Mutual Benefits of Proximity: Economic Interactions between Refugees and Hosts in Uganda

Rama Dasi Mariani[a,b,*], Furio Camillo Rosati[b] and Pasquale Scaramozzino[b,c]

[a]Department of Law, Roma Tre University

[b]Centre for Economic and International Studies (CEIS),Fondazione Tor Vergata, University of Rome "Tor Vergata"

[c]School of Finance and Management, SOAS University of London

Abstract

Refugees are mainly hosted in low-income countries, where they often remain for a long time. Therefore, it is important to assess how they integrate with the local economy and to what extent their presence can contribute to the transition to a more dynamic economic environment. Proximity between refugees and hosts might improve the welfare of both groups by increasing opportunities for mutually beneficial economic exchanges. In particular, welfare gains might be generated through the availability of a greater variety of commodities.

In this paper we propose an empirical analysis which makes use of a unique dataset covering around 80% of the refugee population living in Ugandan settlements and the cohabiting host households. The results show that proximity between groups increases the food expenditure and the variety of food consumption of both groups. We also found that exposition to inter-group interactions increases the non-food expenditure and the probability to run a farm and a non-farm enterprise by refugee households, while hosts are not crowded out from production.

Our empirical results are consistent with a theoretical model where the *love for variety* hypothesis can be used to interpret the mutual benefits arising from the interaction between hosts and refugees facilitated by geographical proximity.

Keywords: Forced Migrations; Love of Variety; Inter-Group Exchange; Distance

JEL Codes: F12, F63, I30, O19, O55

Acknowledgments: We would like to thank Nicola D. Coniglio, Stefano Gagliarducci, Anna Yeganeh Hillingdon, Anna Maria Mayda, Rebecca Pietrelli, Francesco Sobbrio, Federico Trionfetti and participants to the Internal Seminars of the University of Rome "Tor Vergata" and of SOAS University of London, to the SITES-GLO 2022 Annual Conference in Naples, to the workshop "Improving Impact in the Resilience and Social Protection Agenda in Uganda", to the IFAD 2022 conference in Rome, to the 13th International Conference and Summer School "Economics of Global Interactions" in Bari, to the International Conference on Development Economics at Aix-Marseille Univeristy, and to SITES-GLO 2024 Annual Conference in Naples for valuable comments and inputs. We are extremely grateful to Marco d'Errico for his unvaluable work of data collection as previous Team Leader of the Agrifood Economics Division (ESA) at the Food and Agriculture Organization of the United Nations. Any errors or omissions are the sole responsibility of the authors.
Funding declaration: Financial support from the Italian Minister of University, Grant PRIN 2017 (Prot.2017KHR4) is gratefully acknowledged.
Declarations of interest: none.
*Corresponding author. Email address: ramadasi.mariani@uniroma3.it. Postal address: via Ostiense, 161/165 00154 Rome (Italy)

1. Introduction

A large number of refugees is hosted in low-income countries, where they often remain for a long time. This has raised concerns about how the refugee communities can support themselves in the long run, especially as aids from national and, particularly, international organizations are far from sufficient and over time attention tends to be diverted to new emergency situations. It is therefore crucial to evaluate how effectively refugees can integrate with local communities to develop economic activities to support their long-term well-being. Importantly, this integration should occur without harming local communities and ideally generating benefits for them, in order to prevent economic or social tensions.

One possible mechanism through which the presence of refugees can generate a (at least partly) self-sustaining economic environment is the possibility that the arrival of refugees and their interaction with the local communities generate incentives to an expansion and a diversification of economic activities.

This paper investigates the possible gains arising from market exchanges between refugee and host households in Uganda. Ugandan settlements are a model of progressive approach to refugee hosting: they are incorporated into a more general national development plan, and the delivery of essential services and infrastructure is intended to benefit both host and refugee households. Moreover, in Uganda, refugees are free to work and are not secluded in camps but integrated with the host communities: such conditions are likely to facilitate market interactions due to their physical proximity. This allows us the explore the role of proximity between host and refugee households in generating profitable exchanges. A topic that, to the best of our knowledge, has never been studied before.

There is evidence, especially from middle-income countries like Türkiye and Jordan, that letting refugees free to integrate and to participate to the labour market can ease the burden on public (national or international) resources by making the refugees self-reliant. This, of course, potentially has an impact on local communities. However, in the case of middle-income countries, Tumen (2016) finds small but statistically significant losses in informal employment among natives in Türkiye associated with a reduction in the price level in the two years following the mass inflows of refugees from Syria, while Fallah et al. (2019) do not identify any significant impact of Syrian refugees on natives' labour market outcomes in Jordan. Aracı et al. (2022) complement these results by showing that the level of local development is key in determining the size of the effect.

In low-income countries attention has been paid predominantly, if not exclusively, to the impact on local communities. Among others, Alix-Garcia and Saah (2010), Alix-Garcia et al. (2018), Coniglio

et al. (2023), Kreibaum (2016), Maystadt and Verwimp (2014), Maystadt et al. (2019), Taylor et al. (2016), Tsuda (2022), Walelign et al. (2022) and Zhou et al. (2023) have shown that in most cases the presence of refugees has a non-negative (and often positive) impact on local communities who live relatively close to the camps.[1] More specifically, in a very recent paper Kadigo and Maystadt (2023) look at the impact of the presence of refugees on the welfare of the host households in Uganda using panel data at district level. They show that the welfare of the hosting-community households improves thanks to the presence of the refugee settlements and that the group benefiting the most is that of farmers who are able to sell their products in the market.

We argue that these findings have two limitations. The first is that they only refer to the impact on the hosting community but do not consider whether the interactions between the two groups also benefit refugees. The second is that they fail to disentangle the overall effect of refugee presence – due to the benefits deriving from the infrastructures, services and programs made available by the support agencies – from the creation of a more favourable and dynamic economic environment – due to the direct interaction of the host communities with refugees.

More in detail, most of the existing literature on the impact of refugees has been focussing on hosts' welfare in context where refugees are hosted in camps or camp-like structures with limited possibilities of economic interactions with the host communities. Otherwise, like in the case of Kadigo and Maystadt (2023), the available data do not allow to disentangle the channels of the estimated effect. Therefore, the estimated effects are also due to the expansion of infrastructures, social protection programs, and increased demand generated by the need to provide goods and services to the refugees.

In this paper we analyse whether hosting refugees not in camps but in large settlement areas where host communities cohabit, and where refugees are free to be involved in spontaneous economic activities and to interact with the local population, generates benefits for both hosts and refugees due to the possibility of direct economic exchanges. In this way we assess to what extent such an approach to refugee hosting produces additional benefits with respect those analysed in the literature and can contribute at the same time to refugees' self-reliance and hosts' welfare.

d'Errico et al. (2022) have shown that host communities benefit from the presence of refugees because of such direct interaction with the refugees. In this paper we focus more broadly on how the interaction between the hosts and refugees' communities can generate a market creation process which improves the living standards of both, and which ensures that refugees can, at least partly, self-sustain in the long run. We frame our analysis within a *love of variety* approach to identify the

[1] For very comprehensive surveys about the impact of refugee presence on hosting countries see Ruiz and Vargas-Silva (2013) and Verme and Schuettler (2021).

economic mechanisms at the base of beneficial effects for both hosts and refugees of the arrival of the latter (a simplified formal model is presented in the Appendix). Therefore, our work complements and extends the analysis of Kadigo and Maystadt (2023) by looking at the mechanisms through which the documented improvement takes place. In particular, we aim to show how, by creating new opportunities for mutually beneficial exchanges, proximity between hosts and refugees can increase the self-reliance of refugees, whilst also improving the welfare of the hosts.

We focus our analysis on Uganda. This country hosts the largest number of refugees in Sub-Saharan Africa and has adopted one of the most progressive approaches towards integration. Refugees in Uganda are granted almost the same rights as natives and are free to work and to start a business. Refugee settlements are not regulated closed enclaves but are open to the creation of economic activities and to interaction with local communities.

Refugees arrive in the host countries mainly endowed with their human capital only, bringing with them their own traditions in terms of goods and of sector of activity specialization. We hypothesize and provide descriptive evidence that the presence of refugees enriches the offer of goods and services available in the market and that, similarly, the households benefit by the opportunity of supplying such goods and services in a broader market. Similar to a *love of variety* model of international trade, where access to a wider range of goods enhances welfare and production in both trading countries, the opening up of exchange possibilities between refugee and hosting households leads to similar results.

Transportation costs, moreover, have an important role in this class of models through their effects both on the relative price and on the overall price index. Given the setting of our analysis, the transportation costs will be assumed to be proportional to the physical distance among the households, so that physical proximity is going to play a central role in our analysis.

We are able to empirically assess the role of proximity between households since we can make use of a unique dataset which collects a variety of information on refugee and host households living in the area of the settlements, covering more than 80 per cent of the population of the Ugandan settlements.

In low-income countries, and especially in Uganda where refugees are hosted in remote areas, access to market is difficult for both lack of formal markets and the difficulties in transportation. For these reasons, the possibility of benefiting from the interaction with other households depends on their proximity (easily calculated as all households in the sample are geo-localised) which, as mentioned, will proxy for the transportation costs.

Our estimates indicate that the lower the transportation costs, *i.e.* the closer the households, the higher is their level of consumption and the broader the set of goods they consume. Data limitation restricts

our analysis to food expenditure, which however represents the largest share of expenditure for poor households in rural areas. On the production side, we observe an increase of non-farm activities mainly due to new micro enterprises created by refugees.

Our identification strategy relies on two crucial assumptions: (1) households do not sort themselves according to the distance to the other group; and (2) the distance between households is not correlated with the access to other services or benefits provided by national and international organizations. Regarding the first assumption, we show that the distance between the households can be considered as exogenous as the refugees are assigned to a specific location by the authorities in charge of them and that by far the vast majority of the host households did not move during their lifetime. The second assumption is necessary to ensure that we capture the effect of the direct interaction of households independently from those due to the delivery of services provided in the settlements. We address this second concern mainly by showing that the distance among households is not related to their probability of accessing the services available in the area.

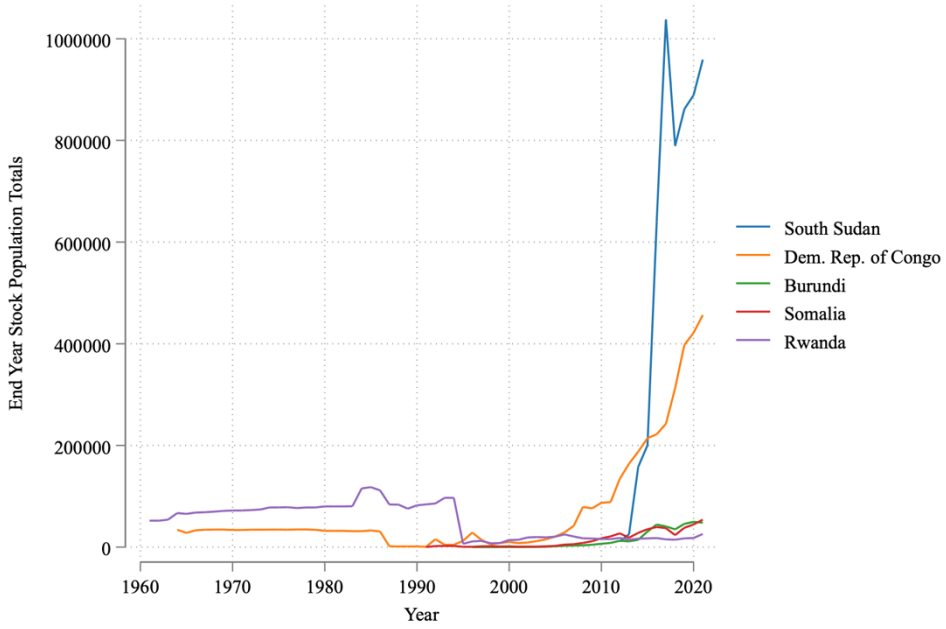
The paper is organised as follows. In section 2 we describe the historical and institutional background of Ugandan settlements. In section 3 we describe some stylized facts which motivate the analysis. Section 4 details the empirical analysis and its results, while section 5 show the robustness tests. Section 6 concludes. In Appendix A we develop a theoretical model that frames and guides the interpretation of the empirical results.

2. Refugees in Uganda

The first forced migration flows to Uganda started with the decolonization process of Sub-Saharan Africa, initiated in the 1960s. Displaced people originated mainly from Rwanda and were, nevertheless, in small numbers.

Grievous conflicts have burst in the region after 2000. Starting in the late 2010s, hundreds of thousands have fled from the conflict in Kivu in the Democratic Republic of Congo (DRC) and arrived in Uganda. The civil war in South Sudan, begun in 2014 and officially ended in 2020, has brought about the largest influx of refugees and, nowadays, South Sudanese is the most numerous group among the refugee population of Uganda (see Figure 1).

Figure 1: Refugees and Asylum Seekers in Uganda (1960-2020)

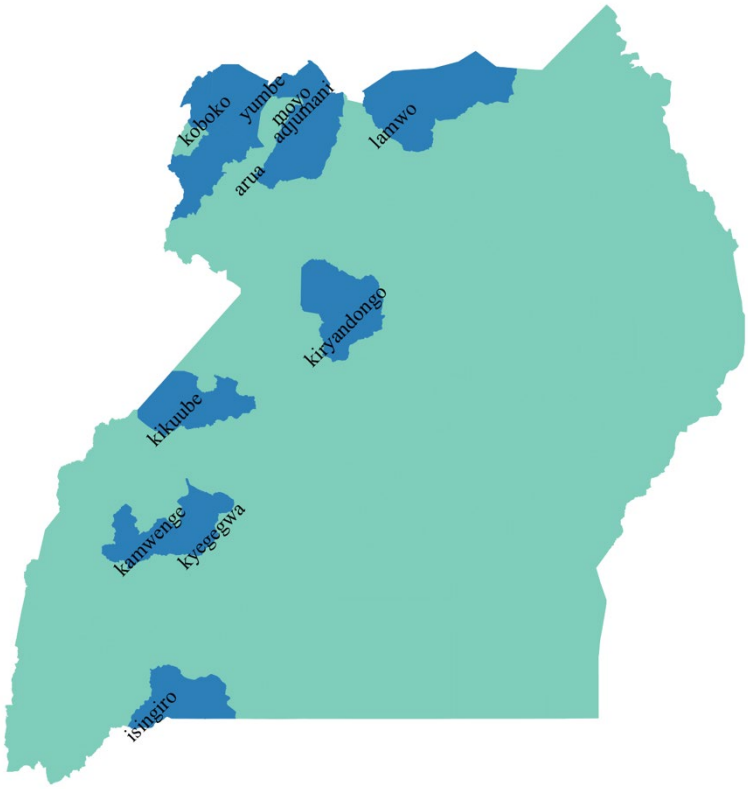


Notes: Authors' elaboration on UNHCR data (accessed online in July 2022, https://www.unhcr.org). End-year stocks are computed as the sum of refugees (or in a refugee-like situation) and asylum seekers by country of origin.

Refugees and asylum seekers living in Uganda are currently more than 1,7 million (UNHCR, 2024) and represent the largest displaced population in Sub-Saharan Africa. They mainly come from the bordering countries: in addition to South Sudan and DRC, they mostly come from Burundi, Somalia and Rwanda.

Refugees and asylum seekers are hosted in settlements – managed by the Office of Prime Minister (OPM) in cooperation with the United Nation High Commissioner for Refugees (UNHCR) – where they co-exist with the local community. The settlements are located in the Northern and in the Southwestern districts and most of them are proximate to the borders with the neighbouring countries. Currently, there are 13 settlement-hosting districts (see Figure 2).

Figure 2: Settlement-Hosting Districts



Notes: Authors' elaboration on UBOS data (accessed online in October 2019, https://www.ubos.org). In blue settlement-hosting districts – *i.e.*, Adjumani, Arua, Isingiro, Kampala (not shown in the map), Kamwenge, Kikuube, Kiryandongo, Kyegegwa, Koboko, Lamwo, Madi-Okollo (originally included in the Arua district), Obongi (carved out from Moyo district in 2019) and Yumbe.

The organisation of the settlements is seen as a model of a progressive approach to refugee governance. The 2006 Refugees Act and the 2010 Refugees Regulations have established the following fundamental pillars. First, there are no restrictions in terms of number or of origin and all asylum seekers, irrespective of their nationality or ethnic affiliation, are allowed to enter the country and to receive assistance. Second, all refugees are granted freedom of movement and the right to seek employment. Third, refugee households are provided with a plot of land for their own (mainly agricultural) use.

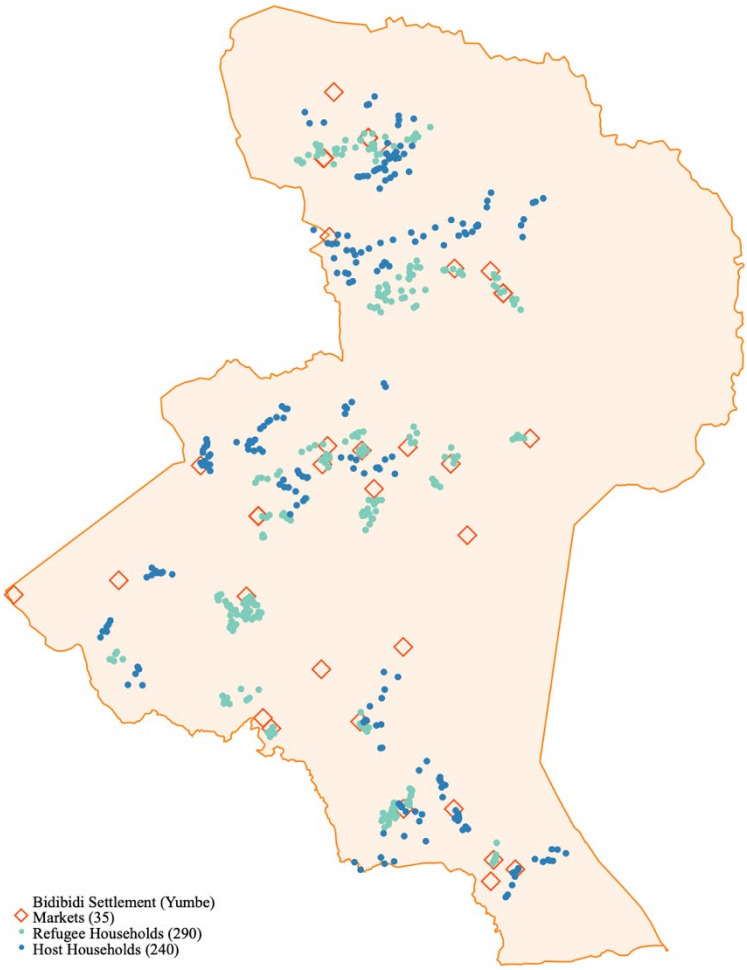
More in detail, once asylum seekers enter the country, they are received and registered in the hotspots managed by the UNHCR distributed along the borders. Thereupon, asylum seekers can ask for assistance, in which case they are allocated to a settlement, or they can choose not to live in the settlements, in which case they do not receive the assistance of the national and international institutions. Once refugees arrive in a settlement, they are provided with essential utensils and assigned to a specific plot of land.

The latter point implies that, within the settlements, the location of the recipients of the assistance is decided by the institutions, namely OPM and UHNCR. However, the settlements should not be seen as closed camps, containing refugees and services without room for interaction with the hosting communities.

Bjørkhaug (2020) describes the historical factors that have caused hosting-community households to reside within the areas of the settlements. Focussing the qualitative analysis on Nakivale Settlement, in the Southwestern District of Isingiro, Bjørkhaug (2020) documents that the area was already inhabited by autochthonous pastoralists and farmers and that the settlement has developed around them.

As an example of the spatial extension and spread of the Ugandan settlements, Figure 3 shows a detailed map of the Bidibidi Settlement in Northern district of Yumbe.

Figure 3: Refugee and Host Households Sampled in Bidibidi (Yumbe)



Notes: Authors' elaboration with RIMA (FAO) and UBOS data (accessed online in October 2019, https://www.ubos.org). As of the 2019, Bidibidi settlement covered 250 km^2 of the Eastern half of the Yumbe district.

The Bidibidi settlement was established in 2016, primarily to host the increasing inflows of South Sudanese arriving from the North. Initially formed of small agglomerates in rural areas, Bidibidi has rapidly transformed into a dynamic place where a plethora of economic exchanges occur daily.

Several public services or infrastructures have been built – hospitals, schools, mosques and churches, country roads and means of transports among others. Figure 3 shows the location of the markets, along with the refugee and hosting-community households sampled in the data that we use for the empirical analysis. The potential of interactions between households is easily grasped from the illustration showing how both groups of households – refugee and host – are intermingled and all share similar possibilities of access to the settlement infrastructures. The latter point is crucial because it exemplifies (and this will be formally tested later on) how access to markets and other services is unlikely to be confounding the effects due to the proximity between hosts and refugees, as it is a factor common to all the households in the sample.

There are some, albeit few, qualitative studies focused on the economic activities that take place in the settlements in Uganda (see *e.g.*, Alloush et al., 2017 and Betts et al., 2014). They show that refugees bring with them different experiences and skills which are productively applied in their new reality. They engage in the production of goods and services (agricultural or not) that add to and differ from the products locally supplied. At the same time, refugees benefit from the presence of hosts as they can acquire goods and services which they could not otherwise obtain.

Moreover, the different communities of refugees have themselves different specialization. For example, while Congolese and Rwandan specialize mainly in agricultural works (own or for a wage), Somali appear to be mainly involved in trade and food and South Sudanese beside agriculture specialize in brewery (Betts et al., 2014).

It appears, therefore, that the arrival of refugees does not affect the economy of the areas of settlement only through changes in the level of the supply and the demand of goods and labour. The inflow of refugees expands the range of goods and services available to the local community, while access to the host community allows refugees to obtain goods which they would not be able to produce under autarky.

3. Descriptive Analysis

3.1 The Data

With the objective of assessing economic and social needs of households living in refugee-hosting districts of Uganda, the Food and Agriculture Organization of the United Nation (FAO) – in coordination with the Office of the Prime Minister (OPM) – carried out a survey that targets the refugee households and the hosting-community households who live within the area of the Ugandan settlements. The data collection started in 2017 and, as of 2020, the survey was implemented in three rounds. The survey covered all the districts,[2] located in the North-West and in the South-West of the country, that host a refugee settlement.[3]

In 2020, the survey re-interviewed all the refugee and host households interviewed in the first wave (administered in two times between November 2017 and April 2018) with the addition of the households living in Koboko, visited for the first time during the second wave of the survey (administered in December 2019). The survey was based on a two-stage cluster sampling method, with the settlement as the Primary Sampling Unit (PSU) and the households as the Secondary Sampling Unit (SSU). The selection of the households was random and based either on lists provided by the local authorities or by walking through the settlement. The survey was designed to be representative of refugee and hosting-community households that live in the area of Ugandan settlements.[4] The final sample consists of 14,810 observations.

The survey questionnaire is broad and includes questions about socio-demographic details of the households, as well as the single household members, durable and agricultural assets, access to services (*i.e.*, schools, health facilities, transports, churches or mosques, and markets), expenditure and loans, agriculture and livestock production, consumption, coping strategies, assistance received, social cohesion, employment and business activities.[5] The main limitation of the survey is that the survey questions changed slightly from wave to wave and sometimes information is not available for all the observations of the sample.[6] We decided to always keep as many observations as possible for each regression so as to gain more power in our fixed-effect estimates.

[2] Arua, Yumbe, Moyo, Adjumani, Lamwo, Kiryandongo, Kyegegwa, Kamwenge, Isingiro, Kikuube and Koboko.
[3] All the districts have only one settlement, but Adjumani, which hosts 18 settlements, Isingiro, which hosts 2 settlements, and Arua, which hosts 2 settlements.
[4] For more details see FAO and OPM (2018) and FAO and OPM (2019).
[5] We provide the summary statistics of the variables used in the empirical analysis in Appendix C.
[6] For each estimate we specify the years for which information are missing.

To collect the data, Computer Assisted Personal Interviewing (CAPI) technology has been used with the aids of digital tablets, so that the household data are georeferenced. This allows us to know the exact location and to compute the distance of each household to all other households in the sample, as well as to the services and infrastructures of the settlement.

Table 1 shows, separately for refugee and host households, some summary statistics about the spatial information available in the data.[7]

Table 1: Spatial information (First Round of the Survey)

	Hosts				**Refugees**			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Distance to Refugees (1-NN)	1.572	1.306	0	6.998	0.144	0.129	0	1.039
Refugees in 1-km Radius	2.895	5.739	0	43.00	19.80	15.75	0	81.00
Distance to Hosts (1-NN)	0.190	0.175	0	1.080	1.633	1.259	0	6.867
Hosts in 1-km Radius	14.96	14.73	0	74.00	2.248	4.554	0	68.00
Distance to a Petty Market	1.477	1.386	0	12.00	1.133	1.267	0	13.00
Distance to Means of Transport	1.625	2.279	0	35.00	1.520	2.383	0	25.00
Distance to a Church/Mosque	1.236	1.137	0	6.500	0.735	0.846	0	6.000
Observations	1,632				2,107			

Notes: Authors' elaboration on RIMA (FAO) dataset. Only observations which GPS coordinates are correctly measured are included. The distance between refugee and hosting-community households is computed based on GPS coordinates. The distance to the settlement infrastructures is self reported. 1-NN stands for 1-Nearest Neighbour.

To measure the refugee-host proximity we borrow from a widespread method in space data, the k-nearest neighbours, first studied by Cover and Hart (1967). Nearest neighbour methods are based on the k nearest observations, with distance metrics varying according to the context and k being a number of observations chosen heuristically and usually lower than 10. We adapt this method to our context by considering the mean distance of the k out-group households closest to a given household. Fukunaga and Hostetler. (1975) have shown that the variance of finite sample estimates may be minimized by the proper choice of the local distance metrics. We argue that the best way to measure local distance in our context is the earth-arc distance – the distance between two points on a sphere – which approximates the distance to be covered on foot, with $k = 1$. We also employed other values of k (namely 5 and 10): the results are substantially unchanged, albeit less precise.[8]

3.2 Stylized Facts

During the first wave of the data collection, FAO administered a very rich module of food consumption. Specifically, households have been asked about the consumption of 72 food items, grouped in 12 categories. Each of the 12 categories contains several items considered by FAO close

[7] In Appendix C we provide more summary statistics about the socioeconomic characteristics of sampled households.
[8] Results available on request

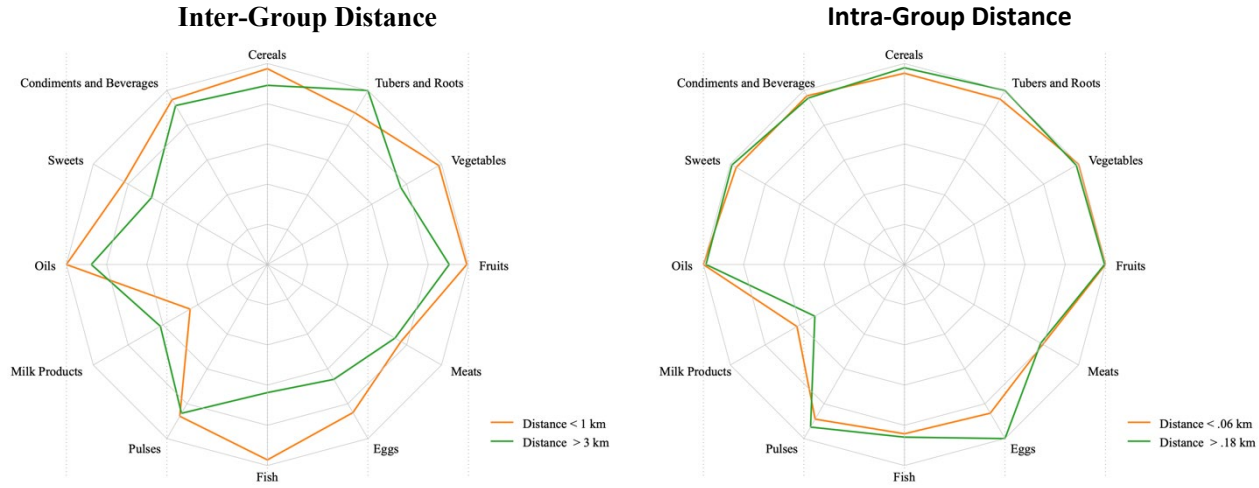
substitute of each other as they serve similar nutritional needs. For instance, in the category "Tubers and Roots" are included: Matoke, Potatoes, Yams, Cassava, Sweet Potatoes/Irish Potatoes, Other. In the category "Meats", instead, are included: Chicken and Poultry, Red Meat (Goat, Beef, Mouton), Organ Meat (Offal, Liver, Blood), Game Meat, Pigeon, Insects, Other.

We count the number of varieties consumed by the households within each category. Then we standardise the count so that it ranges from zero (no variety consumed in the category) to one (all the varieties within a category are consumed) and, as *prima facie* evidence of the benefits arising by market interactions facilitated by proximity, we compute the average number of varieties consumed by the households who live close to or far from the potential trading partners. We plot the results of these computations in Figure 4, where points at the vertices of the polygons represent the maximum number of varieties and points at the centre depict the zero-consumption level.

If we consider the intra-group distance – *i.e.*, we consider the distance of refugees from other refugees and hosts from other hosts (see the right-hand panel of Figure 4) – it does not appear to be any differences in terms of the number of varieties consumed between households who are in the first or in the fifth quintile of the intra-group distance distribution. By contrast, in the left-hand panel of Figure 4, we show the average of varieties consumed by the households who are in the first or in the fifth quintile of the inter-group distance distribution – *i.e.*, the distance of refugees from hosts and vice versa. Left-hand side of Figure 4 shows that households who live within a distance of less than 1 kilometre have a more varied consumption than households who live more than 3 kilometres away, as the average number of varieties consumed within each category is larger for almost all the categories.

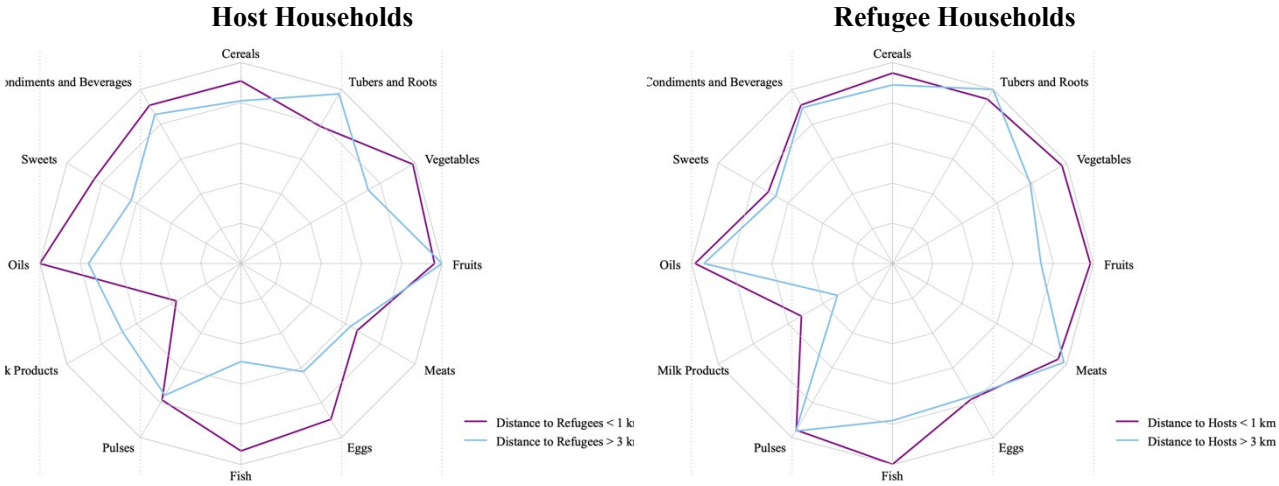
In Figure 5, we consider only the inter-group distance, and we compute separately by refugee status the number of varieties consumed on average by the households. Again, households who live relatively close to the households of the other community present a relatively more varied food consumption. This pattern is more evident for hosts (left-hand panel of Figure 5) than for refugees (right-hand panel of Figure 5).

Figure 4: Between-Group and Within-Group Distance and Variety of Food Consumption



Notes: Authors' elaboration on RIMA (FAO) data. The two graphs are drawn considering the quintiles of the specific distribution, *i.e.* the inter-group distance distribution in the left-hand panel and the intra-group distribution in the right-hand panel.

Figure 5: Refugee-to-Hosts and Hosts-to-Refugees Distance and Variety of Food Consumption



Notes: Authors' elaboration on RIMA (FAO) data. The two graphs are drawn considering the quintiles of the distribution of the inter-group distance.

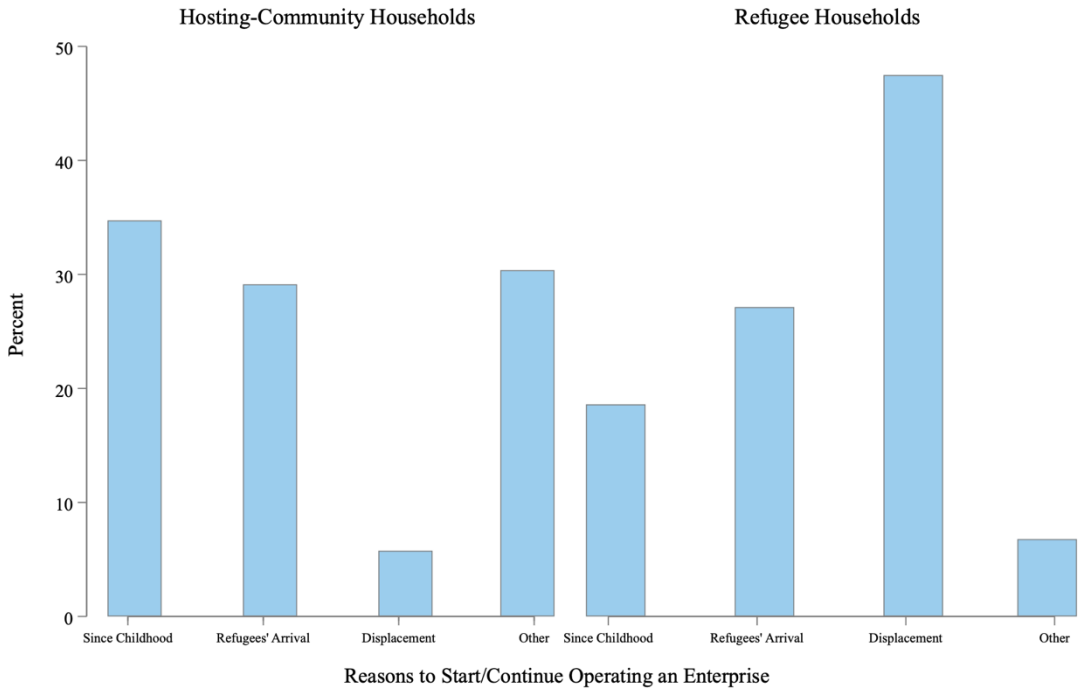
The fact that Agencies might provide part of their support in kind, and in particular by distributing specific food items, does not alter our conclusions. While food distribution might change the varieties available to refugees (and possibly to hosts), there is no reason to believe that this is affected by the distance between households other than because proximity increases exchanges: exactly the point we are trying to make.

Finally, as illustrative evidence of the improvement brought about by the interactions with newly arrived refugees on the production side, Figure 6 shows that many refugees households started to be

involved in non-farm activities as a consequence of their displacement or of the arrival of other refugees. The host households also appear to have started a non-farm activity in response to the refugee arrival, although to a more limited extent. Altogether, this points to the fact that the inflow of refugees was associated with an increase in the economic dynamism in the non-farm sector, especially for the refugees.

We conceptualize these stylized facts in a theoretical model inspired by the results of the literature on international trade – formally developed in Appendix A – and we further assess these stylized facts with a more rigorous empirical analysis in the following section.

6: Reasons to Start/Continue Operating an Enterprise (2020)



Notes: Authors' elaboration on RIMA (FAO) data. The graph is computed by using information from the question "Why have you started operating / continuing the [ENTERPRISE]?" which answers are 1 = "Working on the same enterprise since childhood" 2 = "New opportunities after refugees' arrival" 3 = "New opportunity after displacement" 4 = "Other [specify]".

4. Empirical Analysis

4.1 The Empirical Strategy

The arrival of refugees produces to a certain extent effects similar to that of opening an economy to international trade. As described by Betts et al. (2014), not only do refugees bring with them different

cultures, but also offer goods and services which are different from those offered by the hosts and which are appreciated by them.

The literature on international exchanges predicts that a reduction in transportation costs and a larger size of the market enhance the level and composition of consumption and increase the production of goods and services. With this in mind, we can reasonably expect that the exchange between the two groups is affected by the trading costs and consequently limited by the geographical distance. Therefore, we look at the impact of the distance on the level of consumption. To evaluate if and how households' consumption changes in response to the opportunity of exchange between communities, we estimate the following linear regression model:

$$Y_{i,j,z}^{t} = \alpha + \beta \cdot Dist_{i,j,-z} + \boldsymbol{\delta} \cdot \boldsymbol{X_i^t} + \varphi_j + \rho_t + \varepsilon_{i,j}^{t} \quad (1)$$

$Y_{i,j,z}^{t}$ is the outcome of interest for household i, living in district j, belonging to the community $z = \{h; r\}$: host or refugee, observed during round t. $Dist_{i,j,-z}$ represents the distance between host and refugee households, computed as the earth-arc distance between two points using the nearest neighbour method, detailed in section 3. $\boldsymbol{X_i}$ is a vector of household controls – namely, the number of household members, the number of female adults and male adults, the average years of education of adult household members, the age and gender of the household head, if the household head did ever move from the current residence (dummy), and the density of the own-group population proxied by the number of households within the radium of 1 kilometre.[9] The latter controls for the potential confounding effect of more populated areas where households live closer. Nonetheless, when we compute the number of households in a radium of 1 kilometre, we do not consider the households of the other group, since proximity to the other group (and therefore the presence of the other-group households in the proximity) is the effect that we aim at estimating. φ_j and ρ_t are district and round fixed effects respectively. $\varepsilon_{i,j}^{t}$ is the idiosyncratic error term. The estimation method is Pooled Ordinary Least Squares and standard errors are computed with a Bootstrap procedure.

With a similar approach, we also look at other outcomes beside consumption. In particular, the variety of consumption, the probability of running a farm and a non-farm enterprise, and the revenues from farm and non-farm business.

The estimation strategy we apply allows us to estimate the effect of the distance between the two groups, controlling for household, district and year characteristics that might all confound the results.

[9] See Appendix B for a further description of the data and the summary statistics of the control variables.

Nonetheless, when implementing this strategy, one must be careful to address a number of potential concerns.

First, the location of the households, and consequently the distance among them, could be endogenously determined by households' residence choices. Both host and refugee households might in principle choose where to establish their dwelling to stay closer to the other group and benefit from smaller trading costs. However, we have information on household head's date of arrival in the current residence and this allow us to assess whether host households have moved following the arrival of refugees. In fact, more that 90 percent of host household heads never moved from where they were born and of the 10 per cent that moved only a small number moved after the arrival of the refugees.[10] On the other hand, according to the Ugandan dispersal policy described in section 2, refugees' location is decided by the institutions that manage the settlements according, primarily, to family reunification and the settlement capacity.[11]

Furthermore, to rule out concerns regarding a potential endogeneity of the distance between households, due to a self-selection of host households or to a non-random refugee assignment design, we perform a balance test on households' characteristics, on the probability to receive assistance and on the access to settlement infrastructures conditional on the distance to the other group. The results are shown in Table 2 where we divide the households on the basis of their distance to the other group, using the median of the inter-group distance distribution as threshold. As the Table 2 shows, hosting-community households who live far away from refugees are not systematically different from hosts who live close by. Refugee households who live far away from the hosts do not have different characteristics with respect to those who live close to them, with the exception of a small but significant difference in household size. If anything, we would expect this to attenuate our estimates, as larger households are more likely to diversify their consumption and, in any case, we control for such a characteristic in the estimations.

Another concern regards the fact that the effects of the proximity to the refugees on hosts' outcomes might be correlated, and therefore be confounded, with the proximity to settlement's service centres. If this were the case, the estimated coefficient would not only represent the effect of the interaction between refugees and hosts, but it would also capture the effect of the services and infrastructures provided by the national and international agencies operating in the area. However, the sample consists only of host households who live within the settlements, all of whom have access to the services and the infrastructures related to the settlements, irrespective of how close they are to refugee

[10] See the Robustness section for a more details.

[11] To check the degree of homogeneity of the settlements, which is one of the main criteria to allocate refugees, we show the origin of refugees by district. See the Robustness section for more details.

households. Table 2 also shows that the relative location of hosts is not correlated with the probability to receive assistance, especially cash transfers, and to the settlement infrastructures.

Table 2: Balance Test on Households' Characteristics and Probability to Receive Assistance

	Close by		**Far away**		**Difference**	
Panel A:			**Hosting-Community Households**			
Number of Household Members	6.63	[3.18]	6.62	[3.20]	0.00	(0.12)
Share of Children in the Household	0.42	[0.21]	0.41	[0.22]	0.01	(0.01)
Share of Male Adults	0.48	[0.20]	0.48	[0.21]	0.00	(0.01)
Average Years of Education of Adults	4.16	[2.69]	4.33	[2.71]	-0.17	(0.10)
Household Head's Age	45.2	[14.8]	44.80	[14.61]	0.46	(0.55)
Female Household Head	0.27	[0.44]	0.25	[0.43]	0.02	(0.02)
Years Since Arrival	12.9	[13.2]	13.83	[13.49]	-0.88	(1.31)
Stayer Household Head	0.89	[0.31]	0.79	[0.41]	0.10***	(0.01)
Probability to Receive Cash Transfers	0.50	[0.50]	0.51	[0.50]	-0.01	(0.02)
Probability to Receive In-kind Transfer	0.06	[0.23]	0.06	[0.23]	-0.00	(0.01)
Formal Transfers (Thousand ugx)	57.8	[292]	74.8	[430]	-170	(136)
Informal Transfers (Thousand ugx)	56.1	[666]	53.5	[225]	267	(241)
Distance to a Petty Market	4.81	[52.4]	2.90	[37.2]	1.91	(1.69)
Distance to a Church/Mosque	1.84	[13.2]	3.99	[52.3]	-2.15	(1.42)
Distance to Means of Transport	7.24	[73.9]	8.67	[82.9]	-1.43	(2.91)
Observations	1453		1453		2906	
Panel B:			**Refugee Households**			
Number of Household Members	6.05	[3.08]	6.48	[3.23]	-0.43***	(0.10)
Share of Children in the Household	0.46	[0.23]	0.47	[0.22]	-0.01	(0.01)
Share of Male Adults	0.43	[0.27]	0.41	[0.26]	0.01	(0.01)
Average Years of Education of Adults	4.03	[3.03]	4.02	[3.02]	0.01	(0.10)
Household Head's Age	39.2	[13.2]	39.74	[13.5]	-0.55	(0.43)
Female Household Head	0.51	[0.50]	0.54	[0.50]	-0.03	(0.02)
Years Since Arrival	3.09	[3.35]	3.13	[4.07]	-0.04	(0.12)
Probability to Receive Cash Transfers	0.53	[0.50]	0.51	[0.50]	0.02	(0.01)
Probability to Receive In-kind Transfer	0.05	[0.22]	0.06	[0.24]	-0.01	(0.02)
Formal Transfers (Thousand Ugx)	98.5	[845]	81.9	[282]	16.6	(0.01)
Informal Transfers (Thousand Ugx)	55.0	[286]	495	[264]	54.5	(200)
Distance to a Petty Market	8.48	[83.4]	6.05	[67.4]	2.43	(118)
Distance to a Church/Mosque	3.11	[45.1]	3.00	[45.0]	0.11	(2.40)
Distance to Means of Transport	13.0	[104]	9.75	[89.2]	3.23	(1.43)
Observations	1990		1990		3980	

Notes: Authors' elaborations with RIMA (FAO) data. Households are divided on the basis of the median of the distribution of the distance to the other group.

The result of the balance test is consistent with Figure 3 in section 2 which is an illustration of the location of the households living in a settlement, along with the official markets of a settlement, and

which shows that living close to or far from the refugee households does not imply that hosts also live close to or far from the markets, and vice versa.

In section 5.4 we also show the results of a robustness analysis aiming at assessing the potential confounding effect of the presence of settlement services. For that reason, we correlate the outcomes of the main analysis to the distance to the market. We also check whether the probability of being involved in exchanges with traders is affected by the distance between households.

4.2 Empirical Results: Consumption

We first look at the demand side and the outcomes of interest are the food expenditures and the variety of consumption in twelve categories of goods – *i.e.*, cereals, tubers and roots, vegetables, fruits, meats, eggs, fish, pulses, milk products, oils, sweets, condiments and beverages. For each household, we compute the total expenditure on food, measured in Ugandan Shillings (adjusted for price changes across waves), and the number of food varieties consumed. The results of the estimates are presented in Tables 3 and 4.

Table 3: OLS Regressions for Food Expenditure

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	-3,374*** (302.4)	-2,772*** (609.0)	-1,611*** (581.0)
Observations	6,685	2,831	3,854
R-squared	0.287	0.267	0.324
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Sample Mean	41,953	52,328	33,941

Notes: Authors' elaborations with RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the food expenditure over the past 7 days and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. The dependent variable is available for the first and second round. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the number of male adults, adults' average years of education, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * $p < .1$, ** $p < .05$ *** $p < .01$

The results of Table 3 indicate that consumption is higher for those households who live relatively close to households of the other group. This is true both when we consider all households together and when we distinguish between hosts and refugees. In general, a reduction in 1 kilometre[12] of the

[12] When interpreting the results it is important to keep in mind the remote context of the analysis, where distances greater than 1 kilometre are covered on foot every day.

inter-group distance is associated to an increase in the food expenditure by 3,374 Ugandan Shillings (corresponding to about 8 percent of the sample mean).

The results shown in Tables 4 point at a significant negative effect of the distance to the other group on consumption variety: a reduction in 1 kilometre in the distance between hosts and refugees increases the food-consumption variety index by 0.006 (corresponding to about 1 percent of the sample mean). A similar effect is obtained when we consider host and refugee households separately.

Table 4: OLS Regressions for Food-Consumption Variety

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	-0.00626*** (0.00184)	-0.00468* (0.00255)	-0.00422** (0.00202)
Observations	6,745	2,855	3,890
R-squared	0.207	0.198	0.163
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Sample Mean	0.544	0.588	0.509

Notes: Authors' elaborations with RIMA (FAO) data. Models (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the number of food-consumption goods and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. The dependent variable is available for the first and second round and they are coherent with the estimation on food expenditure. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

Since we have data on non-food expenditure in the last month, we study the effect of the distance also on this outcome. However, we cannot replicate the analysis on the variety of consumption for non-food expenditure, as the data do not allow us to decompose consumption into sub-categories – *i.e.*, we do not have information on varieties of footwear like socks, shoes, boots, sandals, *etc*. Therefore, in Table 5 we show the results for the level of consumption only.

According to the estimated coefficients, a reduction in the distance to hosts by 1 kilometre is associated with an increase in refugees' non-food expenditure by 2,453 Ugandan Shillings, corresponding to 5 percent of the variable sample mean. We do not find any significant result for host households, which is a pattern that we are also going to observe in the analysis of the production side, shown in the next section.

Table 5: OLS Regressions for Non-Food Expenditure

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	-1,107 (800.7)	916.0 (1,523)	-2,501** (1,200)
Observations	3,476	1,499	1,977
R-squared	0.099	0.080	0.109
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	No	No	No
Sample Mean	54,110	68,916	42,798

Notes: Authors' elaborations with RIMA (FAO) data. Model (1) is estimated for the full population – *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the non-food expenditure over the past 30 days and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. The dependent variable is available for the first round. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the number of male adults, adults' average years of education, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

4.3 Empirical Results: Production

On the production side, we consider the probability of doing agriculture and of running a non-farm enterprise, as well as the revenues from both activities. The results for agriculture are shown in Table 6.

According to the coefficients shown in the Table 6, living 1-kilometre closer to the host households is associated with a 1.4 p.p. higher probability of refugees being involved in farming in the twelve months preceding the survey. The opposite effect emerges for the hosts, who are less likely to be involved in agriculture the closer they live to refugee households. On aggregate, we do not observe an increase in the number of households doing crop production, as the coefficient estimated for the whole population is not significantly different from zero.

We also looked at the impact of the inter-group distance on revenues from agriculture and results are reported in Table 7. Even if the coefficient is positive, it is very poorly determined. As most of the households are involved in farming it is likely that a marginal change in the extensive margin does not have an identifiable effect on the intensive margin.

Table 6: OLS Regressions for the Probability of Doing Agriculture

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	-0.00396 (0.00453)	0.0163*** (0.00544)	-0.0143** (0.00647)
Observations	3,510	1,525	1,985
R-squared	0.066	0.072	0.102
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	No	No	No
Sample Mean	0.927	0.969	0.896

Notes: Authors' elaborations with RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the household probability of doing agriculture and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. To be coherent with the estimation on non-farm enterprise, the analysis is restricted to the first wave only. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

Table 7: OLS Regressions for the Revenues from Agriculture (in Logarithms)

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	0.0101 (0.0656)	0.0347 (0.134)	-0.0117 (0.0796)
Observations	3,343	1,397	1,946
R-squared	0.163	0.093	0.246
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	No	No	No
Sample Mean	40,599	62,232	25,003

Notes: Authors' elaborations with RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the value of crop sells and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. To be coherent with the estimation on non-farm enterprise, the analysis is restricted to the first wave only. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

We now look at the changes in non-farm activities both along the intensive and the extensive margin. The results of the regressions presented in Table 8 are consistent with the stylized facts of section 3. The proximity to households of the other group increases on average the probability that a

household runs a non-farm business. This effect, however, is due to the increase in the participation in the non-farm sector by refugee households, while host households do not appear to be affected.

Table 8: OLS Regressions for the Probability of Non-Farm Enterprise

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	-0.0127** (0.00497)	0.00515 (0.0125)	-0.0213*** (0.00784)
Observations	3,510	1,525	1,985
R-squared	0.063	0.044	0.055
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	No	No	No
Sample Mean	0.355	0.439	0.220

Notes: Authors' elaborations with RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the household probability of running a business enterprise in the last week and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. Data are available for the first round. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the number of male adults, the average years of education of adults, household head's age, and two dummies indicating whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

Not surprisingly, then, the same pattern is observed when we look at the intensive margin measured by the revenues. Proximity to host households appears to increase the revenues from non-farm activities (see Table 9).

Table 9: OLS Regressions for the Revenues from Business (in Logarithms)

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	-0.142** (0.0722)	0.0402 (0.116)	-0.204*** (0.0768)
Observations	3,317	1,382	1,935
R-squared	0.055	0.046	0.054
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	No	No	No
Sample Mean	18,504	26,460	12,757

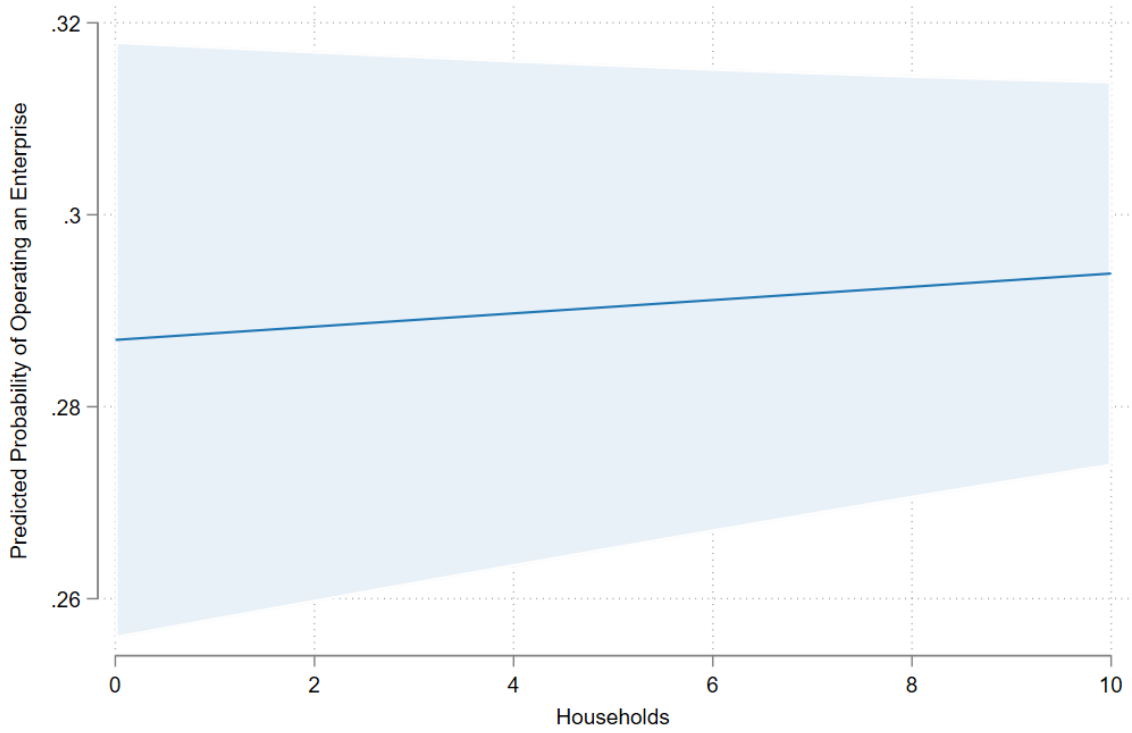
Notes: Authors' elaborations with RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the value of sales, and the main explanatory variable is the inter-group distance, measured as the earth-arc distance to the 1st-Nearest Neighbour. Data are available for the first round. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

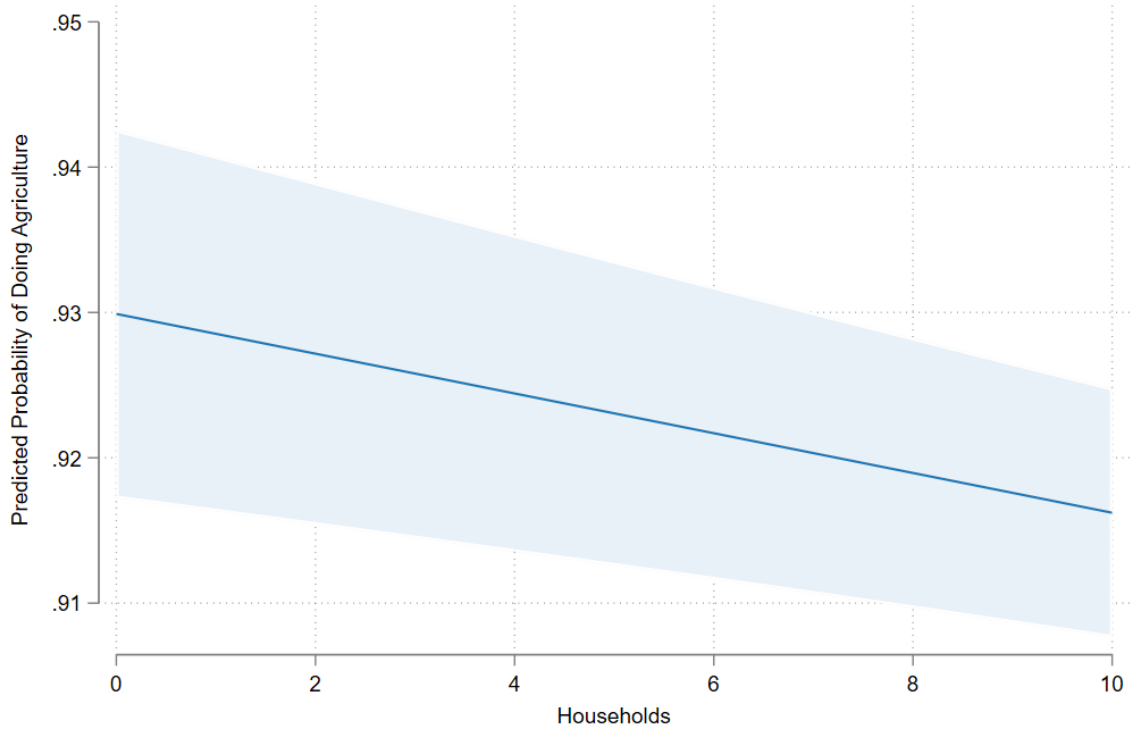
These results are consistent with those by d'Errico et al. (2022) showing that the probability to be employed in the private sector by host households' members is positively affected by the proximity to refugees.

So far, we focused on the main prediction of the theoretical model – *i.e.*, distance as the main obstacle to the trade between groups. Nonetheless, one could also check the assumptions of the theoretical model. In particular, in Appendix A, we show that the output produced by each firm is a function of the number of other firms operating in the market (see equation 14). If positive network externalities prevail, the production by a single firm grows along with the number of other firms. By contrast, if there are negative congestion externalities, firm' output is a decreasing function of the number of other firms.

In order to have a preliminary test of the type of external effect taking place among producers, we compute the marginal effect of the number of households within a radius of 1-kilometer on the probability to run a non-farm enterprise and on the probability of doing agriculture. The results are plotted in Figure 7. The marginal effect of household density is positive for the probability of doing a non-farm enterprise, but negative for the probability of doing agriculture. Nonetheless, the linear prediction does not offer a clear-cut result. This might depend on the fact that we do not have information on the number of producers, but only on the number of households in the proximity and, with more data, this could be an interesting area for further research.

Figure 7: Marginal Effects of Households' Density on Production





Notes: Authors' elaboration on RIMA (FAO) data. Marginal effects of a linear regression model estimated by OLS. The dependent variable is household's probability to run a business enterprise (upper panel) and household's probability to do agriculture (bottom panel). The main explanatory variable in both regressions is the number of households in a 1-km radius. Controls for the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence are included, as well as year and district fixed effects. Bootstrapped standard errors are computed with 50 replications. Regression outputs are presented in the Appendix.

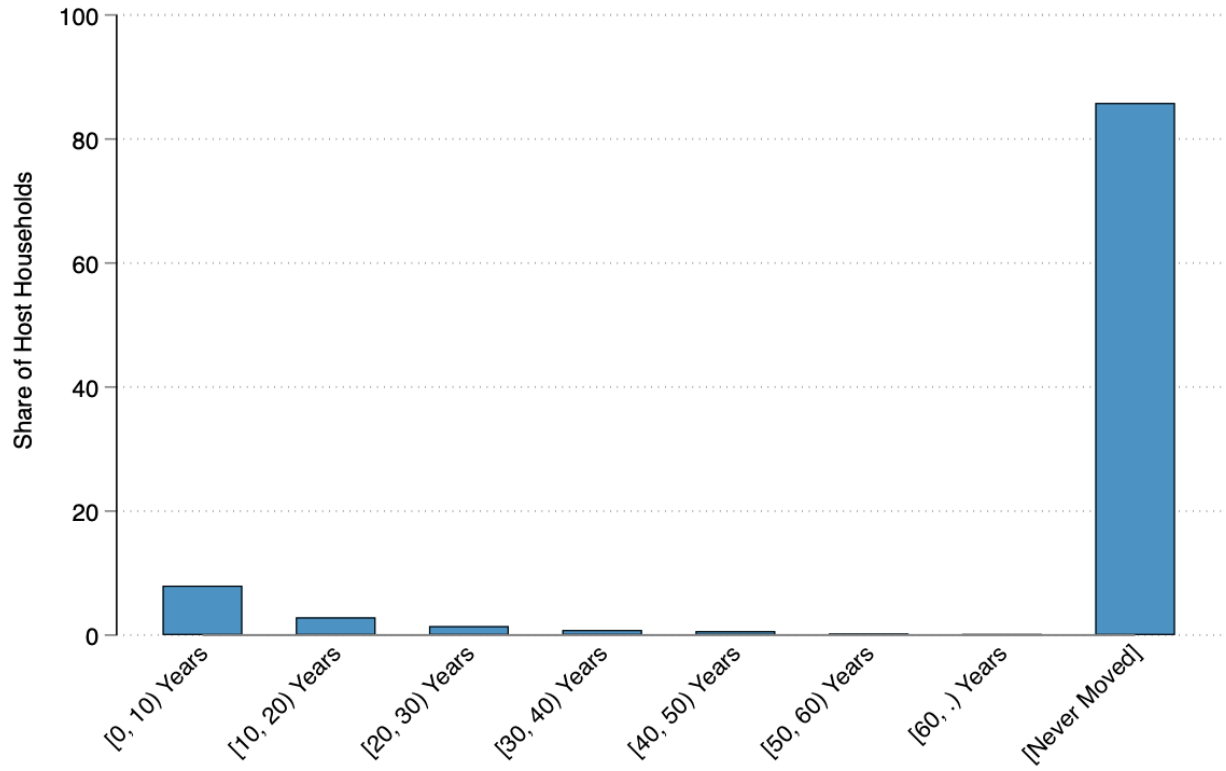
5. Robustness analysis

A concern that our empirical approach might arise is that households' location, and consequently the distance among them, might be endogenously determined by households' residence choices. Both host and refugee households might choose where to establish their dwelling to stay closer to the other group and benefit from smaller trading costs.

However, we exploit the information we have on household head's date of arrival in the current residence. This allows us to assess whether host households have moved following the arrival of refugees.

As shown in Figure 8, more that 90 percent of host household heads never moved from where they were born and of the 10 per cent that moved only a small number moved after the arrival of the refugees.

Figure 8: Years since Arrival of Hosting-Community Household Heads

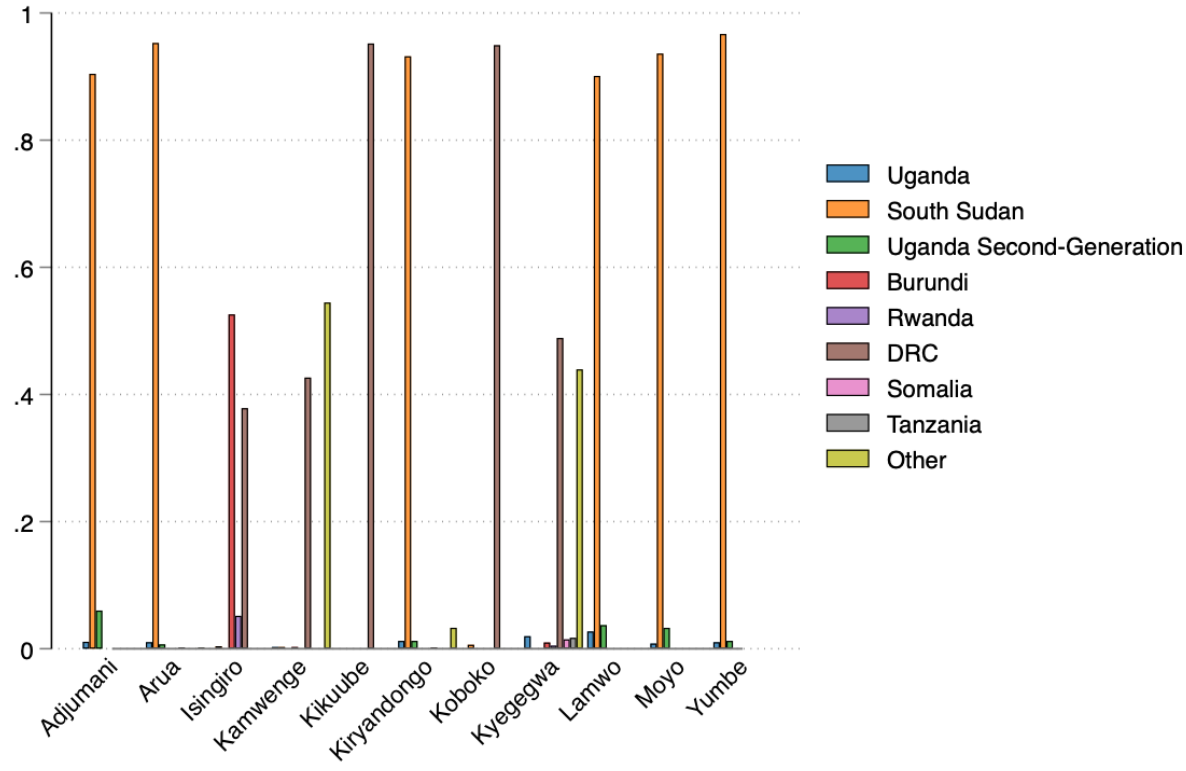


Notes: Authors' elaboration on RIMA (FAO) data

Furthermore, according to the Ugandan dispersal policy described in section 2, refugees' location is decided by the institutions that manage the settlements according, primarily, to the settlement capacity or to family reunification.

To give an idea of the degree of homogeneity of the settlements, which is one of the main criteria to allocate refugees, we show the origin of them by district (see Figure 9).

Figure 9: Origin Composition of Districts' Refugee Population



Notes: Authors' elaboration on RIMA (FAO) data

The only district that hosts two groups of different origin is Isingiro, whose settlement is Nakivale. To test whether this heterogeneity might drive our results, we run the main regressions by excluding Nakivale (Isingiro). Results are presented in Table 10 and do not change substantially from the results presented in the main-result section.

Table 10: OLS Regressions without Nakivale (Isingiro)

Panel A: Refugees	(1) Food Expenditure	(2) Food Variety	(3) Non-Food Expenditure	(4) Agriculture	(5) Agri. Revenues	(6) Enterprise	(7) Revenues
Dist. to Hosts	-1,652***	-0.418*	-2,501**	-0.0143**	4,643*	-0.0213***	-2,430
	(448.8)	(0.226)	(1,065)	(0.00658)	(2,481)	(0.00821)	(1,509)
Observations	3,784	3,820	1,977	1,985	1,979	1,985	1,979
R-squared	0.328	0.161	0.109	0.102	0.215	0.055	0.029
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No	No
Sample Mean	33,941	0.509	42,798	0.847	19,832	0.22	20,885
Panel B: Hosts	(1) Food Expenditure	(2) Food Variety	(3) Non-Food Expenditure	(4) Agriculture	(5) Agri. Revenues	(6) Enterprise	(7) Revenues
Dist. to Refugees	-2,772***	-0.468*	916.0	0.0163***	4,432	0.00515	-1,496
	(664.7)	(0.243)	(1,441)	(0.00432)	(4,583)	(0.0101)	(2,590)
Observations	2,831	2,855	1,499	1,525	1,494	1,525	1,491
R-squared	0.267	0.198	0.080	0.072	0.175	0.044	0.036
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No	No
Sample Mean	52,328	0.588	68,915	0.962	69,622	0.403	58,598

Notes: Authors' elaborations with RIMA (FAO) data. All the models exclude households living in Nakivale (Isingiro). The dependent variable is, in column (1), the value of household's food expenditure, in column (2), the variety of food consumption, in column (3) the value of household's non-food expenditure, in column (4), the probability of doing agriculture, in column (5) the revenues from agriculture, in column (6), the probability of running a business enterprise, and, in column (7), the revenues from business enterprise. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

As anticipated in the introduction, the concern about other mechanisms at work might arise. In particular, the increased volume of exchanges observed might happen through market infrastructures rather than be due to the proximity between the households. We consider here two possible channels that might be at work and compromise the interpretation of our results. The first is linked to the presence of organized marketplaces: the distance between the households of different groups might be correlated with the distance to the markets. If this were the case, the estimated coefficients might capture not only the effect of the interaction between the two groups, but also the presence of the settlement infrastructures.

In order to provide some evidence of the soundness of our empirical results, we run the same regressions as in the previous section, but we substitute the distance to the other group – *i.e.*, our main

explanatory variable – with the distance to the local market. The results are shown in Table 11, and we can safely conclude that the distance to the infrastructure does not drive our main results.

Table 11: OLS Regressions on the Distance to the Market

Panel A: Refugees	(1) Food Expenditure	(2) Food Variety	(3) Non-Food Expenditure	(4) Agriculture	(5) Agri. Revenues	(6) Enterprise	(7) Revenues
Distance to the Market	16.657	-0.005*	13.742	0.000	2.277	-0.000	9.665
	(13.979)	(0.003)	(13.983)	(0.000)	(8.553)	(0.000)	(23.687)
Observations	4,147	3,952	1,998	2,006	2,000	2,006	2,000
R-squared	0.309	0.154	0.107	0.101	0.214	0.053	0.027
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No	No
Sample Mean	33,941	0.509	42,798	0.847	19,832	0.22	20,885
Panel B: Hosts	(1) Food Expenditure	(2) Food Variety	(3) Non-Food Expenditure	(4) Agriculture	(5) Agri. Revenues	(6) Enterprise	(7) Revenues
Distance to the Market	-5.033	-0.006	52.351	0.000	-19.122	-0.000*	-18.781
	(8.271)	(0.008)	(56.503)	(0.000)	(511.349)	(0.000)	(36.416)
Observations	3,207	3,239	1,538	1,566	1,530	1,566	1,532
R-squared	0.237	0.216	0.080	0.055	0.175	0.047	0.035
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No	No
Sample Mean	52,328	0.588	68,915	0.962	69,622	0.403	58,598

Notes: Authors' elaborations with RIMA (FAO) data. In all models the main explanatory variable is the distance to the market. The dependent variable is, in column (1), the value of household's food expenditure, in column (2), the variety of food consumption, in column (3) the value of household's non-food expenditure, in column (4), the probability of doing agriculture, in column (5) the revenues from agriculture, in column (6), the probability of running a business enterprise, and, in column (7), the revenues from business enterprise. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

In a similar vein, we look at the possibility that exchanges take place not among households but through professional traders. If traders where the channel though which increased exchange takes places and the distance between households of different groups were correlated with the probability of using traders, again, the estimated coefficients presented in the previous section might be biased. As it can be easily grasped from Table 12 and 13, where all the coefficients are positive and non significant, we can discard this last mechanism and conclude that the between-group interaction has

a positive effect on households' welfare, especially of refugee households that appear to be the most economically disadvantaged.

Table 12: OLS Regressions for the Probability to Sell Crop Production to Traders

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	0.00487	0.00293	0.00716*
	(0.00350)	(0.00551)	(0.00403)
Observations	6,526	2,942	3,584
R-squared	0.189	0.114	0.147
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Sample Mean	0.28	0.468	0.114

Notes: Authors' elaboration on RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the probability of selling the crop production to professional traders, and the main explanatory variable is the inter-group distance. Data are available for the second and third rounds. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

Table 13: OLS Regressions for the Probability to Sell Enterprise Output to Traders

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group (1-NN)	0.0234	0.00902	0.0276
	(0.0178)	(0.0296)	(0.0239)
Observations	419	187	232
R-squared	0.121	0.083	0.151
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Sample Mean	0.244	0.338	0.16

Notes: Authors' elaboration on RIMA (FAO) data. Model (1) is estimated for the full population - *i.e.*, considering both host and refugee households. Model (2) refers to the hosting-community households only and Model (3) refers to the refugee households only. The dependent variable is the probability of selling the enterprise output to professional traders, and the main explanatory variable is the inter-group distance. Data are available for the second and third rounds. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Bootstrapped standard errors are computed with 50 replications. * p < .1, ** p < .05 *** p < .01

All the results presented so far are obtained by computing the standard errors with a Bootstrap procedure. This allows to avoid any *a priori* assumption on the correlation pattern of the errors. According to Bertrand et al. (2004) and Abadie et al. (2023), this non-parametric procedure works well – and in some cases better than any other parametric method – when the number of sampling units are a sizeable fraction of the number of clusters in the population. More precisely, Abadie et al. (2023) demonstrate that clustering standard errors by factors such as geography might be, on the one hand, arbitrary and, on the other hand, too conservative, when "*researchers observe units from all the clusters they are interested in*".[13] In our dataset all the refugee-hosting districts are covered.

Nonetheless, since in our case study the treatment is defined by the geographic location of the households, a cautious hypothesis to compute the standard errors would be to cluster them at the local (district) level. Therefore, we present also the results obtained when we adjust for possible spatially correlated shocks (Conley, 1999). The results with clustered standard errors at the district level are shown in Table 14 confirm those presented in the main text, even though the standard errors a bit more inflated as predicted by Abadie et al. (2023).

[13] See Abadie et al. (2023) p. 4

Table 14: OLS Regressions with Clustered Standard Errors

Panel A: Refugees	(1) Food Expenditure	(2) Food Variety	(3) Non-Food Expenditure	(4) Agriculture	(5) Agri. Revenues	(6) Enterprise	(7) Revenues
Distance to Hosts (1-NN)	-1,611**	-0.00422	-2,501*	-0.0143	4,643	-0.0213*	-2,430
	(723.6)	(0.00259)	(1,172)	(0.0179)	(4,548)	(0.00946)	(2,208)
Observations	3,854	3,890	1,977	1,985	1,979	1,985	1,979
R-squared	0.324	0.163	0.109	0.102	0.215	0.055	0.029
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No	No
Sample Mean	33,941	0.509	42,798	0.847	19,832	0.22	20,885
Panel B: Hosts	(1) Food Expenditure	(2) Food Variety	(3) Non-Food Expenditure	(4) Agriculture	(5) Agri. Revenues	(6) Enterprise	(7) Revenues
Distance to Ref.s (1-NN)	-2,772***	-0.00468*	916.0	0.0163**	4,432	0.00515	-1,496
	(956.8)	(0.00265)	(1,418)	(0.00623)	(5,997)	(0.0146)	(3,843)
Observations	2,831	2,855	1,499	1,525	1,494	1,525	1,491
R-squared	0.267	0.198	0.080	0.072	0.175	0.044	0.036
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No	No	No	No
Sample Mean	52,328	0.588	68,915	0.962	69,622	0.403	58,598

Notes: Authors' elaborations with RIMA (FAO) data. The dependent variable is, in column (1), the value of household's food expenditure, in column (2), the variety of food consumption, in column (3) the value of household's non-food expenditure, in column (4), the probability of doing agriculture, in column (5) the revenues from agriculture, in column (6), the probability of running a business enterprise, and, in column (7), the revenues from business enterprise. The household controls include: the population density (*i.e.*, the number of households of the same group living in a 1-kilometre radius), the number of household members, the number of female adults, the average years of education of adults, household head's age, and two dummies indicating, respectively, whether the household has a female head and whether the household head ever moved from the current residence. Errors are clustered at the district level. * $p < .1$, ** $p < .05$ *** $p < .01$

6. Conclusions

Refugees are present in large number and for protracted periods of time mainly in low- and middle-income countries. The sustainability of their presence both in economic and political terms is therefore of particular relevance. Such sustainability requires that, at least in the medium run, refugees are able to support themselves and that local communities benefit from their presence.

The different abilities and specialization of refugees with respect to the hosts provide us with a potential channel through which self-reliance of refugees and improved welfare for hosts can be

achieved. The arrival of refugees, as in a *love of variety* framework, can widen the set of products and services available in the market and, as shown in the theoretical model, increase consumption as well as production.

Transport costs play an important role in affecting the probability of exchange between households and in the rural setting we are considering, characterized by very limited transport infrastructures where most of the movements are on foot, physical distance plays a very important role.

We have shown that proximity between hosts and refugees, coherently with the theoretical model, substantially increases the level as well as the breadth of consumption and the involvement of households, especially refugees, in non-farm activities.

In particular, we have seen that proximity between refugees and hosts increases the food expenditure and the variety of food consumed by both groups. We also found that inter-group interactions raise both the non-food expenditure and the probability to run an enterprise (farm or non-farm) by refugee households, while host households are not crowded out from production.

Giving an initial partial answer to question about effective integration policies raised by Kadigo and Maystadt (2023), our results show that an effective and simple integration policy is just to let refugees and hosts mix and live close together. Favouring the integration, also physical, of the different communities helps therefore to exploit the potential for market creation due to the expansion of the set of goods and services available.

Our results do not exclude that other factors might be at play in the areas where the refugees are hosted, benefiting both hosts and refugee households. Infrastructures as well as social protection programs can improve the welfare of the residents of the area. However, we have identified a mechanism that appears to be independent, to a certain extent, of external interventions and that generates endogenously an increase in welfare through the market interaction of hosts and refugees.

Of course, other interventions, like the one mentioned above, might be instrumental and affect the outcomes stemming from the interaction, for example by supporting household demand through social protection and/or by facilitating exchanges by improving roads or other communication infrastructures. The extent to which this might be relevant remains an open question, as the data available do not allow us to test for such possible complementarity.

As a final remark, we notice that the economic improvements for hosts and refugees that we have highlighted in our study are not the only benefits stemming from the interactions among the households. Betts et al. (2023), for example, points to the broader positive impact of refugee-host interactions on social cohesions. This is an important point which deserve further analysis.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering?.*The Quarterly Journal of Economics*, *138*(1), 1-35.

Alix-Garcia, J., & Saah, D. (2010). The effect of refugee inflows on host communities: Evidence from Tanzania. *The World Bank Economic Review*, 24(1), 148-170, https://doi.org/10.1093/wber/lhp014

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

Alloush, M., Taylor, J. E., Gupta, A., Valdes, R. I. R., & Gonzalez-Estrada, E. (2017). Economic life in refugee camps. *World Development*, 95, 334-347, https://doi.org/10.1016/j.worlddev.2017.02.030

Aracı, D. T., Demirci, M., & Kırdar, M. G. (2022). Development level of hosting areas and the impact of refugees on natives' labor market outcomes in Turkey. *European Economic Review*, *145*, 104132, https://doi.org/10.1016/j.euroecorev.2022.104132

Benassy, J. P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters*, *52*(1), 41-47, https://doi.org/10.1016/0165-1765(96)00834-8

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?.*The Quarterly journal of economics*, *119*(1), 249-275.

Betts, A., Bloom, L., Kaplan, J. D., and Omata, N. (2014). Refugee economies: Rethinking popular assumptions (pp. 1-44). Oxford: University of Oxford, Refugee Studies Centre.

Betts, A., Stierna, M. F., Omata, N., & Sterck, O. (2023). Refugees welcome? Inter-group interaction and host community attitude formation. *World Development*, 161, 106088, https://doi.org/10.1016/j.worlddev.2022.106088

Bjørkhaug, I. (2020). Revisiting the refugee–host relationship in Nakivale Refugee Settlement: A dialogue with the Oxford Refugee Studies Centre. *Journal on migration and human security*, *8*(3), 266-281.

Coniglio, N. D., Peragine, V., & Vurchio, D. (2023). The effects of refugees' camps on hosting areas: Social conflicts and economic growth. *World Development*, *168*, 106273, https://doi.org/10.1016/j.worlddev.2023.106273

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of econometrics*, 92(1), 1-45, https://doi.org/10.1016/S0304-4076(98)00084-0

Cover, Thomas, and Peter Hart. (1967). Nearest neighbour pattern classification. *IEEE transactions on information theory*. 13.1: 21-27.

d'Errico, M., Mariani, R. D., Pietrelli, R., & Rosati, F. C. (2022). Refugee-host proximity and market creation in Uganda. *The Journal of Development Studies*, 58(2), 213-233, https://doi.org/10.1080/00220388.2021.1961749

Dixit, A. K., & Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *The American economic review*, 67(3), 297-308.

Fallah, B., Krafft, C., and Wahba, J. (2019). The impact of refugees on employment and wages in Jordan. *Journal of Development Economics*, 139, 203-216, https://doi.org/10.1016/j.jdeveco.2019.03.009

FAO and OPM (2018). Food security, resilience, and well-being analysis of refugees and host communities in Northern Uganda. FAO Resilience Analysis Report 12. FAO Publication available at http://www.fao.org/resilience/resources/resources-detail/en/c/1143820/

FAO and OPM (2019). Food security, resilience, and well-being analysis of refugees and host communities in Southwest Uganda. FAO Resilience Analysis Report 19. FAO Publication available at http://www.fao.org/resilience/resources/resources-detail/en/c/1234011/

Fujita, M., Krugman, P. R., & Venables, A. (2001). *The spatial economy: Cities, regions, and international trade*. MIT press

Fukunaga, Keinosuke, and Larry Hostetler. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on information theory*. 21.1 (1975): 32-40.

Gouel, C., & Sébastien, J. (2021), "Love of Variety and Gains from Trade", CESifo Working Paper, No. 9460, Center for Economic Studies and ifo Institute (CESifo), Munich.

Kadigo, M. M., & Maystadt, J. F. (2023). How to cope with a refugee population? Evidence from Uganda. *World Development*, *169*, 106293, https://doi.org/10.1016/j.worlddev.2023.106293

Kreibaum, M. (2016). Their suffering, our burden? How Congolese refugees affect the Ugandan population. *World Development*, 78, 262-287, https://doi.org/10.1016/j.worlddev.2015.10.019

Maystadt, J. F., Hirvonen, K., Mabiso, A., & Vandercasteelen, J. (2019). Impacts of hosting forced migrants in poor countries. *Annual Review of Resource Economics*, *11*, 439-459, https://doi.org/10.1146/annurev-resource-090518-095629

Maystadt, J. F., & Verwimp, P. (2014). Winners and losers among a refugee-hosting population. *Economic development and cultural change*, 62(4), 769-809, https://doi.org/10.1086/676458

Ruiz, I., & Vargas-Silva, C. (2013). The economics of forced migration. *The Journal of Development Studies*, *49*(6), 772-784, https://doi.org/10.1080/00220388.2013.777707

Taylor, J. E., Filipski, M. J., Alloush, M., Gupta, A., Valdes, R. I. R., & Gonzalez-Estrada, E. (2016). The economic impact of refugees. *Proceedings of the National Academy of Sciences*, 113(27), 7449-7453, https://doi.org/10.1016/j.worlddev.2017.02.030

Tsuda, S. (2022). Refugee inflows, surplus farm labor, and crop marketization in rural Africa. *Journal of Development Economics*, 155, 102805, https://doi.org/10.1016/j.jdeveco.2021.102805

Tumen, S. (2016). The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey. *American Economic Review*, *106*(5), 456-60, https://doi.org/aer.p20161065

UNHCR. (2024). Refugee Population Statistics Database. Available at: https://www.unhcr.org/refugee-statistics/, last accessed on September 16th, 2024.

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

Walelign, S. Z., Wang Sonne, S. E., & Seshan, G. (2022). Livelihood Impacts of Refugees on Host Communities, Policy Research Working Paper n. 10044

Zhou, Y. Y., Grossman, G., & Ge, S. (2023). Inclusive refugee-hosting can improve local development and prevent public backlash. *World Development*, *166*, 106203, https://doi.org/10.1016/j.worlddev.2023.106203

Appendix A. Theoretical Model

One of the main advantages of the interaction of a local with an immigrant community is the new availability of a wider range of consumption commodities, which are imperfect substitutes for the locally produced consumption items and which expand the choice sets for both communities. It is well known that a wider set of consumption commodities can be associated with welfare gains (the *love of variety* effect: see Gouel and Jean, 2021, for a recent assessment and discussion). The market structure and welfare implications of horizontally differentiated commodities have been studied by Dixit and Stiglitz (1977). Their analysis of monopolistic competition has been fruitfully applied to international trade and to the New Economic Geography (Fujita, Krugman and Venables, 1999).

In the present paper we explore the possibility that the creation of a refugee settlement close to a local village may enhance the welfare of both communities through an increase in their consumption opportunity sets. Under quasi-concave preferences there is a utility gain from allocating consumption among a wider variety of commodities (Benassy, 1996). The marginal production cost of firms producing horizontally differentiated and imperfectly substitute commodities can in turn be a function of the total number of firms in the market. In principle, the marginal cost could be a decreasing function of the number of firms when positive externalities across firms prevail: these can be associated to learning-by-doing and/or to positive network effects. On the other hand, the marginal production cost could be an increasing function of the number of firms when negative congestion externalities dominate. Our empirical analysis will help us assess which of these types of externalities prevails. Indeed, in section 4 we show that our findings are consistent with net positive externalities and with decreasing marginal production costs.

Following Dixit and Stiglitz (1977), we assume that the utility function of households in both the host and the refugee communities is separable into a (composite) commodity, q_0, and a vector of commodities which are imperfect substitutes of each other, $(q_1, q_2, \dots, q_n)$:

$$U = U(q_0, V_n(q_1, q_2, \dots, q_n)) \tag{A1}$$

The utility function $V_n(\cdot)$ is assumed to be quasi-concave, *i.e.* its indifference surfaces are convex and therefore there is a utility gain from diversifying consumption across the commodities $q_1, q_2, \dots, q_n$. The objective function (A1) must be maximised subject to a budget constraint:

$$I = \sum_{j=0}^{n} p_j q_j \tag{A2}$$

Both $U(\cdot)$ and $V_n(\cdot)$ are assumed to be CES in what follows. Furthermore, the utility function $V_n(\cdot)$ is linearly homogeneous and is normalised so that $V_1(q_1) \equiv q_1$. Following Benassy (1996), the utility gain from consuming *n* differentiated products instead of only one variety is given by:

$$v(n) = \frac{V_n(q,q,\ldots,q)}{V_1(nq)} = \frac{V_n(1,1,\ldots,1)}{n} \tag{A3}$$

with $v'(n) > 0$: the utility gain from differentiation is an increasing function of the number of varieties available to consumers. The elasticity of the function $v(n)$ measures the marginal taste for variety:

$$\mu(n) = \frac{nv\prime(n)}{v(n)} > 0 \tag{A4}$$

In a symmetric equilibrium $q_1 = q_2 = \cdots = q_n = q$. The utility function thus becomes:

$$U = U(q_0, V_n(q, q, \ldots, q)) \tag{A5}$$

with

$$\begin{aligned} V_n(q, q, \ldots, q) &= q \cdot V_n(1,1, \ldots,1) \\ &= q \cdot n \cdot v(n) \end{aligned} \tag{A6}$$

Thanks to the separability of the utility function (A5), the optimisation programme can be solved through a two-stage maximisation procedure (see Appendix B for details). The composite index of the differentiated commodity Q_1 is defined as

$$Q_1 = V_n(q_1, q_2, \ldots, q_n) = \left(\sum_{j=1}^{n} q_j^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \tag{A7}$$

and the corresponding dual price index P_1 as

$$P_1 = \sum_{j=1}^{n} p_j^{1-\theta} \tag{A8}$$

where $\theta > 1$ is the elasticity of substitution between the differentiated commodities $(q_1, q_2, \ldots, q_n)$. The relative expenditure shares for q_0 and Q_1 are:

$$\frac{s_1}{s_0} = \left(\frac{\alpha_1}{\alpha_0}\right)^{\sigma} \left(\frac{p_0}{P_1}\right)^{\sigma-1} \tag{A9}$$

with $\alpha_0 = 1/[1 + n \cdot v(n)]$, $\alpha_1 = [n \cdot v(n)]/[1 + n \cdot v(n)]$, and where $\sigma > 1$ is the elasticity of substitution between q_0 and Q_1. When the number of differentiated commodities *n* increases, the weight parameter $\alpha_1 = [n \cdot v(n)]/[1 + n \cdot v(n)]$ also increases and the price index P_1 declines, consistent with Bertrand competition among producers of the differentiated commodities facing marginal costs which are decreasing in the number of firms (see section 3.1 below). Hence, the share on consumption of the differentiated commodity index must increase as well.

The Marshallian demand functions are finally:

$$q_j = P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} \tag{A10}$$

A.1 No trading costs

Identical firms produce the differentiated commodities under monopolistically competitive conditions. Let their fixed production cost be *a* and their marginal cost $c(n)$. The marginal cost is a decreasing function of the number of firms *n* when positive externalities associated with either learning-by-doing *à la* Arrow or positive network effects prevail. By contrast, the marginal cost would be an increasing function of the number of firms when negative congestion externalities prevail. The absolute elasticity of demand for each producer is given by θ.

Profit maximisation requires that the marginal revenue of each firm be equal to its marginal cost:

$$p_i \cdot \left(1 - \frac{1}{\theta}\right) = c(n) \tag{A11}$$

By symmetry, all firms must charge the same price: $p_1 = p_2 = \cdots p_n = p_e$. Hence,

$$p_e = \left(\frac{\theta}{\theta - 1}\right) \cdot c(n) \tag{A12}$$

and the output produced by each firm is:

$$x_e = (\theta - 1) \cdot \frac{a}{c(n)} \tag{A13}$$

Assuming that positive network externalities prevail and $c'(n) < 0$, the equilibrium quantity produced by each firm is an increasing function of the number of commodities/firms, *n*. An increase in the size of the market measured by the number of firms *n* is therefore associated with an increased demand for the output of each firm.

A.2 Trading costs

We now consider the possibility that purchasing from the other community is costly to consumers. To this purpose and without loss of generality, the vector of commodities $(q_1, q_2, \dots, q_n)$ can be decomposed into two sub-vectors, $(q_1, q_2, \dots, q_{n_1})$ and $(q_{n_1+1}, q_{n_1+2}, \dots, q_n)$ respectively, where the first $n_1 < n$ components are produced by the local community and the remaining $n_2 = (n - n_1)$ are produced by the other community. Consumers incur positive iceberg-type trade costs parameterized by $\tau > 1$ when purchasing commodities from the other community. In our setting, the trading costs are related to the physical distance between the two communities which is usually covered on foot. The vector of prices faced by consumers is therefore:

$$(p'_0, p'_1, p'_2, \dots, p'_n) = (p_0, p_1, p_2, \dots, p_{n_1}, \tau p_{n_1+1}, \tau p_{n_1+2} \dots, \tau p_n) \quad \text{(A14)}$$

The dual price index (A8) becomes:

$$P'_1 = \tau^{(n_2/n)} \sum_{j=1}^{n} p_j^{1-\theta} = \tau^{(\frac{n_2}{n})} P_1 \quad \text{(A15)}$$

The relative expenditure shares are:

$$\frac{s_1}{s_0} = \left(\frac{\alpha_1}{\alpha_0}\right)^{\sigma} \left(\frac{p_0}{\tau^{(\frac{n_2}{n})} P_1}\right)^{\sigma-1} \quad \text{(A16)}$$

The consumption share of the differentiated commodity is now a decreasing function of the trade cost τ.

The Marshallian demand functions for the local and for the other community are now:

$$\begin{aligned} q_j &= P'_1 Q_1 \cdot \frac{(p')_j^{-\theta}}{\sum_{j=1}^{n} (p')_j^{1-\theta}} \\ &= P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} && \text{if } j = 1, 2, \dots, n_1 && \text{(A17a)} \\ &= \tau^{-\theta} P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} && \text{if } j = n_1+1, n_1+2, \dots, n && \text{(A17b)} \end{aligned}$$

We therefore obtain that the demand for consumption commodities of the other community is a decreasing function of the distance between the two communities, measured by the trading cost τ.

Appendix B. Derivation of the Marshallian demand function in the theoretical model

The optimisation programme of the household can be solved through a two-stage maximisation procedure. In the first stage the household maximises:

$$\max_{(q_0,Q_1)} U = U(q_0, n \cdot v(n) \cdot Q_1) = \left(\alpha_0 q_0^{\frac{\sigma-1}{\sigma}} + \alpha_1 Q_1^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{B1}$$

with $\alpha_0 = 1/[1 + n \cdot v(n)]$ and $\alpha_1 = [n \cdot v(n)]/[1 + n \cdot v(n)]$, where Q_1 is the composite index of the differentiated commodities and $\sigma > 1$ is the elasticity of substitution between q_0 and Q_1. The budget constraint is:

$$I = p_0 q_0 + P_1 Q_1 \tag{B2}$$

where P_1 is the price index consistent with two-stage maximisation (equation (A8)). The Marshallian demand functions for q_0 and Q_1 are:

$$q_0 = I \cdot \frac{\alpha_0^\sigma p_0^{-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} = \frac{I}{P} \cdot \alpha_0^\sigma p_0^{-\sigma} \tag{B3a}$$

$$Q_1 = I \cdot \frac{P_1^{-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} = \frac{I}{P} \cdot \alpha_1^\sigma P_1^{-\sigma} \tag{B3b}$$

where $P \equiv \alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}$ is the dual price index.

The expenditure shares for q_0 and Q_1 are respectively:

$$s_0 = \frac{\alpha_0^\sigma p_0^{-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} \tag{B4a}$$

$$s_1 = \frac{\alpha_1^\sigma P_1^{1-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} \tag{B4b}$$

The relative expenditure shares are therefore obtained as:

$$\frac{s_1}{s_0} = \left(\frac{\alpha_1}{\alpha_0}\right)^\sigma \left(\frac{p_0}{P_1}\right)^{\sigma-1} \tag{B5}$$

In the second state of the optimisation programme, households solve:

$$\max_{(q_1,q_2,\dots,q_n)} Q_1 = V_n(q_1, q_2, \dots, q_n) = \left(\sum_{j=1}^n q_j^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \tag{B6}$$

with $\theta > 1$, subject to:

$$P_1 Q_1 = \sum_{j=1}^{n} p_j q_j \tag{B7}$$

where P_1 is the dual price index:

$$P_1 = \sum_{j=1}^{n} p_j^{1-\theta} \tag{B8}$$

Note also that:

$$v(n) = n^{\frac{1}{\theta-1}}$$
$$\mu(n) = \frac{1}{\theta-1} n^{\frac{1}{1-1/\theta}}$$
$$\mu'(n) = \frac{1}{(\theta-1)^2} n^{\frac{1}{\theta-1}} > 0$$

The Marshallian demand functions are thus:

$$q_j = P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} \tag{B9}$$

Appendix C

The data used in the empirical analysis have been collected by the Food and Agriculture Organization of the United Nation (FAO), in coordination with the Office of the Prime Minister (OPM). The survey targets the refugee households and the hosting-community households that live in the area of the refugees' settlements. The first round of the data collection was in 2017 and, as of 2020, the survey was implemented in three rounds: in December 2017 and April 2018 the first round, in December 2019 the second round, and in December 2020 the third round.

In the following Table shows the summary statistics of the socio-demographic characteristics of the sampled households. The summary statistics relating the hosting-community households are shown in the upper panel, while the summary statistics relating the refugee households are shown in the lower panel.

Table C1: Socio-Demographic Characteristics of the Households (Full Sample)

Panel A	**Host Households**			
	Mean	St. Dev.	Min.	Max.
Number of Household Members	6.524	3.112	1	28
Number of Children	2.815	1.984	0	20
Number of Adult Members	3.709	2.032	1	15
Number of Male Adults	1.828	1.339	0	10
Number of Female Adults	1.881	1.193	0	10
Average Years of Education of Adults	5.141	2.795	0	19
Average Years of Education of Male Adults	5.607	3.409	0	20
Average Years of Education of Female Adults	4.646	3.035	0	21
Household Head's Age	45.789	14.817	0	100
Female Household Head	0.249	0.433	0	1
Years Since Arrival (Missing if Never Moved)	13.543	13.295	0	70
Observations	5476			
Panel B	**Refugee Households**			
	Mean	St. Dev.	Min.	Max.
Number of Household Members	6.230	3.206	1	52
Number of Children	3.147	2.169	0	26
Number of Adult Members	3.084	1.849	0	26
Number of Male Adults	1.391	1.215	0	11
Number of Female Adults	1.693	1.138	0	15
Average Years of Education of Adults	4.593	3.058	0	18
Average Years of Education of Male Adults	5.398	3.693	0	21
Average Years of Education of Female Adults	3.986	3.218	0	21
Household Head's Age	40.105	13.521	1	105
Female Household Head	0.505	0.500	0	1
Years Since Arrival (Missing if Never Moved)	3.741	3.959	0	44
Observations	6470			

Notes: Authors' elaboration on RIMA (FAO) data. The computation of the summary statistics relating the years since arrival of the household heads exclude those household heads that never moved from the current residence.

Refugee households have on average more children, one year less education, are more likely to be female headed and the latter is relatively younger.

In all the regressions we estimate, we control for the household composition – *i.e.*, the number of household members, the number of male and female adults – the average years of education of adults, the household head's age, whether the household is female headed and, for hosts, if it is a stayer household.

We further control for the average number of own-group households in a 1-kilometre radius, and for district and survey-wave fixed effects.