1. Introduction

A large number of refugees is hosted in low-income countries, where they remain for a long time. This has raised concerns about how the refugee communities could support themselves in the long run, especially as aid from national and, particularly, international organizations is far from sufficient and, as time passes, attention tends to be diverted to new emergency situations. It is therefore necessary to assess to what extent the integration of refugees with the local community might lead to the development of economic activities that will facilitate the long-run livelihood of the refugees. Moreover, this should happen without harming and possibly benefiting local communities in order to avoid tensions.

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

This paper investigates the possible gains arising from market exchanges between refugee and hosting-community households. By disentangling these from the benefits of the intergovernmental assistance, we detect the importance of avoiding marginalization in a forward-looking governance of refugee crises.

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 a small but statistically significant informal employment loss 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 especially to the impact on local communities. Among others, Alix-Garcia and Saah (2010), Alix-Garcia et al. (2018), Coniglio et al. (2023), Kadiko and Maystadt (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 (often positive) impact on local communities that live relatively close to the camps.[1]

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 benefit also 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-community with refugee households. In a previous paper, we have shown that hosting communities benefit from the presence of refugees because of such direct interaction with the refugees (d'Errico et al., 2022).

In this paper we focus more broadly on how the interaction between the host and refugees' communities can generate a market creation process able to improve the living standards of both and to ensure that refugees can, at least partly, self-sustain in the long run. We develop a theoretical model based on a *love of variety* approach to identify the economic mechanisms at the base of beneficial effects for both hosts and refugees of the arrival of the latter. 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, while improving also the welfare of the hosts.

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 that the presence of refugees enriches the offer of goods and services available to the host community and that, similarly, the refugees benefit by the opportunity of supplying such goods and services. Similar to a *love of variety* model of international trade, where access to a wider range of goods increases welfare and production in 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.

After developing a theoretical framework based on the class of *love of variety* models, we focus our empirical analysis on Uganda. This country hosts the largest number of refugees among the low-income countries and has 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.

[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).

Refugee settlements are not closed enclaves, but open to movement and to interaction with local communities. Moreover, we can make use of a unique dataset that collects a variety of information on refugees and hosts households in the vicinity of the refugee's settlements, covering more than 80 per cent of the refugee population and of the hosts residing in the areas close to the settlements.

In low-income countries, and especially in Uganda where refugees are hosted in remote areas, access to market is difficult for the lack both of formal markets and for 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 localized) that, 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 restrict our analysis to food expenditure, that however represent 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: that the distance between refugees and hosts is exogenous to households' location choice and that the distance between households is not correlated with the access to other services or benefits provided by national and international organizations. The latter assumption is necessary to ensure that we capture the effect of the direct interaction of households unconfounded by the delivery of services provided in the settlements. 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. We address the second concern by showing that the distance among households is not related to their probability of accessing the services available in the area.

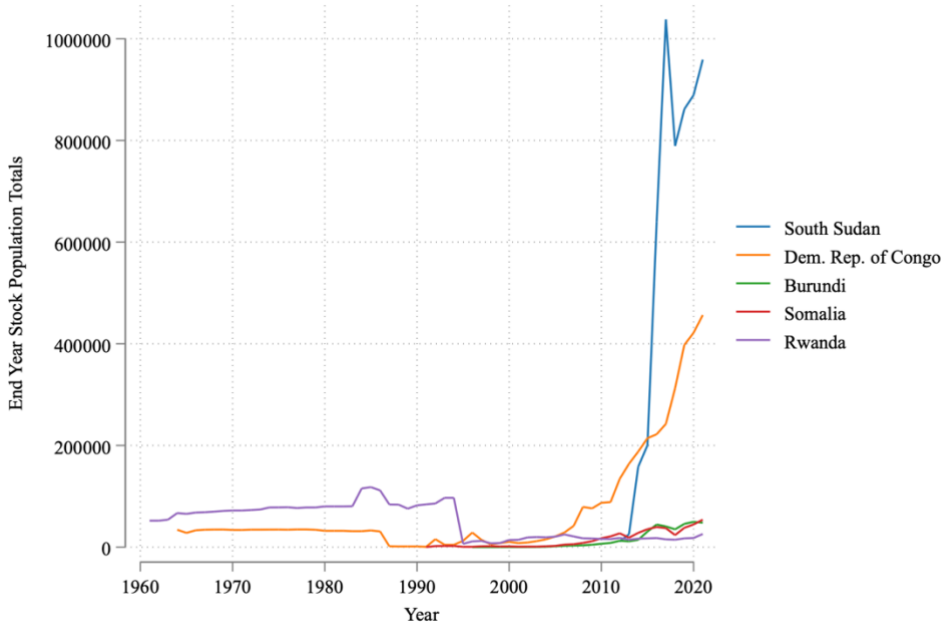
The paper is organised as follows. In section 2 we describe the historical and institutional background of Ugandan settlements. In section 3 we develop the theoretical model. Sections 4 and 5 detail the empirical analysis along with its results and section 6 concludes.

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.

In the region, grievous conflicts have burst after 2000. In particular, 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 most massive arrival 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)

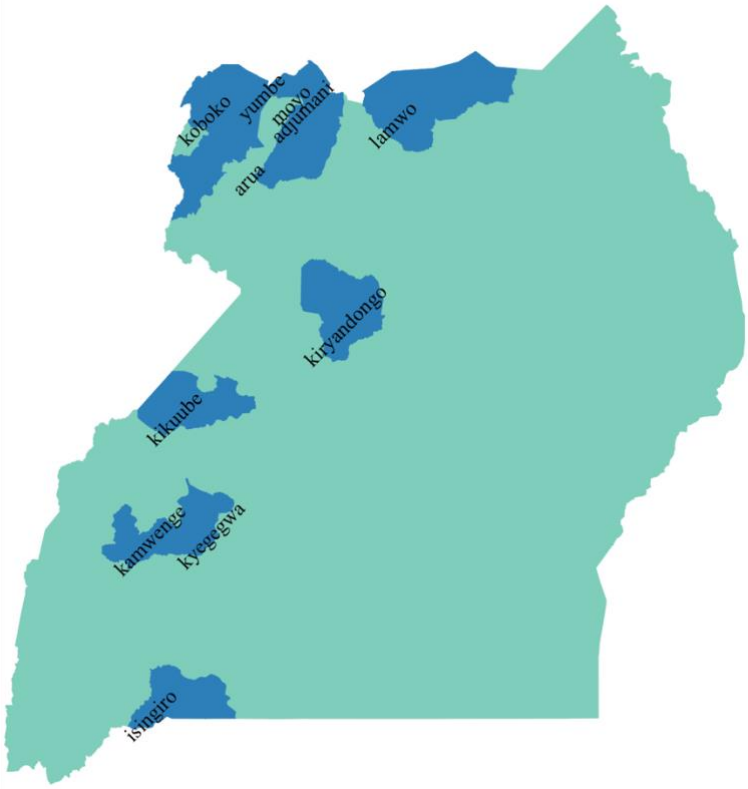


Notes: Authors' elaboration on UNHCR data (accessed online in July 2022, https://www.unhcr.org). End-year stocks are computed as the sum of refugees (or in a refugee-like situation) and asylum seekers by country of origin.

Refugees and asylum seekers living in Uganda are currently more than 1,4 million (UNHCR, 2021) and they represent the biggest refugee community in Sub-Saharan Africa. They come from the bordering countries, some of which have been already mentioned: South Sudan, DRC, 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). The settlements are located in the Northern and in the South-Western districts and most of them are proximate to the borders with the neighbouring countries. Currently, there are 31 settlements in 13 districts (see Figure 2).

Figure 2: Refugee-hosting Settlements



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.

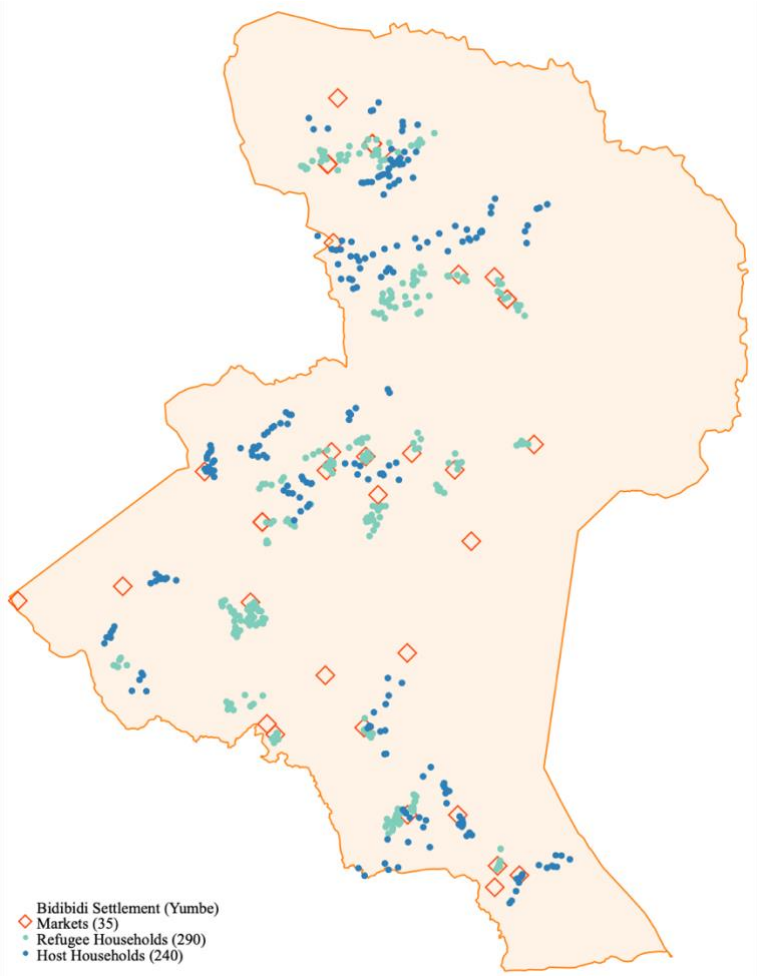
The organisation of the settlements is seen as a model of a progressive approach to refugee governance. The 2006 Refugees Act and the 2010 Refugees Regulations have established the following fundamental pillars. First, there are no restrictions in terms of number or of origin and all asylum seekers, 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.[2]

The latter point implies that the location of refugee households is decided by the institutions and the agencies involved in the management of the settlements, 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. As an example of the spatial extension and spread of the

[2] For a detailed description of Ugandan refugee dispersal policy see d'Errico et al. (2022).

settlements, we show in Figure 3 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 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. From 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. In the Figure 3, we show 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.

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 that 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.

On the basis of a large number of interviews carried out in the settlements of Nakivale and Kyangwali it emerges, for example, that "*Ugandans are important customers for settlement-based refugees. Each day, a significant number of Ugandans visit refugee settlements from neighbouring villages and cities […] to purchase products and services. […] Such visits play a central role in the economic life of many refugees. […] Ugandan vendors make regular trips to settlement markets in Nakivale and Kyangwali from nearby villages, where they seek refugee customers. Other refugees travel outside of Nakivale and Kyangwali to purchase basic necessities from Ugandans which they cannot get from shops within the settlements."* (Betts et al. 2014).

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 that they would not be able to produce under autarky.

3. 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 argue 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, $(q_1, q_2, \dots, q_n)$:

$$U = U(q_0, V_n(q_1, q_2, \dots, q_n)) \tag{1}$$

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 (1) must be maximised subject to a budget constraint:

$$I = \sum_{j=0}^{n} p_j q_j \tag{2}$$

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:

$$v(n) = \frac{V_n(q,q,\dots,q)}{V_1(nq)} = \frac{V_n(1,1,\dots,1)}{n} \tag{3}$$

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:

$$\mu(n) = \frac{nv'(n)}{v(n)} > 0 \tag{4}$$

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

$$U = U(q_0, V_n(q, q, \dots, q)) \tag{5}$$

with

$$\begin{aligned} V_n(q, q, \dots, q) &= q \cdot V_n(1,1, \dots,1) \\ &= q \cdot n \cdot v(n) \end{aligned} \tag{6}$$

Thanks to the separability of the utility function (5), the optimisation programme can be solved through a two-stage maximisation procedure (see Appendix A for details). The composite index of the differentiated commodity Q_1 is defined as

$$Q_1 = V_n(q_1, q_2, \dots, q_n) = \left(\sum_{j=1}^{n} q_j^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}} \tag{7}$$

and the corresponding dual price index P_1 as

$$P_1 = \sum_{j=1}^{n} p_j^{1-\theta} \tag{8}$$

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

$$\frac{s_1}{s_0} = \left(\frac{\alpha_1}{\alpha_0} \right)^{\sigma} \left(\frac{p_0}{P_1} \right)^{\sigma-1} \tag{9}$$

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:

$$q_j = P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} \tag{10}$$

3.1 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:

$$p_i \cdot \left(1 - \frac{1}{\theta}\right) = c(n) \tag{11}$$

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

$$p_e = \left(\frac{\theta}{\theta - 1}\right) \cdot c(n) \tag{12}$$

and the output produced by each firm is:

$$x_e = (\theta - 1) \cdot \frac{a}{c(n)} \tag{13}$$

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.

3.2 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 must be usually covered on foot. The vector of prices faced by consumers is therefore:

$$(p_0', p_1', p_2', \dots, p_n') = (p_0, p_1, p_2, \dots, p_{n_1}, \tau p_{n_1+1}, \tau p_{n_1+2} \dots, \tau p_n) \tag{14}$$

The dual price index (8) becomes:

$$P_1' = \tau^{(n_2/n)} \sum_{j=1}^{n} p_j^{1-\theta} = \tau^{(\frac{n_2}{n})} P_1 \tag{15}$$

The relative expenditure shares are:

$$\frac{s_1}{s_0} = \left(\frac{\alpha_1}{\alpha_0}\right)^{\sigma} \left(\frac{p_0}{\tau^{(\frac{n_2}{n})} P_1}\right)^{\sigma-1} \tag{16}$$

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:

$$\begin{aligned} q_j &= P_1' Q_1 \cdot \frac{(p')_j^{-\theta}}{\sum_{j=1}^{n} (p')_j^{1-\theta}} \\ &= P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} && \text{if } j = 1, 2, \dots, n_1 && (17a) \\ &= \tau^{-\theta} P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} && \text{if } j = n_1+1, n_1+2, \dots, n && (17b) \end{aligned}$$

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.

4. The Empirical Analysis

4.1 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 that live in the proximity 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. The survey covered all the districts,[3] located in the North-West and in the South-West of the country, that host a refugee settlement.[4]

In 2020, the survey re-interviewed all the refugee and host households interviewed in the first wave of the survey (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).[5] 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. 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]

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 nearest settlement.

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

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

	Hosts				**Refugees**			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
Distance to Refugees	1.567	1.307	0	6.998	0.163	0.296	0	5.913
Refugees in 5-Km Radius	78.129	79.395	0	318	118.625	91.294	0	317
Distance to Hosts	0.203	0.263	0	5.281	1.628	1.260	0	6.867
Hosts in 5-Km Radius	58.685	31.957	0	161	60.660	51.978	0	200
Distance to Petty Markets	5.475	60.754	0	999	11.038	97.066	0	999
Distance to Transports	9.660	88.753	0	999	9.283	86.731	0	999
Distance to Church/Mosque	3.822	49.368	0	999	4.008	53.599	0	999
Observations	1,632				2,107			

Notes: Authors' elaboration on RIMA (FAO) dataset. Only observations which GPS coordinates are correctly observed are included. 999 indicates no access.

[3] Arua, Yumbe, Moyo, Adjumani, Lamwo, Kiryandongo, Kyegegwa, Kamwenge, Isingiro, Kikuube and Koboko.

[4] All the districts have only one settlement, but Adjumani, which hosts 18 settlements, Isingiro, which hosts 2 settlements, and Arua, which hosts 2 settlements.

[5] In the preceding rounds, the sampling strategy was a two-stage cluster sampling method, with settlement block as the Primary Sampling Unit (PSU) and the households as the Secondary Sampling Unit (SSU). For more details see FAO and OPM (2018) and FAO and OPM (2019).

[6] For each estimate we specify the years for which information are missing.

4.2 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), refugees not only bring with them different cultures, but also offer goods and services that are different with respect to those offered by the hosts and that are appreciated by them.

The theoretical model developed in the previous section predicts that a larger size of the market and a reduction in transportation costs enhance the level and composition of consumption and increase the production of goods and services. In particular, if marginal costs are decreasing along with the size of the market, we expect to observe an increase in the supply of goods and services and also in the level and variety of consumption. Furthermore, as shown by the theoretical model in the previous section, 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:

$$Y_{i,j,k}^{t} = \alpha + \beta \cdot Dist_{i,j,-k} + \boldsymbol{\delta} \cdot \boldsymbol{X_i^t} + \varphi_j + \rho_t + \varepsilon_{i,j}^{t} \quad (18)$$

$Y_{i,j,k}^{t}$is the outcome of interest for household i, living in district j, belonging to the community $k = \{h; r\}$: host or refugee, observed during round t. $Dist_{i,j,-k}$ is the distance between host and refugee households, computed as the earth-arc distance between two points.[7] $\boldsymbol{X_i}$ is a vector of household controls, and φ_j and ρ_t are district and round fixed effects respectively.[8] $\varepsilon_{i,j}^{t}$ is the idiosyncratic error term. The estimation method is by Pooled Ordinary Least Squares.

With a similar approach, we then look at other outcomes. In particular, the variety of consumption, as detailed in the next section, the probability of running a farm and a non-farm enterprise and the revenues.

The estimation strategy we implement 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. Nonetheless, this strategy might raise more than one concern.

[7] The inter-group distance is the distance between groups. An example might help to clarify. Let us consider a refugee household living in Yumbe district. Based on GPS coordinated we can calculate the distance to all the host households living in Yumbe. To compute the inter-group distance, we consider the distance to the closest host household – *i.e.*, the host household for which the distance is the shortest.

[8] See Appendix B for a detailed description of the data and a full list of the control variables.

First, the location of the households, and consequently the distance among them, could in principle 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 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.[9] 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 the settlement capacity or to family reunification.[10]

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 and on the probability to receive assistance 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. Treated households are considered those whose distance to the other group is lower than the median distance. As the Table shows, hosting-community households that live far away from refugees – *i.e.*, the "controls" – are not systematically different from hosts that live close by. Instead, refugee households that live far away from the hosting community are significantly larger, but this is the only attribute in which control and treated refugees are different. 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 the settlement and the services provided within them. 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 reflect 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 in the proximity of 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 households.

[9] See Appendix C for a more details.

[10] For security reasons, we are not allowed to disclaim refugees' ethnic origin. Nevertheless, 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 without giving information on their location. See Appendix C for details.

Table 2 also shows that the relative location of hosts is not correlated with the probability to receive assistance and, in particular, cash transfers.

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

	Controls		Treated		Difference	
Panel A:	**Hosting-Community Households**					
Number of Household Members	6.62	[3.20]	6.63	[3.18]	0.00	(0.12)
Share of Children in the Household	0.41	[0.22]	0.42	[0.21]	0.01	(0.01)
Share of Male Adults	0.48	[0.21]	0.48	[0.20]	0.00	(0.01)
Average Years of Education of Adults	4.33	[2.71]	4.16	[2.69]	-0.17	(0.10)
Household Head's Age	44.80	[14.61]	45.25	[14.77]	0.46	(0.55)
Female Household Head	0.25	[0.43]	0.27	[0.44]	0.02	(0.02)
Years Since Arrival (Missing if Never Moved)	13.83	[13.49]	12.95	[13.16]	-0.88	(1.31)
Probability to Receive Assistance (Cash Transfers)	0.51	[0.50]	0.50	[0.50]	-0.01	(0.02)
Observations	1453		1453		2906	
Panel B:	**Refugee Households**					
Number of Household Members	6.48	[3.23]	6.05	[3.08]	-0.43***	(0.10)
Share of Children in the Household	0.47	[0.22]	0.46	[0.23]	-0.01	(0.01)
Share of Male Adults	0.41	[0.26]	0.43	[0.27]	0.01	(0.01)
Average Years of Education of Adults	4.02	[3.02]	4.03	[3.03]	0.01	(0.10)
Household Head's Age	39.74	[13.51]	39.20	[13.22]	-0.55	(0.43)
Female Household Head	0.54	[0.50]	0.51	[0.50]	-0.03	(0.02)
Years Since Arrival (Missing if Never Moved)	3.13	[4.07]	3.09	[3.35]	-0.04	(0.12)
Probability to Receive Assistance (Cash Transfers)	0.51	[0.50]	0.53	[0.50]	0.02	(0.02)
Observations	1990		1990		3980	

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

More importantly, as an illustration of the location of the households living next to and in a settlement, along with the official markets of a settlement, Figure 3 in section 2 shows a detailed map of Bidibidi. Living close to or far from the refugee households does not imply that hosts also live close to or far from the markets. In what follows, we qualify this circumstantial evidence by showing that there is not a significant correlation between the outcomes considered and the proximity to the market.

d'Errico et al. (2022) provides additional evidence on the orthogonality between refugees' location and hosts' access to several services and infrastructures. In particular, they exploit the baseline information of hosts' access to several services and infrastructures and their distance to the settlement

reception centres. From the results of their analysis, it appears that there is no correlation between the two.[11]

In section 5.3 we also show the results of a robustness analysis aiming at assessing the potential confounding effect of the presence of settlement services or of national and international agencies at work in the area. 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.

5. Empirical Results

5.1 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, 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	-3,283*** (402.6)	-2,613*** (601.9)	-1,528*** (530.7)
Observations	6,685	2,841	3,854
R-squared	0.290	0.265	0.327
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. The dependent variable is available for the first and second round. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), the number of household members, the number of female adults, the number of male adults, adults' average years of education, household head's age, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis. * p < .1, ** p < .05 *** p < .01

[11] See d'Errico et al. (2022) p. 8.

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 when we consider all households together and when we distinguish between hosts and refugees. In general, a reduction in 1 kilometre of the inter-group distance is associated to an increase in the food expenditure by 3,283 Ugandan Shillings (corresponding to about 8 percent of the sample mean).

To assess whether the proximity to households of the other group increases the variety of consumption, we can rely on a very rich module of food consumption in the questionnaire administered by FAO during the first wave of the data collection. 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 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.

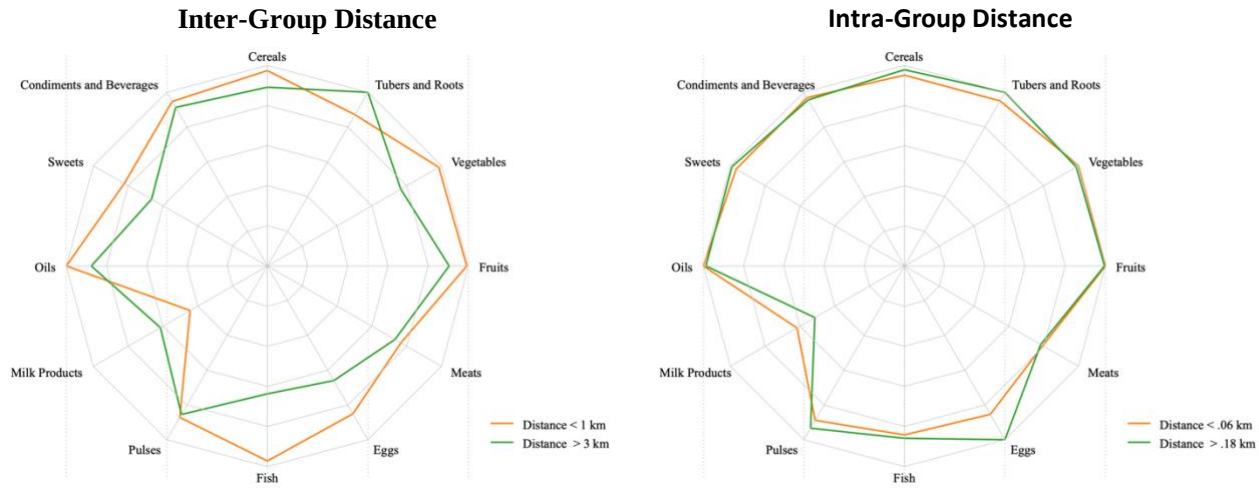
As illustrative evidence of the predictions of the theoretical model we count the number of varieties consumed by the households within each category. We standardise the count so that it ranges from zero (no variety consumed in the category) to one (all the varieties within a category are consumed) and compute the average number of varieties consumed by the households that live close 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.

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

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 that live relatively close to the other community households presents 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).

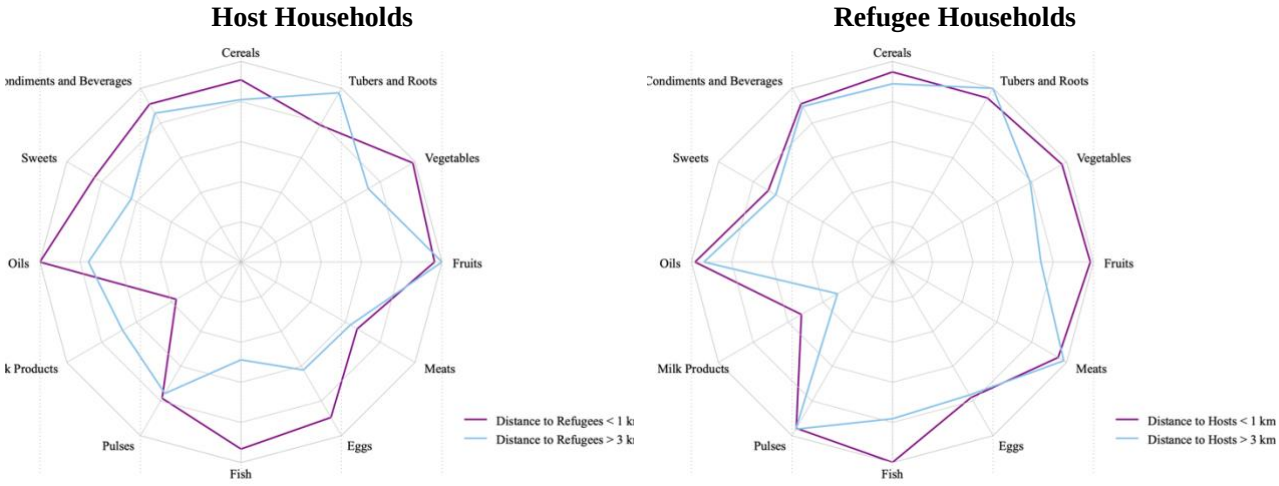
We take this descriptive evidence as a pattern which deserves an inferential empirical analysis.

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.

More formally, then, we regress the number of varieties consumed on the inter-group distance and the other controls mentioned above.

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 0.6 p.p. 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	-0.00590*** (0.00162)	-0.00451* (0.00244)	-0.00408* (0.00224)
Observations	6,745	2,865	3,890
R-squared	0.211	0.197	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 main explanatory variable is the inter-group distance. To be coherent with the estimation on food expenditure, the analysis is restricted to the first and second wave only. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), the number of household members, the number of female adults, the average years of education of adults, household head's age, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis. * 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. Nonetheless, we cannot replicate the analysis on the variety of consumption for non-food expenditure, as data do not allow us to decompose the consumption in sub-categories – *i.e.*, we do not have information on varieties of footwear like socks, shoes, boots, sandals. Therefore, in Table 5 we show the results for the level of consumption only.

Table 5: OLS Regressions for Non-Food Expenditure

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group	-1,006 (883.2)	765.4 (1,497)	-2,453** (1,050)
Observations	3,476	1,499	1,977
R-squared	0.101	0.079	0.109
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	No	No	No
Sample Mean	54110.246	68915.657	42798.009

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. The dependent variable is available for the first round. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), the number of household members, the number of female adults, the number of male adults, adults' average years of education, household head's age, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis. * p < .1, ** p < .05 *** p < .01

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 going to observe also in the analysis for the production side, shown in the following section.

5.2 Production

We then move the analysis to the production side and 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.

Table 6: OLS Regressions for the Probability of Doing Agriculture

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group	-0.00374 (0.00437)	0.0168*** (0.00487)	-0.0142** (0.00672)
Observations	3,510	1,525	1,985
R-squared	0.081	0.071	0.114
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. To be coherent with the estimation on non-farm enterprise, the analysis is restricted to the first wave only. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), the number of household members, the number of female adults, the average years of education of adults, household head's age, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis. * p < .1, ** p < .05 *** p < .01

According to the coefficients shown in the Table 6, living 1-kilometre closer to the host households is associated for refugees to a 1.4 p.p. higher probability of 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 distance on revenues from agriculture and results are reported in Table 7. Even if the coefficients are positive, they are 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.

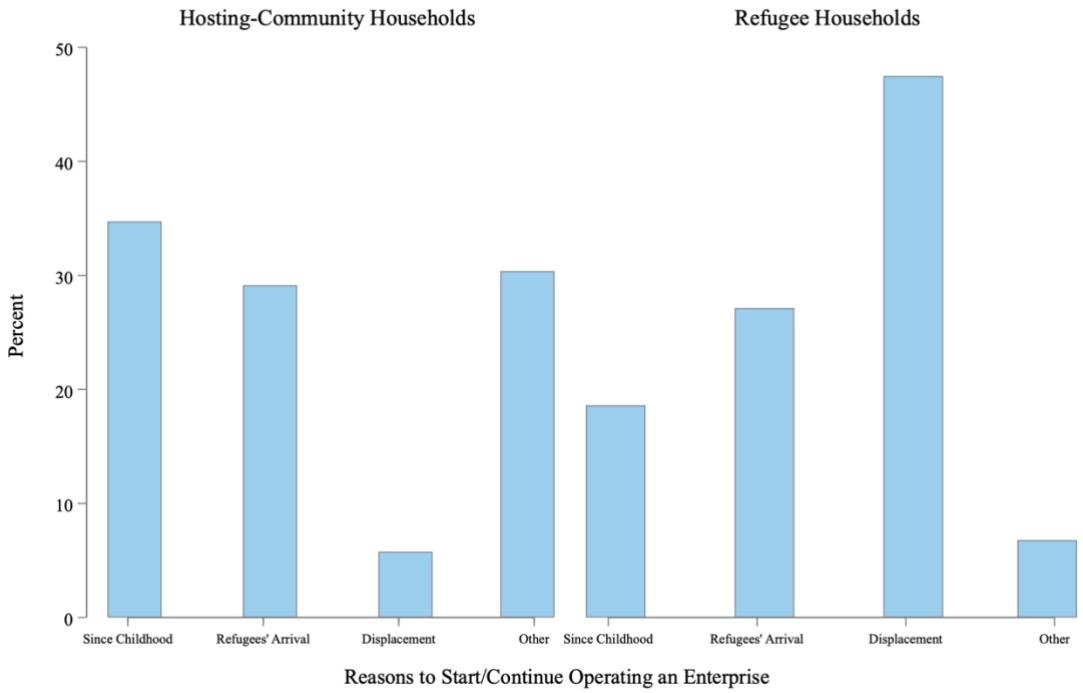
Table 7: OLS Regressions for the Revenues from Agriculture (in Logarithms)

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group	0.0204 (0.0732)	0.0244 (0.125)	-0.000697 (0.0862)
Observations	3,343	1,397	1,946
R-squared	0.168	0.092	0.248
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
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. To be coherent with the estimation on non-farm enterprise, the analysis is restricted to the first wave only. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), the number of household members, the number of female adults, the average years of education of adults, household head's age, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis. * p < .1, ** p < .05 *** p < .01

We now look at the changes in non-farm activities both along the intensive and the extensive margin. As shown in Figure 6, 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 appear to have started a non-farm activity in response to the refugee arrival to a much 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.

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



Notes: Authors' elaboration on RIMA data

The results of the regressions presented in Table 8 are consistent with the descriptive evidence. 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	-0.0116* (0.00660)	0.00609 (0.0109)	-0.0208*** (0.00788)
Observations	3,510	1,525	1,985
R-squared	0.068	0.044	0.056
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. Data are available for the first round. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), 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, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis. * 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 household appear 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	-0.132* (0.0710)	0.0496 (0.123)	-0.200** (0.0803)
Observations	3,317	1,382	1,935
R-squared	0.059	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. Data are available for the first round. The household controls include: the number of households of the same group living in a 1-kilometre radius (population density), the number of household members, the number of female adults, the average years of education of adults, household head's age, a dummy indicating whether the household has a female head and the months since arrival to the current residence (999 if the household head never moved). Robust standard errors in parenthesis.
* 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.

5.3 Robustness Analysis

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 market places: the distance between the households of different groups might be correlated with the distance to the market. 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. As shown in Table 9, none of the outcomes is correlated with the distance to the market, but the food variety index. Nonetheless, the estimated coefficients are very small and only marginally significant, therefore we can safely conclude that the distance to the infrastructure does not drive our results.

Table 10: OLS Regressions on the Distance to the Market

	(1) Food Exp.	(2) Food Variety	(3) Non-Food Exp.	(4) Farm	(5) Farm Revenues	(6) Enterprise	(7) Revenues
Distance to a Petty Market	9.486	-0.000**	21.91	0.000	-1.738	-0.000*	-0.676
	(9.680)	(2.73e-05)	(17.34)	(5.46e-05)	(11.20)	(7.82e-05)	(20.37)
Observations	7,354	7,427	3,536	3,572	3,530	3,572	3,532
R-squared	0.268	0.210	0.101	0.080	0.180	0.066	0.041
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	41,953	0.544	54,110	0.897	41,288	0.299	37,177

Notes: Authors' elaboration on RIMA (FAO) data.

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 11 and 12, 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 11: OLS Regressions for the Probability to Sell Crop Production to Traders

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group	0.00534	0.00313	0.00710
	(0.00423)	(0.00680)	(0.00464)
Observations	6,526	2,952	3,584
R-squared	0.189	0.115	0.146
HH Controls	Yes	Yes	Yes

District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Sample Mean	0.279	0.468	0.114

Notes: Authors' elaboration on RIMA (FAO) data.

Table 12: OLS Regressions for the Probability to Sell Enterprise Output to Traders

	(1) All	(2) Hosts	(3) Refugees
Distance to the Other Group	0.0236	0.00914	0.0274
	(0.0183)	(0.0294)	(0.0240)
Observations	419	188	232
R-squared	0.118	0.084	0.147
HH Controls	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Sample Mean	0.244	0.338	0.160

Notes: Authors' elaboration on RIMA (FAO) data.

Since, in our case study, the treatment is defined by the geographic location of the households, a more cautious hypothesis to compute the standard errors would be to cluster them at the local (district) level. By doing so, we adjust for possible spatially correlated shocks (Conley, 1999). Therefore, we estimate the regressions for food expenditure and food variety with clustered standard errors. and the results shown in Table 13 confirm that living close by the other group is associated with a significantly higher outcome.

Table 13: OLS Regressions with Standard Errors Clustered at the District Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Food Expenditure			**Food Variety**		
	All	Hosts	Refugees	All	Hosts	Refugees
Distance to the Other Group	-3,283***	-2,772**	-1,528**	-0.00590***	-0.00459	-0.00408
	(773.4)	(1,069)	(599.4)	(0.00174)	(0.00306)	(0.00275)
Observations	6,685	2,831	3,854	6,745	2,855	3,890
R-squared	0.290	0.267	0.327	0.211	0.198	0.163
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Mean	41,953	52,328	33,941	0.544	0.588	0.509

Notes: Authors' elaboration on RIMA (FAO) data.

Finally, even if our theoretical model points at the distance as the main obstacle to interaction and, consequently, to the trade between groups, we also test a different measure of the exposition to the other group – *i.e.*, the number of households of the other group living a radius of 1 kilometre. This measure of the potential of exchange is likely to have a positive effect on households' outcomes and Table 14 shows that is the case for food expenditure and food variety.

Table 14: OLS Regressions on the Density in 1-km Radius

	(1)	(2)	(3)	(4)	(5)	(6)
	Food Expenditure			**Food Variety**		
	All	Hosts	Refugees	All	Hosts	Refugees
Num. of HHs (other group)	448.7***	144.3***	191.9***	0.000957***	0.000651***	0.000482**
	(121.8)	(49.93)	(69.56)	(0.000323)	(0.000172)	(0.000198)
Observations	7,354	3,207	4,147	7,427	3,239	4,188
R-squared	0.269	0.240	0.313	0.210	0.219	0.163
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample Mean	41,953	52,328	33,941	0.544	0.588	0.509

Notes: Authors' elaboration on RIMA (FAO) data. The Density is measured as the number of the households belonging to the other group and living in a 1-kilometre radius from the observation.

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. In a *love of variety* framework, the arrival of refugees can widen the set of products and

services available in the market and, as shown in the theoretical framework, 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 consumption of 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.

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 or by facilitating exchanges by improving roads or other communication infrastructures. To what extent 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 at some positive impact of refugee-host interactions on social cohesions that deserve further analysis.

Acknowledgments: We would like to thank FAO Uganda, the Uganda Bureau of Statistics (UBOS) and the United Nations Refugee Agency (UNHCR) for their support on data collection. We also thank Stefano Gagliarducci, Anna Hillingdon, Rebecca Pietrelli, Francesco Sobbrio and participants to the Internal Seminars in University of Rome "Tor Vergata", to the SITES 2022 Annual Conference in Naples, to the workshop "Improving Impact in the Resilience and Social Protection Agenda in Uganda", and to the IFAD 2022 conference in Rome for valuable comments and inputs. Any errors or omissions are the sole responsibility of the authors.

Bibliography

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

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

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

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

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. (2021). Refugee Population Statistics Database. Available at: https://www.unhcr.org/refugee-statistics/

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

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:

$$\max_{(q_0,Q_1)} U = U(q_0, n \cdot v(n) \cdot Q_1) = \left(\alpha_0 q_0^{\frac{\sigma-1}{\sigma}} + \alpha_1 Q_1^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \quad \text{(A1)}$$

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:

$$I = p_0 q_0 + P_1 Q_1 \quad \text{(A2)}$$

where P_1 is the price index consistent with two-stage maximisation (equation (8)). The Marshallian demand functions for q_0 and Q_1 are:

$$q_0 = I \cdot \frac{\alpha_0^\sigma p_0^{-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} = \frac{I}{P} \cdot \alpha_0^\sigma p_0^{-\sigma} \quad \text{(A3a)}$$

$$Q_1 = I \cdot \frac{P_1^{-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} = \frac{I}{P} \cdot \alpha_1^\sigma P_1^{-\sigma} \quad \text{(A3b)}$$

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:

$$s_0 = \frac{\alpha_0^\sigma p_0^{-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} \quad \text{(A4a)}$$

$$s_1 = \frac{\alpha_1^\sigma P_1^{1-\sigma}}{\alpha_0^\sigma p_0^{1-\sigma} + \alpha_1^\sigma P_1^{1-\sigma}} \quad \text{(A4a)}$$

The relative expenditure shares are therefore obtained as:

$$\frac{s_1}{s_0} = \left(\frac{\alpha_1}{\alpha_0}\right)^\sigma \left(\frac{p_0}{P_1}\right)^{\sigma-1} \quad \text{(A5)}$$

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

$$\max_{(q_1,q_2,\dots,q_n)} Q_1 = V_n(q_1, q_2, \dots, q_n) = \left(\sum_{j=1}^n q_j^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}} \quad \text{(A6)}$$

with $\theta > 1$, subject to:

$$P_1 Q_1 = \sum_{j=1}^{n} p_j q_j \tag{A7}$$

where P_1 is the dual price index:

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

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:

$$q_j = P_1 Q_1 \cdot \frac{p_j^{-\theta}}{\sum_{j=1}^{n} p_j^{1-\theta}} \tag{A9}$$

Appendix B

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 proximity 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 B 1: 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 the years since arrival.

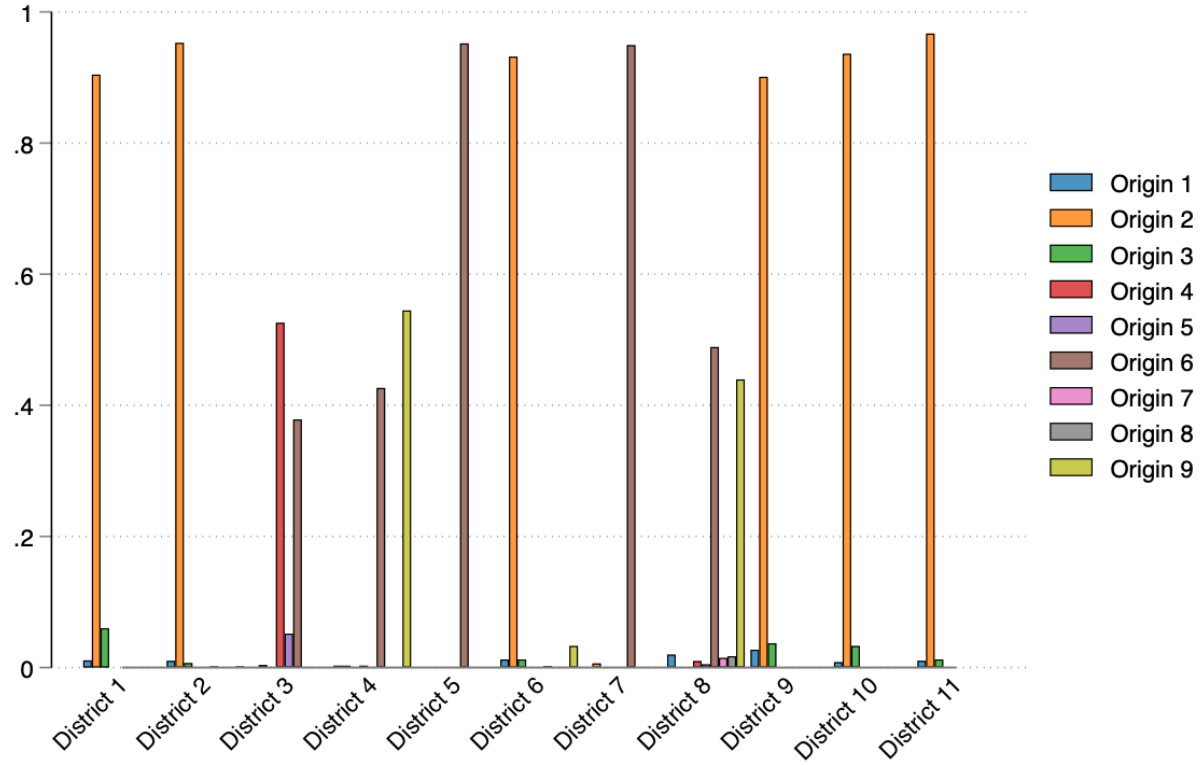
We further control for the average number of households in a 1-kilometre radius and for district and survey-wave fixed effects.

Appendix C

As discussed in the main text, 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, 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.

For security reasons, we are not allowed to disclaim refugees' ethnic origin. Nevertheless, 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 without giving information on their location (see Figure C1).

Figure C 1: Origin Composition of Districts' Refugee Population

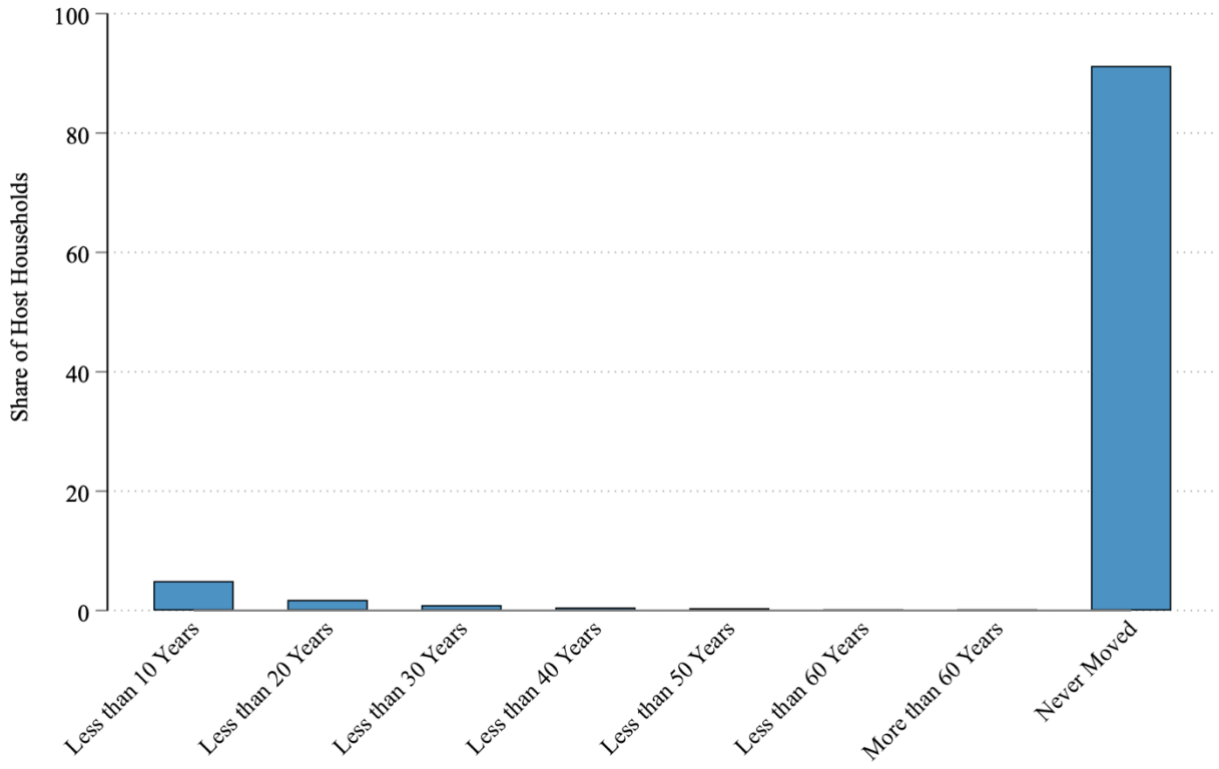


Notes: Authors' elaboration on RIMA (FAO) data

Furthermore, 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 C 2, 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 C 2: Years since Arrival of Hosting-Community Household Heads



Notes: Authors' elaboration on RIMA (FAO) data