	(1) HHFE (2SLS)	(2) HHFE (2SLS)	(3) HHFE (2SLS)	(4) HHFE (2SLS)	(5) HHFE (2SLS)	(6) HHFE (2SLS)
Treat	RC	RC	RC	RI(closest)	RI(closest)	RI(closest)
	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.
Treat	-0.0161	0.0565	-0.0175	0.000533***	-0.00684***	-0.00807***
	(0.0600)	(0.103)	(0.0776)	(0.000191)	(0.00139)	(0.00202)
$Treat^2$					1.04e-05**	1.28e-05**
					(4.56e-06)	(6.28e-06)
Subs.		-0.115***	-0.110***		-0.128***	-0.120***
		(0.0340)	(0.0323)		(0.0325)	(0.0298)
Sales		-0.0705*	-0.0629		-0.0593	-0.0502
		(0.0413)	(0.0454)		(0.0361)	(0.0398)
Non-agri		0.171***	0.143***		0.140***	0.111***
L		(0.0320)	(0.0351)		(0.0315)	(0.0304)
Agri		0.119***	0.115***		0.108**	0.105**
L		(0.0427)	(0.0447)		(0.0483)	(0.0506)
Business		0.223***	0.222***		0.212***	0.211***
		(0.0779)	(0.0735)		(0.0728)	(0.0675)
LVST		0.0128	0.0184		0.0102	0.0153
		(0.0613)	(0.0586)		(0.0585)	(0.0564)
Subs.		-0.0542	-0.00510		0.00475***	0.00530***
*Treat		(0.0892)	(0.0739)		(0.000805)	(0.000758)
Sales		0.153**	0.169***		0.00201***	0.00196***
*Treat		(0.0749)	(0.0439)		(0.000613)	(0.000629)
NonAgri.		-0.121**	-0.105***		0.00408***	0.00454***
L*Treat		(0.0524)	(0.0364)		(0.000529)	(0.000472)
Agri		-0.122	-0.128		-0.00176***	-0.00185***
L*Treat		(0.108)	(0.113)		(0.000292)	(0.000311)
Business		-0.106	-0.0713		-0.00206***	-0.00202***
*Treat		(0.110)	(0.0796)		(0.000267)	(0.000309)
LVST		-0.0230	-0.0263		0.000890***	0.000945**
*Treat		(0.0851)	(0.0793)		(0.000293)	(0.000375)
Log(land)	-0.00362	0.000494	0.00251	-0.00361	0.000479	0.00268
	(0.00520)	(0.00573)	(0.00456)	(0.00522)	(0.00566)	(0.00434)
Prop.Lit.	0.179**	0.151**	0.134**	0.180***	0.151**	0.137**
	(0.0703)	(0.0671)	(0.0541)	(0.0693)	(0.0693)	(0.0556)
Size	-0.0380***	-0.0400***	-0.0413***	-0.0379***	-0.0397***	-0.0409***
	(0.0123)	(0.0110)	(0.0100)	(0.0124)	(0.0112)	(0.0102)
Split-off	0.135***	0.119***	0.111***	0.135***	0.117***	0.109***
	(0.0486)	(0.0402)	(0.0390)	(0.0485)	(0.0423)	(0.0410)
Migr.(Kag)	0.263	0.315	0.303	0.272	0.332	0.300
	(0.269)	(0.282)	(0.406)	(0.271)	(0.279)	(0.422)
Prop.	0.206	0.329	0.430	0.187	0.352	0.445
Migr.(out)	(0.381)	(0.378)	(0.400)	(0.392)	(0.385)	(0.417)
Cluster D			sign.			sign.
α_t	0.0406	0.0337	0.0526	0.0298	0.0353	0.0508
	(0.0634)	(0.0837)	(0.0537)	(0.0681)	(0.0829)	(0.0514)
Obs.	3674	3674	3674	3674	3674	3674
R^2	0.181	0.232	0.285	0.181	0.233	0.287
F-Test	3.66*	1.28	4.13*	53.65***	4.80*	4.43*

**** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance. Robust standard errors in parentheses and clustered at the district level. RC is a dummy variable indicating whether the community leader has replied positively to the question "Were there any refugee settlement closed by?". RI(Closest) is similar to the RI index but restricted to the closest refugee camp.*

Table 11: Household Fixed Effect (excl. migrants) with alternative treatment variables

	(1) HHFE (2SLS)	(2) HHFE (2SLS)	(3) HHFE (2SLS)	(4) HHFE (2SLS)	(5) HHFE (2SLS)	(6) HHFE (2SLS)
Treat	Space	Space	Space	Scale	Scale	Scale
Treat	53.72**	-579.5***	-702.8***	-1.20e-06	-5.73e-06***	-5.64e-06***
	(25.85)	(90.55)	(143.8)	(9.72e-07)	(1.25e-06)	(1.52e-06)
$Treat^2$		92742**	111293**		4.49e-11***	4.34e-11***
		-36914	-47519		1.48e-11	1.60e-11
Subs.		-0.129***	-0.121***		-0.123***	-0.114***
		(0.0326)	(0.0300)		(0.0294)	(0.0280)
Sales		-0.0593	-0.0502		-0.0626*	-0.0545
		(0.0363)	(0.0399)		(0.0322)	(0.0348)
Non-agri		0.138***	0.108***		0.133***	0.109***
L		(0.0312)	(0.0302)		(0.0328)	(0.0319)
Agri		0.106**	0.104**		0.120***	0.116***
L		(0.0486)	(0.0509)		(0.0391)	(0.0418)
Business		0.213***	0.212***		0.210***	0.210***
		(0.0729)	(0.0676)		(0.0699)	(0.0669)
LVST		0.00970	0.0146		-0.00828	0.000720
		(0.0591)	(0.0571)		(0.0614)	(0.0618)
Subs.		437.1***	511.2***		-1.38e-07	2.64e-07
*Treat		(78.89)	(65.61)		(7.08e-07)	(1.03e-06)
Sales		242.7***	237.7***		1.79e-06***	2.33e-06***
*Treat		(58.82)	(57.15)		(6.43e-07)	(3.68e-07)
NonAgri.		435.9***	498.0***		6.90e-07	8.19e-07
L*Treat		(38.84)	(35.13)		(9.50e-07)	(6.52e-07)
Agri		-123.9***	-133.6***		-2.66e-06**	-2.98e-06
L*Treat		(38.39)	(41.99)		(1.34e-06)	(2.00e-06)
Business		-218.9***	-210.8***		-1.53e-06***	-1.39e-06**
*Treat		(35.23)	(41.38)		(4.84e-07)	(5.60e-07)
LVST		105.2***	115.6**		1.36e-06*	1.09e-06
*Treat		(37.56)	(48.55)		(6.97e-07)	(7.49e-07)
Log(land)	-0.00356	0.000459	0.00267	-0.00384	-0.000223	0.00208
	(0.00525)	(0.00565)	(0.00433)	(0.00506)	(0.00536)	(0.00439)
Prop.Lit.	0.180***	0.151**	0.137**	0.177**	0.153**	0.136**
	(0.0694)	(0.0695)	(0.0560)	(0.0700)	(0.0719)	(0.0575)
Size	-0.0379***	-0.0398***	-0.0410***	-0.0379***	-0.0396***	-0.0413***
	(0.0124)	(0.0112)	(0.0102)	(0.0124)	(0.0112)	(0.0100)
Split-off	0.135***	0.117***	0.109***	0.132***	0.116***	0.104***
	(0.0485)	(0.0422)	(0.0407)	(0.0466)	(0.0399)	(0.0383)
Migr.(Kag)	0.273	0.334	0.304	0.258	0.289	0.226
	(0.272)	(0.279)	(0.419)	(0.266)	(0.268)	(0.351)
Prop.	0.185	0.352	0.444	0.244	0.329	0.404
Migr.(out)	(0.392)	(0.386)	(0.416)	(0.395)	(0.351)	(0.360)
Cluster D			sign.			sign.
α_t	0.0291	0.0350	0.0507	0.0604	0.0735	0.0813
	(0.0685)	(0.0833)	(0.0518)	(0.0655)	(0.0831)	(0.0554)
Obs.	3674	3674	3674	3674	3674	3674
R^2	0.180	0.232	0.287	0.184	0.238	0.293
F-Test	21.89***	61.71***	98.19***	164.74***	3.71*	5.2**

**** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance. Robust standard errors in parentheses and clustered at the district level. Space represents the spatial dimension of the RI(closest) index, i.e $exp(d_{v,closest})^{-1}$ while scale is computed as the number of refugees in the closest refugee camp.*

Table 12: Household Fixed Effect (excl. migrants): scale or space

	(1) HHFE	(2) HHFE	(3) HHFE	(4) HHFE	(5) HHFE (2SLS)	(6) HHFE (2SLS)	(7) HHFE (2SLS)	(8) HHFE (2SLS)
	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.	split. incl.
Space	163.2	-388.3***	-192.5*	-301.9*	171.7**	-395.9***	-199.8***	-310.6**
	(111.0)	(79.05)	(76.27)	(139.1)	(76.70)	(34.83)	(68.47)	(130.3)
$Space^2$			-127435**	-86992			-126443***	-89784
			-47217	-100381			-39502	-84732
Scale	-2.15e-06	-7.11e-07	-5.38e-06**	-5.30e-06*	-2.15e-06*	-8.08e-07	-5.19e-06***	-5.04e-06***
	(1.91e-06)	(1.57e-06)	(1.70e-06)	(2.37e-06)	(1.16e-06)	(1.00e-06)	(1.03e-06)	(1.37e-06)
$Scale^2$			5.18e-11**	4.52e-11*			4.88e-11***	4.39e-11***
			1.33e-11	2.07e-11			7.18e-12	(1.38e-11)
Subs.		522.9***	553.2***	568.3***		536.7***	568.6***	579.9***
*Space		(51.96)	(75.44)	(61.66)		(35.31)	(70.81)	(47.62)
Sales		118.2*	147.8**	120.3*		101	134.3**	97.47
*Space		(47.45)	(46.08)	(49.41)		(67.52)	(61.17)	(73.63)
						[0.135]		[0.186]
NonAgri.		472.0***	515.0***	583.0***		470.7***	509.1***	574.7***
L*Space		(76.94)	(104.7)	(65.68)		(46.24)	(63.61)	(44.82)
Agri		115.7	99.30	119.7		107.5	91.10	113.6
L*Space		(237.7)	(214.1)	(351.5)		(141.8)	(129.5)	(204.8)
Business		-169.5	-162.7*	-180[0.114]		-183.7***	-178.9***	-187.4***
*Space		(86.13)	(77.59)	(94.05)		(52.37)	(52.56)	(54.59)
		[0.106]						
LVST		87.1	54.70	83.29*		98.30**	64.38*	98.24**
*Space		(52.69)	(42.70)	(38.48)		(43.26)	(32.90)	(38.39)
		[0.159]						
Subs.		-1.15e-06	-1.91e-06***	-1.09e-06		-1.10e-06**	-1.85e-06***	-1.16e-06
*Scale		(6.47e-07)	(3.30e-07)	(7.04e-07)		(4.65e-07)	(3.46e-07)	(8.26e-07)
		[0.135]		[0.183]				[0.160]
Sales		1.53e-06**	1.20e-06*	1.71e-06**		1.86e-06**	1.52e-06*	2.05e-06***
*Scale		(5.63e-07)	(5.32e-07)	(4.94e-07)		(8.32e-07)	(7.85e-07)	(4.83e-07)
NonAgri.		-3.40e-07	-1.06e-06	-1.04e-06		-2.53e-07	-9.21e-07	-1.03e-06*
L*Scale		(9.67e-07)	(1.16e-06)	(9.03e-07)		(6.11e-07)	(7.03e-07)	(5.71e-07)
							[0.190]	
Agri		-4.06e-06	-3.76e-06	-4.16e-06		-3.90e-06*	-3.60e-06*	-4.02e-06
L*Scale		(3.62e-06)	(3.28e-06)	(5.54e-06)		(2.13e-06)	(1.96e-06)	(3.18e-06)
Business		-9.81e-07	-8.45e-07	-4.20e-07		-7.81e-07	-6.40e-07	-4.10e-07
*Scale		(1.39e-06)	(1.37e-06)	(1.78e-06)		(8.24e-07)	(8.30e-07)	(1.12e-06)
LVST		2.36e-07	7.37e-07	3.59e-07		1.89e-07	6.81e-07	2.61e-07
*Scale		(1.47e-06)	(1.35e-06)	(1.41e-06)		(8.56e-07)	(7.69e-07)	(7.38e-07)
Migr.(Kag)	0.150	0.225*	0.216*	0.0668	0.279	0.332	0.322	0.241
	(0.0823)	(0.0916)	(0.0901)	(0.0859)	(0.268)	(0.267)	(0.264)	(0.338)
Prop.	0.326	0.351*	0.338	0.343*	0.197	0.352	0.294	0.381
Migr.(out)	(0.180)	(0.172)	(0.171)	(0.158)	(0.378)	(0.360)	(0.343)	(0.358)
α_t	0.0962**	0.0808	0.0927*	0.0888	0.0546	0.0578	0.0635	0.0804
	(0.0324)	(0.0425)	(0.0370)	(0.0581)	(0.0662)	(0.0801)	(0.0840)	(0.0555)
Cluster D				sign.				sign.
Obs.	3803	3803	3803	3803	3674	3674	3674	3674
R^2	0.190	0.241	0.244	0.300	0.185	0.238	0.241	0.296
F-Test	5.2***	31.7***	31.05***	22.32***	53.41***	317.31***	86.12***	86.28***

**** p< 0.01, ** p< 0.05, * p< 0.1; Between brackets, are included the p-value of coefficients closed to significance. Robust standard errors in parentheses and clustered at the district level. Although not presented for clarity reason, the activity and household variables are included in the regression and provide similar results (both in terms of magnitude of the coefficients and significance) than in table 4.*

Table 13: Household Fixed Effect (excl. migrants): Interpretation for scale and space

Tables Treat	(1) Tab4(5) RI	(2) Tab4(6) RI	(3) Tab4(7) RI	(4) Tab4(8) RI	(5) Tab11(2) Space	(6) Tab11(3) Space	(7) Tab11(5) Scale	(8) Tab11(6) Scale	(9)(a) Tab12(7) Space	(9)(b) Tab12(7) Scale	(10)(a) Tab12(8) Space	(10)(b) Tab12(8) Scale
Minimum			9.24	8.05	241.45	306.61	3278.8	2718.69	257.86	53176.23	3581,00	57403.19
APE(Treat)	0.0005	-0.0007	-0.0031	-0.0039	-212.8	-274.5	-0.0000	-0.0000	263.81	-0.0000	3596.88	-0.0000
$\Delta\sigma$	0.0132	-0.0181	-0.0792	-0.0991	-0.0465	-0.05996	-0.1193	-0.1116	0.0576	-0.1411	0.7856	-0.1292
MaxPE(Treat)	0.0005	-0.0007	0.0021	0.0025	172.94	188.36	0.0000	0.0000	-262.12	0.0000	3223.43	0.0000
$\Delta\sigma$	0.0132	-0.0181	0.0541	0.0633	0.0378	0.0411	0.1147	0.1146	-0.0572	0.1132	0.704	0.0996
APE(Treat, subs.=1)	0.0005	0.0005	-0.002	-0.0027	-136.54	-184.57	-0.0000	-0.0000	360.81	-0.0000	5393.63	-0.0000
$\Delta\sigma$	0.0132	0.0128	-0.0518	-0.0688	-0.0298	-0.0403	-0.1205	-0.1097	0.0788	0.0010	1.178	-0.0279
MaxPE(Treat, subs.=1)		0.0005	0.0032	0.0037	249.21	278.34	0.0000	0.0000	-165.12	0.0000	5020.18	0.0000
$\Delta\sigma$		0.0128	0.0816	0.0936	0.0544	0.0608	0.1135	0.1165	-0.0361	-0.0008	1.096	0.0217
APE(Treat, sales.=1)		-0.0026	-0.0047	-0.0059	-330.94	-458.07	-0.0000	-0.0000	-73.49	-0.0000	-218.8	-0.0000
$\Delta\sigma$		-0.0669	-0.1203	-0.1519	-0.0723	-0.10004	-0.076	-0.062	-0.016	0.0005	-0.0478	-0.0118
MaxPE(Treat, sales.=1)		-0.0026	0.0005	0.0004	54.81	4.84	0.0000	0.0000	-599.42	0.00001	-592.25	0.000007
$\Delta\sigma$		-0.0669	0.0131	0.0105	0.01197	0.00106	0.15803	0.1643	-0.1309	-0.0013	-0.1293	0.0379
APE(Treat,Noagr.L=1)		-0.0004	-0.0027	-0.0034	-137.74	-197.77	-0.0000	-0.0000	301.31	-0.0000	258.43	-0.0000
$\Delta\sigma$		-0.0102	-0.0689	-0.088	-0.0301	-0.0432	-0.1014	-0.1038	0.0658	0.0009	0.0564	-0.02726
MaxPE(Treat,NoAgr.L.=1)		-0.0004	0.0025	0.0029	248.012	265.1458218	0.000006	0.000005	-224.62	0.0000	-115.02	0.000004
$\Delta\sigma$		-0.0102	0.0644	0.0744	0.0542	0.0579	0.1326	0.1224	-0.0491	-0.0009	-0.0251	0.0224
APE(Treat, Agri.L=1)		-0.00622	-0.0084	-0.0096	-697.54	-829.366	-0.000008	-0.000008	-116.69	-0.00001	-202.67	-0.000008
$\Delta\sigma$		-0.1589	-0.2138	-0.2467	-0.1523	-0.1811	-0.1788	-0.1847	-0.0255	0.0013	-0.04426	-0.0423
MaxPE(Treat, Agri.L=1)		-0.00622	-0.00315	-0.0033	-311.79	-366.45	0.000002	0.000002	-642.62	0.000003	-576.12	0.000001
$\Delta\sigma$		-0.1589	-0.0805	-0.0843	-0.0681	-0.08003	0.0552	0.0415	-0.14035	-0.0005	-0.1258	0.0074
APE(Treat, Business=1)		-0.00656	-0.00865	-0.00979	-792.54	-906.57	-0.000007	-0.000006	-386.69	-0.000005	-503.67	-0.000005
$\Delta\sigma$		-0.1676	-0.221	-0.2503	-0.1731	-0.19799	-0.1527	-0.1479	-0.0844	0.0008	-0.110002	-0.02415
MaxPE(Treat, Bus.=1)		-0.00656	-0.0034	-0.0034	-406.79	-443.65	0.000003	0.000003	-912.62	0.000006	-877.12	0.000005
$\Delta\sigma$		-0.1676	-0.0876	-0.0879	-0.0888	-0.0969	0.0813	0.0783	-0.1993	-0.00097	-0.1916	0.0255
APE(Treat, Livestock=1)		-0.00359	-0.0058	-0.0069	-468.44	-580.17	-0.000004	-0.000004	-143.41	-0.000004	-218.034	-0.000004
$\Delta\sigma$		-0.0918	-0.1478	-0.1768	-0.1023	-0.1267	-0.0859	-0.0906	-0.0313	0.0006	-0.0476	-0.0208
MaxPE(Treat, LVST=1)		-0.0036	-0.00057	-0.00056	-82.69	-117.25	0.000006	0.000006	-669.34	0.000007	-591.48	0.000006
$\Delta\sigma$		-0.0918	-0.01448	-0.0144	-0.018	-0.0256	0.1481	0.1356	-0.14618	-0.00119	-0.12918	0.0289

APE means Average Partial Effect and is computed as the partial effect of the treatment variable on the dependent variable, taking the average value of the treatment variable. The MaxPE is the maximum partial effect, i.e. the partial effect computed at the maximum value of the treatment variable.

Table 14: Could a refugee camp generate agglomeration economies?

	(1) HHFE (2SLS)	(2) HHFE (2SLS)	(3) HHFE (2SLS)	(4) HHFE (2SLS)	(5) HHFE (2SLS)	(6) HHFE (2SLS)	(7) HHFE (2SLS)
Space	-176.7**	-131.4	-299.5	-34305	-542432**	-162950	-551416*
	(68.87)	(290.9)	(266.2)	-46828	-243677	-105356	-283366
						[0.122]	
$Space^2$	-134017***	-136684	-56164	568693		1.37e+06**	
	-42059	-137641	-124084	-346673		-614214	
	(0.00144)	(0.321)	(0.651)	(0.101)		(0.0258)	
Scale	-3.98e-06***	-9.67e-06*	-7.31e-06	-6.36e-05***	-0.000173***	-0.000118***	-0.000172***
	(1.34e-06)	(5.87e-06)	(5.79e-06)	(1.87e-05)	(4.06e-05)	(4.20e-05)	(4.26e-05)
$Scale^2$	4.31e-11***	8.18e-11**	6.15e-11*	-1.78e-10*	-3.67e-10**	-4.43e-10**	-3.77e-10*
	8.33e-12	3.29e-11	3.42e-11	(1.04e-10)	(1.81e-10)	(1.89e-10)	(2.15e-10)
Log(density)	-0.0987	0.207**	0.218**	-0.134	-0.164	-0.830***	-0.197
	(0.0763)	(0.0921)	(0.0868)	(0.164)	(0.214)	(0.292)	(0.279)
Log(density)				2.02e-05***	4.76e-05***	3.94e-05***	4.76e-05***
*Scale				(6.50e-06)	(1.33e-05)	(1.41e-05)	(1.48e-05)
Log(diversity)		0.148	0.154	-0.0222	0.239	-0.220	0.199
		(0.198)	(0.189)	(0.222)	(0.194)	(0.245)	(0.239)
Log(diversity)				-4.66e-06	3.98e-06	5.80e-07	4.57e-06
*Scale				(5.30e-06)	(7.77e-06)	(5.79e-06)	(7.68e-06)
Log(MP)		0.0682	0.0496	0.0663	0.00875	0.182	0.0269
		(0.162)	(0.144)	(0.159)	(0.0490)	(0.196)	(0.0793)
Log(MP)				2642	44173**	12965	44906*
*Space				-3748	-19846	-8453	-23073
						[0.125]	
Spe		0.0127	0.138	-0.629	0.270	-0.692	0.249
(Subs.)		(0.669)	(0.629)	(0.676)	(0.728)	(0.589)	(0.737)
Spe		0.375	0.401	-0.358	0.427	-0.326	0.411
(Sales)		(0.578)	(0.638)	(0.704)	(0.899)	(0.579)	(0.905)
Spe		-0.141	-0.117	-0.897*	-0.536	-0.678	-0.533
(Non Agr.L)		(0.466)	(0.573)	(0.507)	(0.805)	(0.484)	(0.792)
						[0.161]	
Spe		0.531	0.654	-0.0966	0.567	-0.310	0.569
(Agri L)		(0.514)	(0.485)	(0.678)	(0.829)	(0.648)	(0.846)
			[0.178]	(0.887)	(0.494)	(0.632)	(0.501)
Spe		0.172	0.0691	-0.322	-0.159	-0.0753	-0.131
(bus)		(0.342)	(0.507)	(0.385)	(0.716)	(0.325)	(0.675)
Spe		-3.259	-2.855	-4.797***	-6.785	-3.681***	-6.281
(LVST)		(3.821)	(3.713)	(0.593)	(4.508)	(0.296)	(3.988)
Log(Health)		-0.0361**	-0.0335**		-0.0262**		-0.0240**
(dist)		(0.0143)	(0.0142)		(0.0113)		(0.00998)
Log(Health)					-1286***		-1302***
*Space					(420.6)		(470.3)
Log(School)			0.0544		0.0796		0.0824
(dist)			(0.0516)		(0.0521)		(0.0587)
Log(School)					2.52e-06***		2.30e-06**
*Scale					(6.87e-07)		(9.57e-07)
Subs.	-0.0993***	-0.0832***	-0.0745***	-0.0979***	-0.0684***	-0.0912***	-0.0655***
	(0.0186)	(0.0236)	(0.0207)	(0.0163)	(0.0194)	(0.0160)	(0.0189)
Sales	-0.0654**	-0.0590	-0.0494	-0.0509	-0.0671	-0.0566	-0.0665
	(0.0333)	(0.0438)	(0.0445)	(0.0441)	(0.0445)	(0.0441)	(0.0460)
Non-agri	0.134***	0.137**	0.138**	0.148***	0.120**	0.162***	0.135***
L	(0.0366)	(0.0569)	(0.0567)	(0.0369)	(0.0508)	(0.0406)	(0.0441)
Agri	0.124***	0.109**	0.118**	0.101***	0.119**	0.0857***	0.112**
L	(0.0325)	(0.0480)	(0.0484)	(0.0307)	(0.0502)	(0.0319)	(0.0519)
Business	0.225***	0.238***	0.248***	0.212***	0.216***	0.209***	0.215***
	(0.0708)	(0.0580)	(0.0615)	(0.0558)	(0.0581)	(0.0609)	(0.0530)
LVST	0.00502	-0.00885	-0.000104	-0.00885	-0.00510	-0.000776	-0.00320
	(0.0613)	(0.0609)	(0.0661)	(0.0604)	(0.0724)	(0.0630)	(0.0760)

Table 15: Could a refugee camp generate agglomeration economies? (continued)

	(1) HHFE (2SLS)	(2) HHFE (2SLS)	(3) HHFE (2SLS)	(4) HHFE (2SLS)	(5) HHFE (2SLS)	(6) HHFE (2SLS)	(7) HHFE (2SLS)
Subs.	564.5***	551.4***	536.3***	574.3***	547.9***	580.3***	544.0***
*Space	(79.22)	(190.7)	(190.5)	(100.4)	(164.6)	(113.1)	(166.8)
Sales	130.3**	85.39	22.61	192.9***	73.03	160.1**	65.77
*Space	(66.40)	(207.4)	(173.6)	(51.06)	(181.9)	(62.58)	(190.5)
NonAgri.	504.1***	556.2***	553.0***	491.3***	598.0***	509.6***	597.9***
L*Space	(66.02)	(61.73)	(65.76)	(69.48)	(72.20)	(78.84)	(73.73)
Agri	91.82	-30.81	-31.76	37.39	-7.402	26.90	-7.847
L*Space	(130.0)	(135.8)	(139.1)	(103.6)	(137.8)	(102.3)	(136.0)
Business	-188.7***	-271.3***	-274.5***	-212.0***	-248.6***	-219.4***	-250.0***
*Space	(48.56)	(85.62)	(85.57)	(37.21)	(93.36)	(38.14)	(92.36)
LVST	67.99*	179.9***	219.6***	43.08	198.2**	68.30	203.5**
*Space	(37.52)	(55.66)	(60.24)	(49.24)	(79.43)	(56.37)	(88.93)
Subs.	-2.04e-06***	-2.31e-06	-1.80e-06	-2.74e-06***	-2.97e-06*	-2.87e-06***	-2.99e-06
*Scale	(3.38e-07)	(2.06e-06)	(1.89e-06)	(6.73e-07)	(1.78e-06)	(7.78e-07)	(1.83e-06)
Sales	1.62e-06**	2.14e-06	2.72e-06[0.172]	1.23e-06*	2.30e-06	1.73e-06**	2.37e-06
*Scale	(7.69e-07)	(2.34e-06)	(1.99e-06)	(6.47e-07)	(1.88e-06)	(8.19e-07)	(1.92e-06)
NonAgri.	-9.29e-07	-2.14e-06**	-1.88e-06**	-1.51e-06*	-2.67e-06***	-1.74e-06*	-2.81e-06***
L*Scale	(7.97e-07)	(9.02e-07)	(8.00e-07)	(8.08e-07)	(7.89e-07)	(9.47e-07)	(7.93e-07)
Agri	-3.59e-06*	-1.46e-06	-1.53e-06	-2.59e-06	-1.88e-06	-2.17e-06	-1.80e-06
L*Scale	(1.95e-06)	(2.12e-06)	(2.23e-06)	(1.62e-06)	(2.25e-06)	(1.53e-06)	(2.19e-06)
				[0.109]		[0.157]	
Business	-7.06e-07	2.81e-07	2.47e-07	-3.10e-07	5.47e-08	-3.70e-08	1.12e-07
*Scale	(8.61e-07)	(1.31e-06)	(1.33e-06)	(7.27e-07)	(1.49e-06)	(6.76e-07)	(1.45e-06)
LVST	5.48e-07	-1.32e-06	-1.80e-06	6.83e-07	-1.68e-06	4.20e-07	-1.74e-06
*Scale	(7.85e-07)	(1.15e-06)	(1.30e-06)	(9.57e-07)	(1.44e-06)	(1.02e-06)	(1.55e-06)
Log(land)	0.000215	-0.00124	-0.00179	-4.98e-05	-0.00335	0.000632	-0.00271
	(0.00684)	(0.00806)	(0.00802)	(0.00585)	(0.00797)	(0.00608)	(0.00761)
Prop.Lit.	0.134*	0.217***	0.215***	0.151**	0.227***	0.150**	0.233***
	(0.0690)	(0.0655)	(0.0677)	(0.0639)	(0.0648)	(0.0643)	(0.0659)
Size	-0.0407***	-0.0344***	-0.0346***	-0.0394***	-0.0398***	-0.0402***	-0.0391***
	(0.0113)	(0.00876)	(0.00833)	(0.0110)	(0.00763)	(0.0106)	(0.00802)
Split-off	0.118***	0.125**	0.119**	0.132***	0.105*	0.124***	0.101*
	(0.0424)	(0.0537)	(0.0586)	(0.0453)	(0.0562)	(0.0475)	(0.0608)
Migr.(Kag)	0.326	0.244	0.262	0.336	0.151	0.403	0.161
	(0.287)	(0.307)	(0.308)	(0.267)	(0.294)	(0.308)	(0.283)
Prop.	0.280	0.142	0.330	-0.0647	0.323	0.0450	0.338
Migr.(out)	(0.341)	(0.630)	(0.609)	(0.238)	(0.688)	(0.238)	(0.716)
α_t	0.0872	0.00732	0.0646	0.0839	0.200	0.148	0.203
	(0.0888)	(0.142)	(0.184)	(0.0700)	(0.204)	(0.105)	(0.211)
Obs.	3566	2872	2872	3566	2872	3540	2846
R^2	0.243	0.257	0.263	0.251	0.280	0.246	0.282
F-Test	59.56***	196.4***	495.2***	171.89***	198.8***	29.9***	8.9e+06***