Journal of Development Economics 155 (2022) 102802

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

When do refugees return home? Evidence from Syrian displacement in Mashreq☆

Lori Beaman [a], Harun Onder [b], Stefanie Onder [c,*]

[a] Department of Economics & Institute for Policy Research, Northwestern University, Evanston, IL, United States of America
[b] Macroeconomics, Trade, & Investment Global Practice, World Bank Group, Washington, DC, United States of America
[c] School of International Service, American University, Washington, DC, United States of America

ARTICLE INFO

JEL classification:
F22
D74
J61

Keywords:
Return migration
Forced displacement
Refugees

ABSTRACT

This paper provides an empirical analysis of refugee returns to Syria. Since 2011, about 5.6 million Syrians – more than a quarter of the country's pre-conflict population – have been registered as refugees. By mid-2018, only about 1.8 percent of them had returned to Syria voluntarily. This paper compiles a novel data set with administrative data for 2 million refugees, existing and new household surveys, a new conflict-events database, and nightlights data for Syria to analyze the correlates of these returns. A reduction in conflict intensity and an increase in luminosity in Syria increase the likelihood of spontaneous return. Moreover, the patterns of who returns and when differ between high and low conflict areas of Syria. Finally, we show there is a positive association between better conditions faced by refugees in exile and the likelihood of return to Syria.

1. Introduction

The conditions under which refugees *can* return to their countries of origin, or stay in exile, have been studied extensively under international law.[1] However, the conditions under which they would *choose* to do so have received less attention. On the one hand, an emerging body of literature has focused on the refugees' self-reported return intentions but not on actual returns, which may be different.[2] On the other, the literature on temporary (voluntary) migration has revealed important properties of actual returns — but only that of economic migrants.[3] However, refugees are distinct in many ways, including their human capital investments and dynamic earnings profiles in exile, which is often attributed to differences in their subjective return probability (Cortes, 2004; Galor and Stark, 1991). In this paper, we present an empirical analysis of the early (voluntary and unassisted) return decisions of 2 million Syrian refugees who were displaced between January 2011 and March 2018.

Starting from 2011, more than 5.6 million Syrians fled one of the most destructive wars in recent decades, which led to more than 400,000 direct conflict-related deaths, physical damage to about half of the schools and hospitals in major urban centers, and deepened sectarian divisions (World Bank, 2017). The outflow of refugees peaked in 2013, but never ceased completely. In the meantime, according to the United Nations High Commissioner for Refugees (UNHCR), 103,090 – about 1.8 percent – were verified to return to Syria by mid-2018, the

☆ We would like to thank the Editor-in-Chief Andrew Foster and an anonymous referee for their thoughtful comments. An earlier version of this paper was prepared as a background note for the World Bank report "The Mobility of Displaced Syrians: An Economic and Social Analysis". The findings, interpretations, and conclusions expressed in this study do not necessarily reflect the views of the World Bank, its Board of Directors, the governments they represent, or the UNHCR. Comments from the seminar participants at the American University and at the Joint Research Conference on Forced Displacement by the World Bank, UNHCR, Harvard University, and the Brown University, and from Saroj Kumar Jha, Eric Le Borgne, Kevin Carey, Erhan Artuc, and Caglar Ozden are gratefully acknowledged. Natasha Ahuja, Ola Hisou, Gina Kosmidou, Liam Paul Ellberg Frölund, Michael Morris, and Francesca Truffa provided excellent research assistance. All errors are ours.

* Corresponding author.
E-mail address: onder@american.edu (S. Onder).

[1] See Goodwin-Gill and McAdam (2021) for a comprehensive review of legal issues regarding the mobility of refugees.

[2] Return intentions are often surveyed for monitoring refugee sentiment or for scholarly purposes; see Ghosn et al. (2021) for example. However, self-reported intentions may be cognitively biased as shown by Bertrand and Mullainathan (2001), and they can be more context-sensitive than the actual returns. In the case of Syrian refugees, early surveys on return intentions highlighted political transition as a requirement for return (see Yahya et al. (2018), for example), but more recent surveys do not feature this condition.

[3] For an excellent review of this literature, see Dustmann and Görlach (2016).

https://doi.org/10.1016/j.jdeveco.2021.102802
Received 15 April 2020; Received in revised form 30 November 2021; Accepted 2 December 2021
Available online 22 December 2021
0304-3878/© 2021 Elsevier B.V. All rights reserved.

end point of our analysis.[4] Our sample covers those refugees registered by UNHCR in Lebanon, Jordan and Iraq, including returnees and those who stayed.[5]

An inherent challenge in the literature on conflict and forced migration is the absence of a complete longitudinal data set for conditions in countries of asylum and origin that can be mapped onto refugee characteristics. We make progress by combining different sources and types of data. For demographic characteristics of refugees and their arrival and return information, we use administrative data from the Profile Global Registration System (ProGres) database of UNHCR. For the conditions faced by refugees in exile, we use vulnerability surveys conducted by UN agencies in Jordan and Lebanon, and complement these with a new household survey comprising similar demographic and socioeconomic modules but also including vignettes about the drivers of return. Finally, for conditions in Syria, we have compiled a novel monthly conflict events data set to use along with nighttime light emissions data that proxies access to utilities.

First, we use the temporal and spatial variation of the nightlights and conflict events series to build a monthly panel of conditions *inside* Syria. This is used to analyze the impact of changes in conflict and luminosity patterns on monthly returns from districts in Lebanon, Jordan, and Iraq to sub-districts in Syria using ordinary least squares (OLS) and Poisson quasi-maximum likelihood (PQML) count models. Our results show that security, measured in the refugee's home district, is an important determinant of return. A one standard deviation improvement in security (measured by the change in a composite Conflict Events Index (CEI)[6] between the previous two quarters) increases returns by 5.6 percent when using the PQML model. Improved access to electricity, measured at the refugee's home sub-district level, also encourages returns. In particular, a one standard deviation improvement in luminosity (measured analogously to the CEI, comparing changes between the previous two quarters) increases returns by 2.2 percent. The results have the expected signs aligned with the risk-adjusted expected payoff-based explanations in standard models of migration (e.g., Sjaastad (1962) and Borjas (1987)). While there is a risk of reverse causality,[7] it is unlikely that returns lead to an increase in conflict in this setting since the rate of return migration is low (on average three individuals per month per locality). We also show robustness in the analysis to a number of different specifications with different fixed effects, including governorate (in Syria) by year fixed effects, to weaken the common trends assumption we are making in our analysis (following Del Carpio and Wagner, 2015; Tumen, 2016).

Next, we analyze how conditions in the countries of asylum correlate with return probabilities. We use the vulnerability surveys collected by UNHCR in Jordan and Lebanon to construct a proxy for food security and housing quality among refugees residing in either the same governorate (administrative level 1) in Jordan or the same district (administrative level 2) in Lebanon. Interestingly, our results are not consistent with the common perception that harsh conditions in host communities make refugees return. Refugees residing in locations with a higher rate of food security among refugees are actually more likely to return to Syria. We find a similar relationship for refugees with better housing conditions. This analysis is largely descriptive but we take a number of steps to eliminate some sources of bias. For example, when constructing the proxies for conditions in the countries of asylum, we exclude refugees who ultimately return (by 2018 when our data ends) since those refugees may report differently or alter their consumption choices in anticipation of return. We also include district (in Syria) by district (in country of asylum) fixed effects to address potential sorting of refugees across locations based on time-invariant characteristics that may be correlated with the return propensity.

The number of returnees is still very small, and while the early return decisions are of interest in their own right, we caveat that the decision process behind larger, mass returns may be very different. Nonetheless, these initial returns can also help shed light on subsequent returns as suggested by international experience (Harild et al., 2015). For example, in some refugee cases (similar to households),[8] return decisions are staggered with an individual case member returning, while others remain in exile. Our data suggests that, in such cases, subsequent return is initially much more likely to occur when the first returnee (leader) is prime aged (15–64 years old). In particular, it is rare that there are followers when an older member of the core family returns to high conflict areas.[9] By contrast, for cases returning to relatively low conflict areas, we see a similar rate of followers for leaders who are either prime aged or older. This suggests that the profile of those who return early may be predictive of subsequent flows of refugees.

We also compare the drivers of Syrian refugees' actual return decisions with those of their return intentions. To do this, we administered hypothetical vignettes to a representative sample of 1900 refugees in Lebanon and Jordan. The results based on self-reported intentions to return reinforce the findings relying on actual return patterns. For example, conditions in Syria – in particular whether the family's house in Syria was destroyed and the conditions of schools – are found to have a major effect on return intentions.

Our analysis is closely linked with a rapidly growing literature on the economic analysis of forced displacement. This body of work has largely focused on measuring the economic impact of refugees on host countries[10] in the following areas: employment, wage, entrepreneurship and consumption (Braun and Kvasnicka, 2014; Moser et al., 2014; Ruiz and Vargas-Silva, 2015; Borjas and Monras, 2017; Tumen, 2016; Alix-Garcia et al., 2018), the incidence of crime and hostility, Bell et al. (2013), Hangartner et al. (2019), and electoral outcomes (Dustmann et al., 2019). These papers, however, do not consider the return dynamics of refugees.

This paper also contributes to a second, and more established, strand of literature, which focuses on temporary migrations. An important share of this work focuses on the differences between the economic behavior of those migrants who intend to return and those who do not. For example, Galor and Stark (1990) and Borjas and Bratsberg (1996) argue, respectively, that a positive probability (or intention) of future return to the home country leads to higher savings and higher

[4] The latest figures from UNHCR suggest that the return percentage has increased to 5 percent or 282,283 refugees by May 2021. However, this figure is still small in magnitude with the vast majority of refugees remaining in their respective host countries.

[5] In general, return is promoted by UNHCR as the preferred option among the three durable solutions when conditions permit, the other two being local integration and resettlement (Ghosn et al., 2021). While, as of November 2021, UNHCR did not recommend that refugees return to Syria given security concerns, it supported every individual's right to return to their home country.

[6] The CEI is computed for each district-month using principal components analysis of normalized key conflict events, such as light skirmishes, airstrikes, artillery strikes, and chemical attacks, as well as the casualty-count.

[7] Although there is no conclusive empirical evidence for this, case studies sometimes describe tensions associated with refugee returns. For example, Schwartz (2019) argues that increased local tensions led to subsequent displacements of returning refugees in Burundi. However, the Syrian case is different in important ways. In Burundi, return was large scale (more than half of the refugees returned by some estimates), sometimes involuntary, and, for some Burundians, it took place decades after displacement. In Syria, the return so far has been very small and spontaneous, and it has taken place during an active conflict.

[8] UNHCR registers "cases" of refugees as opposed to households, where a case is defined as a group of individuals who are usually blood relatives and are traveling and staying together while in asylum.

[9] Districts are designated as "high conflict", if they were in the top 10th percentile of the CEI averaged over the time period January 2012–March 2018.

[10] See Becker and Ferrara (2019), Verme and Schuettler (2021), and Ruiz and Vargas-Silva (2015) for detailed reviews of the forced migration literature.

Table 1
Statistics for registered refugees.

	Refugee population	Settlement		Returns
		In camp	Out of camp	(As of end-2019)
Jordan	672,023	130,570	541,453	53,058
Lebanon	851,717	–	851,717	53,286
Iraq	249,733	95,482	154,251	38,117

UNHCR (access date: November 9, 2021). The numbers reported here are only those verified or monitored by UNHCR, the actual return numbers may be higher. All numbers by May 2021 except the returns by country which are not updated as frequently as others.

labor force participation among migrants.[11] The evidence also suggests that return probability is driven by a u-shaped relationship with the migrant's income level (Bijwaard and Wahba, 2014) as well as their family ties and lifestyle choices (Gibson and McKenzie, 2011).

This paper continues with a brief summary of the context around the displacement and return of Syrian refugees. Section 3 describes our data and Section 4 describes the characteristics of individuals and cases that returned within our data. We discuss our empirical strategy in Section 5 and our results in Section 6.

2. Syrian refugees in Mashreq: background[12]

When Syria's first "Arab Spring" protests erupted in 2011, it was a country of 20.7 million people. After a growth spell in the preceding decade, GDP per capita reached USD 2806 in 2010. Despite the seemingly improving economic conditions, however, the social unrest escalated rapidly and by mid-2011 a full-scale armed conflict was already unfolding. Since then, the conflict has led to more than 400,000 direct conflict-induced fatalities and millions of indirect casualties, often by means of severe deprivation.

The brutal conflict in Syria has also created the world's largest forced displacement crisis since World War II. As of 2021, over half of the country's pre-conflict population remains displaced. About 5.7 million Syrians are registered as refugees outside of their country and another 6.2 million persons are displaced within Syria's borders. Of the refugees, most (3.7 million) reside in Turkey, with another 2 million in other countries in the Middle East and North Africa.

In proportion to the host country populations, refugees in Mashreq are among the largest displaced groups in the world, as shown in Table 1. In Lebanon, about 0.85 million registered refugees (close to 20 percent of the country's population before the Syrian conflict) inhabit informal settlements which are spread across a large part of Lebanon, with concentrations near the Syrian border and in Beqaa. In Jordan, about a fifth of the 0.67 million registered Syrian refugees stay in three camps (Zaatari in Mafraq governorate, Azraq and Mrajeeb Al Fhood in Zarqa governorate). Those refugees who live outside camps are largely spread across Amman, Mafraq and Irbid governorates. In Iraq, the number of refugees is small compared to the country's total population (about 0.7 percent), but they constitute 5 percent of Iraq's Kurdistan Region population, where almost all refugees reside with 40 percent staying in camps.

Syrian refugees are an economically active population. In 2018, the labor force participation (LFP) rate of Syrian men was 68 percent in Lebanon and 63.3 percent in Jordan (World Bank, 2019). In comparison to the LFP of Syrian men in Syria (79.1 percent), these rates are lower. However, unemployment and underemployment are likely to be more prevalent in Syria, for which (and for Iraq) we do not have comprehensive data. The female LFP has been typically low in Lebanon, Jordan and Syria, at 10, 13 and 12 percent, respectively.

In both Lebanon and Jordan, Syrian labor is concentrated in the manufacturing, construction and agriculture sectors, mostly on an informal basis. In Lebanon, a decree that dates back to 1964 (Decree No. 17561, Article 9), gives a mandate to the Ministry of Labor to announce sectors that are closed to non-Lebanese nationals annually. With resolutions No. 1/19 in 2013 and No. 1/197 in 2014, the only professions open to Syrians were narrowed down to agriculture, construction, and cleaning. In Jordan, a Ministry of Labor decision dated January 4, 2016 restricted the sectors open to non-Jordanians to manufacturing, construction, and agriculture. The Jordanian government has undertaken a number of measures to formalize the labor force participation of Syrian refugees. In February 2016, the government announced the "Jordan Compact", and committed to issue formal work permits to 200,000 Syrian refugees and eased procedures to obtain them.[13] By June 2018, an estimated 105,404 work permits were issued: 29 percent in the agriculture sector, 43 percent in construction and 11 percent in manufacturing.

In addition to labor market activity, refugees are also eligible for assistance through various programs managed by host country governments and international organizations like UNHCR, the United Nations Children's Fund (UNICEF), and the World Food Program (WFP). Basic-needs support takes the form of winterization assistance, cash assistance, and basic needs kits. Eligibility for assistance is typically determined by a set of common indicators of vulnerability with associated thresholds. A survival minimum expenditure basket (MEB) and family size are used to determine the value of the cash transfers. For example, in Jordan, the MEB for a family of four was estimated at 387 Jordanian dinars (JD, equivalent to USD 546) per person per month, and the total size of the transfers (UNHCR and WFP combined), was 196 JD or USD 276 in 2018. Despite these efforts, however, poverty prevails among Syrian refugees. According to World Bank (2019), the extreme poverty rate of Syrian refugees in Jordan (51–61 percent) and to a lesser extent in Lebanon (37–50 percent) remained close to that in Syria (55–67 percent) in 2018.

Under these conditions, and despite the active conflict in Syria, there has been a small, but non-negligible, number of returns to Syria. Although the exact number is not known, as returnees may wish to remain confidential for security reasons, and the country-breakdown of return numbers is not updated frequently, UNHCR announced 282,283 verified spontaneous returns from all countries, including those from Turkey, between January-2016 and May-2021.

3. Data

Return migration decisions are potentially influenced by expected living conditions in both the country of origin and country of asylum, as well as the individual and case characteristics of refugees. To analyze these factors, a complete longitudinal data set is needed, which was not available due to the active conflict situation in Syria. In what follows, we describe the pragmatic approach we adopted to combine different sources and types of data.

3.1. Refugee attributes

We use the Profile Global Registration System (ProGres) database, which is compiled by UNHCR to record each person of concern who approaches it.[14] Our version comprised 2 million Syrian refugees in Lebanon, Jordan and Iraq, with a cutoff date of March 2018.

[11] Bauer and Sinning (2011) and Dustmann (1997) provide empirical evidence for these arguments.

[12] This section refers to the findings of World Bank (2019) unless noted otherwise.

[13] For example, the foreign passport requirement was abolished and exemptions from a series of medical check-ups were granted.

[14] Registration with UNHCR is not mandatory; Lebanon suspended registration in 2015 but UNHCR continued to collect data for the new arrival cases that approached it in order to include them in their assistance provision process.

The ProGres database is a limited administrative database, which functions like a civil register. It includes a broad set of social and demographic characteristics for each recorded individual: e.g., sex, age, marriage status, occupation, and education. In addition, information on their registration status is recorded, including refugee status, arrival and, where applicable, return date, and sub-district-level location information for last residence in Syria and current residence in the country of asylum. In lieu of families or households, the UNHCR groups individuals together as "cases" who are usually blood relatives and who are traveling and staying together while in asylum. ProGres identifies the relationship of individuals within each case (e.g., familial relationships of everyone within a case to the principal applicant, ranging from members of the core family, such as spouses and children as well as parents and siblings, to extended family, such as in-laws and aunts).

Following the initial registration, entries are updated whenever a UNHCR case worker gains new information about the case. Update frequencies vary from one operation to another, with at least 5 percent of all observations being updated in a given month.[15] Therefore, although information on single-shot events like arrival and return dates is fixed, other information like occupation, education and marital status may change over time. In the case of education and marital status, these changes are largely driven by the aging of the refugee. However, the occupation variable is more problematic, since it could refer to current employment or past employment (including in Syria) depending on when it was last updated. Therefore, while we are able to use the demographic and registration information of the ProGres database, we do not use the occupation variable in the analysis.

3.2. Conditions in countries of asylum

We use vulnerability surveys conducted by UN agencies in Jordan and Lebanon. These surveys assess living conditions of registered refugees at the case, household, and individual levels, and monitor protection, shelter, education, health, water and sanitation, as well as poverty and food coping strategies. Our data from the Vulnerability Assessment Framework (VAF) in Jordan comprises two years: 2015 and 2017 (sampled cases: 2163 and 2001), which are comparable. Samples are weighted by the share of refugees in each governorate, and representative at the 95% confidence interval. Data from the Vulnerability Assessment of Syrian Refugees (VASyR) in Lebanon covers three years: 2015, 2016, and 2017 (sampled households: 4105, 4596, and 4966, respectively). VASyR surveys employ a two-stage cluster sampling approach: first, to ensure geographical representativeness, 30 clusters are randomly selected in proportion to refugee population size and, then, 5 to 6 randomly selected households in each selected cluster are visited.

There are a number of challenges when using the VAF and VASyr surveys. The first is the limited comparability of questions across surveys. This limits the number of variables we can use to measure conditions in Lebanon and Jordan consistently. Nonetheless, we are able to proxy for living conditions and access to employment by computing a composite Food Security Index using a principal components analysis (PCA) of normalized food consumption variables. The latter include the average number of meals per day, and the average number of days a week a case did not have to borrow food, restrict portion sizes, limit the number of meals or restrict consumption of adults. A PCA index is also computed for housing conditions, using dummies for whether the case has an acceptable roof and windows, and access to a (private) latrine.

The second problem is the limited sample size of the two surveys. To take advantage of the much larger ProGres database of 2 million refugees, we compute area averages for the above-mentioned case-level host community conditions, aggregating to the smallest possible geographic unit available (district level for Lebanon and governorate level for Jordan). To limit reporting bias, we exclude cases that ever return from the area averages. The information is then matched with all refugees in the ProGres database that have location information in Lebanon and Jordan, yielding a sample of 1.85 million refugees.

The reporting bias that we are concerned about is the possibility that respondents may have felt they were more likely to receive assistance if they reported worse living conditions. The bias could also go the other way if refugees want to signal their gratitude for the assistance they receive. The problem is even more acute given our research question. Those who intend to return may have systematically different tendencies in reporting. It is possible that those who plan to return no longer feel the need to misreport their income, and this would generate a bias in the correlation between the return decision and asylum country conditions. Removing the respondents who return in the subsequent 1–3 years (up to 2018 which we can observe in ProGres) should at least keep the reporting bias constant across geographic areas. We also employ a fixed effect specification which looks at changes in conditions in Lebanon and Jordan, which will remove differences in reporting bias which are time invariant.

3.3. Conditions in Syria

To capture conflict dynamics, we compiled a novel conflict events data set, covering all districts in Syria between January 2011 and August 2018 at a monthly frequency. This data set provides a record of verified conflict-driven casualties, changes in area of control,[16] and key conflict events (light skirmishes, airstrikes, artillery strikes, and chemical attacks) using more than 7000 news items and multiple databases.[17] Whereas casualties are recorded as a count variable in this case, other conflict events are defined categorically with two or more values, e.g., yes, no for presence of combat activity and low, medium, high for the intensity of it.

To capture a complete picture of the conflict conditions in Syria, we computed a Conflict Events Index (CEI) for each district-month using PCA of normalized conflict activity. The components included key conflict events (light skirmishes, airstrikes, artillery strikes, and chemical attacks) and the casualty-count for the district-month. For each refugee we assign the CEI based on their home district, as we are assuming that refugees return to their home sub-district (in Syria).[18]

For non-security-related conditions in Syria, it was not possible to acquire a comparable and geographically comprehensive time series. Instead, we use nighttime lights measurements from the Suomi National Polar Partnership (SNPP) satellite, which was launched by NASA and NOAA in 2011. The satellite uses a Visible Infrared Imaging Radiometer Suite (VIIRS) instrument to collect low light imaging data in spectral bands covering emissions generated by electric lights, excluding stray light, lightning, lunar illumination, and cloud-cover. Temporal averaging is done on a monthly and annual basis starting from April 2012.

For the purposes of this study, we used the monthly data set of the average luminosity count with zonal statistics aggregation to the sub-district (administrative 3) level. The Luminosity Index is then computed by scaling the average luminosity count by the sub-district's area.

[15] We can assume that the information is up-to-date on average for 2017.

[16] We use an indicator variable which is equal to 1 if an area is either contested or under sole control of non-government forces. The omitted category is sole control by the Government of Syria.

[17] These include the following: ACLED, Carter Center Syria Conflict Resolution Database, Institute of War Syria Events Database, University of Maryland Global Terrorism Database, Syrian Observatory for Human Rights Database, Syrian Shuhada Database, The Uppsala Conflict Data Program and The Violations Documentation Center among others. In addition, activity-specific databases have been consulted, including airwars.org and Arms Control Organization timeline of confirmed chemical weapon use in Syria.

[18] The ProGres database does not provide any information on where refugees return to.

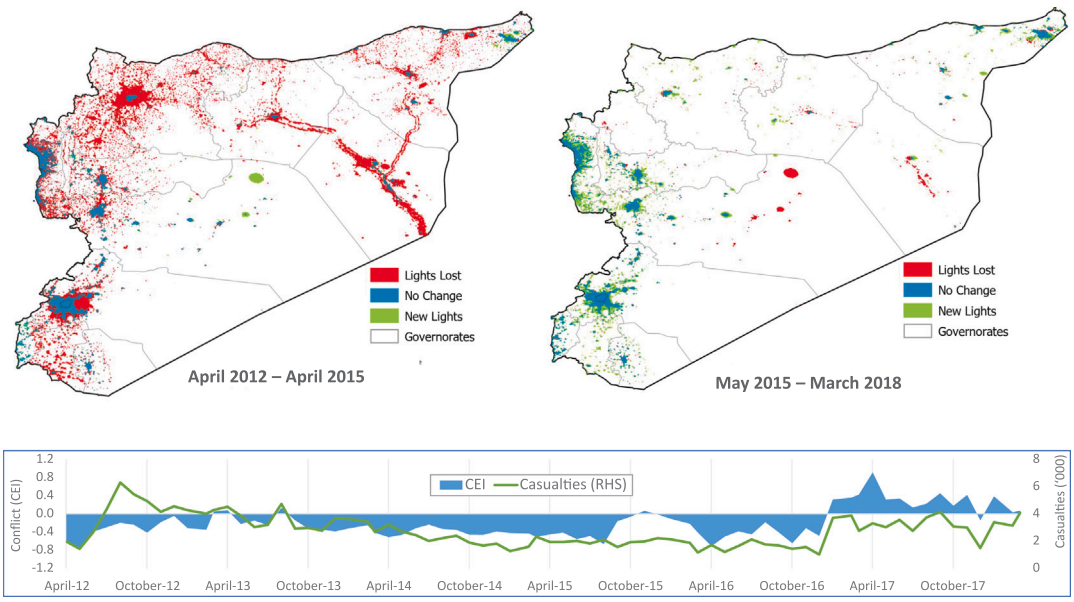

Fig. 1. Trends in night lights, conflict driven casualties and the Conflict Events Index (Apr 2012–Mar 2018). Notes. Night-time low light emissions data are from the Visible Infrared Imaging Radiometer Suites (VIIRS) of the Suomi National Polar Partnership (SNPP). All pixels are equal in size, around 380 m by 380 m (15 arc-seconds by 15 arc-seconds relative to the origin, the mathematical center of the Earth). Casualties show total monthly deaths directly attributed to conflict. The Conflict Events Index (CEI) was computed for each district-month using principal components analysis (PCA) of normalized conflict activity. For this figure, we use the concurrent CEI for that month. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The nightlights in this scheme can be interpreted narrowly as the availability of electricity (grid or generator) or more generally as a proxy measure for the existence of utilities, economic activity or the conflict-driven isolation of a given location. Fig. 1 maps the evolution of visual nightlights and the CEI onto each other. From April 2012 (the first available data point) until April 2015 (midpoint of series), total luminosity decreased by 65% across Syria. By April 2018, about half of these losses were offset, with the exception of areas with persistently intensive conflict like Aleppo and Idleb.

Lastly, Fig. 2 provides an overview of the evolution of the Syrian conflict by year, across districts within Syria. The figure plots the number of casualties per year and shows the variation in conflict intensity both over time and over space. This is the variation we will use in our analysis below. However, it is important to note that several districts have experienced persistently high levels of conflict since its advent. Since return decisions are likely to be impacted by the persistence of conflict, we classify districts into high and low conflict districts, using the top 10th percentile of the mean CEI as a cut-off point.[19] We will use this classification to explore the extent to which return decisions of a case are impacted by the persistence of conflict.

3.4. Survey of refugees in Lebanon and Jordan

In a survey of 1900 Syrian refugees in Jordan and Lebanon, we randomly varied the details of the scenario or vignette presented to a given individual respondent. Some refugee families are certainly more predisposed to wanting to return than others. We instead use hypothetical – but relatable – scenarios, and vary key factors within those scenarios, to help us identify what factors are important to many refugee families when deciding whether to return.

For all respondents in all vignettes, we asked "How likely is this family to return to Syria in the next 2 months?" where the respondent could answer using a Likert scale, ranging from "Very likely" to "Very unlikely". For the analysis below, we use an indicator which is equal to one if a respondent says the family is either very likely or likely to return, and 0 if the respondent says neutral, unlikely or very unlikely.

Each respondent was presented with three vignettes, where key aspects of the scenarios were randomly varied across respondents.[20] These three vignettes were designed to probe the impact of different push and pull factors on the refugees' return decision, allowing us to go beyond the data limitations of the above analysis. That is, the vignettes not only explore the impact of security on return decisions, but also of employment prospects in both the country of asylum and Syria, the status of property in the home community, and the availability of financial assistance.

In particular, the first vignette probes three questions: first, does the ability to work in the host country affect the return decision? Moreover, is the ability to work more or less important among highly skilled workers? Second, we are interested in whether refugees base their return decision on the length of time that security has been stabilized in the origin community. Third, whether financial assistance, and the level of that assistance, affects the return decision.

The second vignette has two key aspects of the scenario which varies across respondents. The first varies whether the wife of a refugee family from Syria, now living in either Lebanon or Jordan (the country was matched to the country where the respondent was currently residing), was working as a housekeeper or stayed home to take care of the family. The second aspect varied the opportunities of the husband of the family to get work back in their home community in Syria. The vignette also sought to understand how a family may decide to send some, but not all, family members to return and elicits the likelihood of each family member to return.

The third and final vignette varied what information a hypothetical family in either Lebanon or Jordan had about their home back in Syria. A respondent was told that the family's house in Syria was either destroyed or intact and unoccupied. The information was provided to

[19] According to this classification, Damascus, Jebel Saman, Deir-ez-Zor, Homs, Al Ma'ra, and Duma are high-conflict districts.

[20] The complete set of vignettes as worded in the survey can be found in the Appendix section "Vignettes of the household survey".

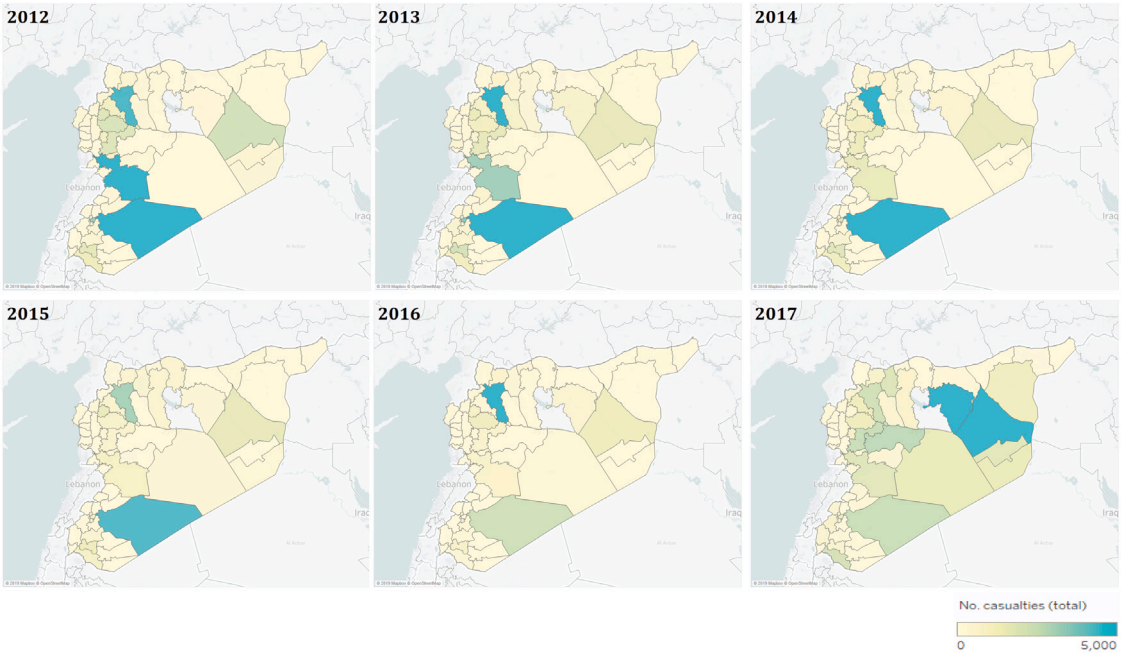

Fig. 2. Casualties by district-year (Jan 2012–Dec 2017). Notes. Data contains district-level casualties for all districts except Dreikish, Safita and Tartous in Tartous governorate (shaded white in the map). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the family either by a resident of the village or from family members who remained in their village in Syria.

Table A.1 provides summary statistics on the sample we interviewed in the survey. Panel A provides descriptive statistics about the households that are comparable to the data available about refugees in the ProGRES database. Panel B describes the statistics of the individuals who responded to the vignette questions themselves, as the opinion of the respondent need not present the opinion of everyone in the household. There are some differences between our sample and the ProGRES data: in particular, the survey has a higher fraction of households with adults with no or informal education, and this is most pronounced in Lebanon.[21]

Overall, our ability to put together a comprehensive data set with key dimensions (micro-characteristics of refugees, conflict dynamics, and the conditions in the countries of asylum and origin) has made the analysis of return decisions possible. The next section will discuss how we leverage the different dimensions of this data set for our purposes.

4. Descriptive analysis of returns

We first begin by providing a descriptive overview of the characteristics of individuals and cases that returned within the ProGRES data. Fig. 3 provides a pairwise comparison between those who returned and those who stayed in terms of their case size, age, and adult education levels. The top panel shows that smaller households constitute a greater share of the returnee sample as compared to the non-returnees. The median case sizes are five for the returnees and 5.3 for the non-returnees. The middle panel shows that whereas children (<15) are a smaller share of the returnee sample, seniors (>55) constitute a larger share, pushing the median age of returnees above that of non-returnees. Finally, the bottom panel shows that the median years of schooling among returnees is lower than that of the non-returnees. Individuals with no schooling comprise about 19 percent of the adult returnee population, while the same category comprise less than 12 percent of the non-returnee population. Table A.2 provides more descriptive statistics on differences between the returnee and non-returnee samples across host countries. We highlight just a few differences here. In Lebanon and Jordan, the returnee sample has more women than the non-returnee sample. By contrast, in Iraq, the returnee sample has more men than women. In all three countries, there are more children in the non-returnee sample than among returnees. There are more older adults (60+) in the returnee sample than in the non-returnee sample; this pattern is in all three countries but is most pronounced in Lebanon.

Individuals and families choose to engage in spontaneous return in many different ways. Qualitative evidence from Harild et al. (2015) suggests that households will sometimes send one family member home to assess the situation on the ground, before other family members return. Fig. 4 decomposes the returns we observe into different scenarios regarding the timing of return. The most common scenario is that an entire case returns: of all returnees, the majority (63%) returned with their entire case at one time.[22] A large share of returns (37%), however, is staggered, i.e., one or more individuals return first, who may then be followed by some or all remaining case members. While group-based return, i.e., more than one member of a case returns first, represent a meaningful share of total returns (16%), there are rarely additional case members (followers) who return to Syria subsequently (1%). By contrast, when single individuals first return, there are relatively more followers who eventually return to Syria (4%).

We focus on the role of early returnees in paving the road for additional case members to return, as this may give us insights into how the spontaneous returns we study in this paper may affect future returns. To further explore this in Fig. 4, we depict the demographic characteristics of "leaders", i.e., the individuals who were the first from their case to return to Syria. In our data, both men and women play important roles as leaders. In particular, prime aged men and women who are core family members within their cases, e.g., spouses, parents, or siblings of the principal applicant, are the most common leaders.

[21] The differences with the ProGRES sample are not driven by the inclusion of unregistered refugees in Jordan: the demographic composition of registered and unregistered refugees in the sample are very similar.

[22] 12% of those cases were cases comprised of single individuals.

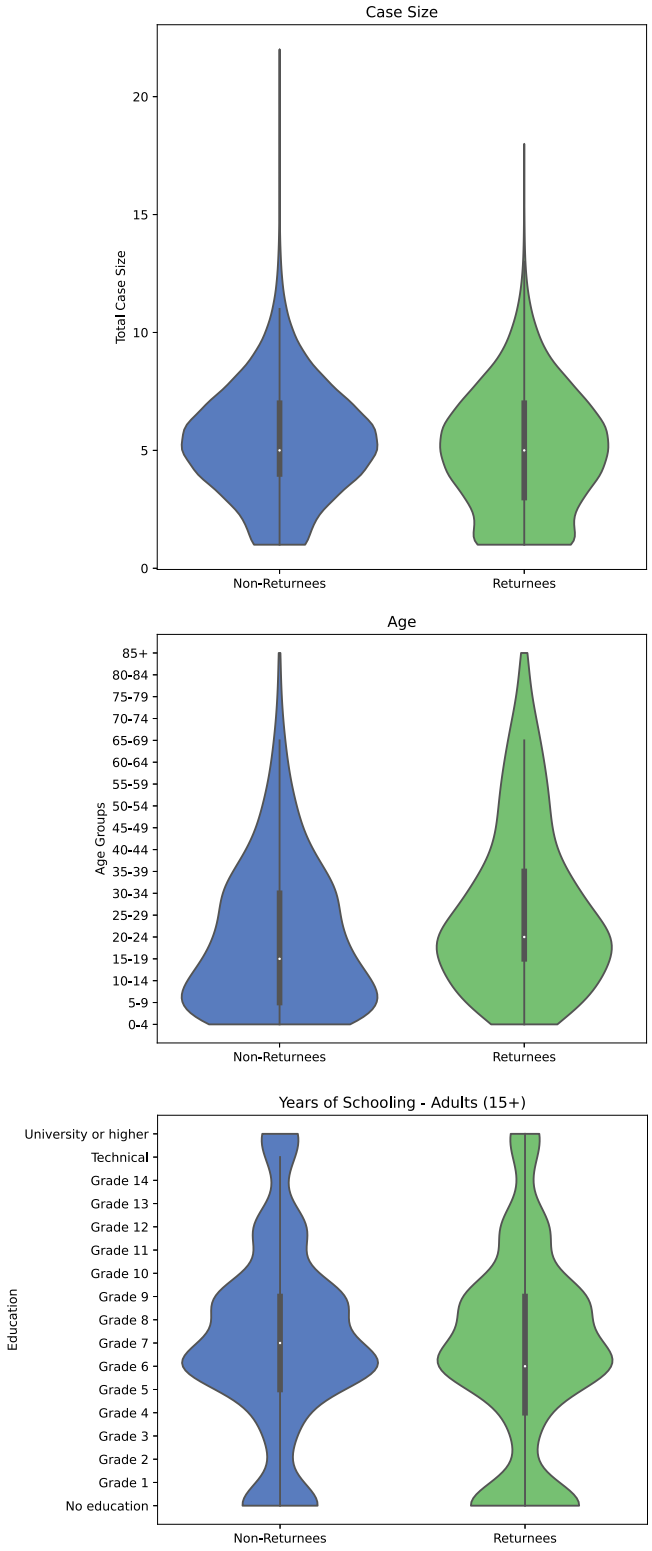

Fig. 3. Case size, age, and adult education: Returnees vs. non-returnees. Notes. The violin plot shows summary statistics of a given indicator. The thin vertical lines denote the full range of observations, over which the sample distributions are show symmetrically on left- and right-hand sides. The thick vertical lines show the range of the middle 50 percent of the observations, and the median values are shown by white dots.

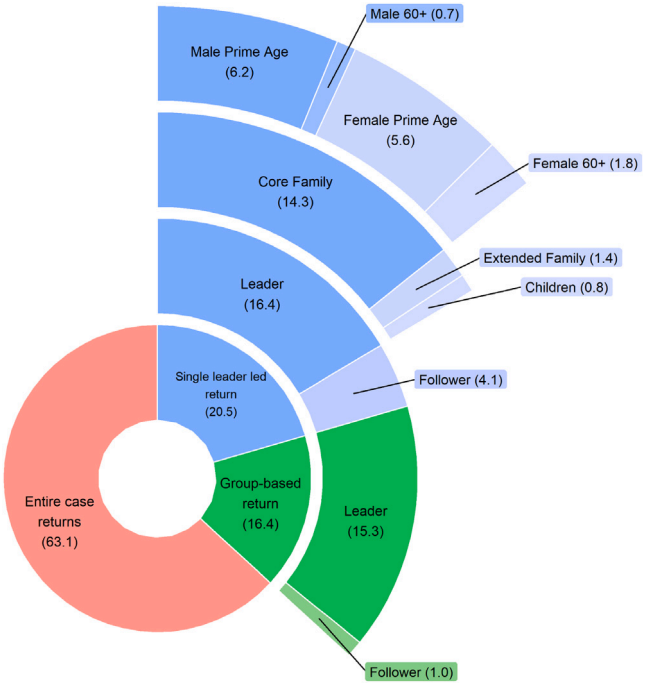

Fig. 4. Breakdown of returnees by timing and demographic characteristics. Notes. For the demographic breakdown in the sunburst chart, the following definitions are used: A "leader" is an individual who was the first from their case to return to Syria. A "follower" is an individual who returned to Syria after a fellow case member returned. Core family members include the principal applicant, his or her spouse and children, as well as his or her parents, siblings, nephews, and nieces. Extended family members include all other blood relatives, such as grandparents, aunts and uncles, as well as in-laws.

In Fig. 5, we show that the likelihood of case members returning differs depending on who within the case is the initial leader and whether the individual's home community in Syria has a high or low conflict intensity. In the figures, we present the estimated Kaplan–Meier survival functions of the likelihood of individuals returning in different scenarios. In Panel A, we look at cases in which there is an initial male returnee who is a core member of the family and who returns to high conflict areas. How likely are other members of the case to return? For the majority of the time period, subsequent return is much more likely to occur when the initial leader is prime aged. We see this also for women leaders in Panel C. Overall, it is rare that there are followers when an older member of the core family returns to high conflict areas. By contrast, for cases that originated in areas in Syria which are relatively low conflict, we see a more similar rate of followers for leaders who are either prime aged or older, for both men and women (Panels B and D).[23]

Subsequent returns do not differ significantly between core family member leaders and extended family leaders in low conflict areas (Panel F); however, in high conflict areas, there is more follow up return after the old age leaders of extended family members (Panel E) compared to the old age leaders of core family members (Panels A and

[23] Table A.3 provides more descriptive statistics on difference between the returnee and non-returnee samples by conflict intensity. There are differences between who leaves their origin community depending on how conflict-affected the area is. However, the differences are fairly small in magnitude. For example, the fraction of refugees who are children is higher from high conflict areas than low conflict areas, a difference which is statistically significant ($p < .01$). However, the magnitude is not: in high conflict areas, the fraction of refugees who are aged 0–14 is .480 while the fraction of refugees who are children originating from low conflict areas is .472. By contrast, in Table A.2 the fraction of returnees who are children is only .13–.31 across Lebanon, Iraq and Jordan.

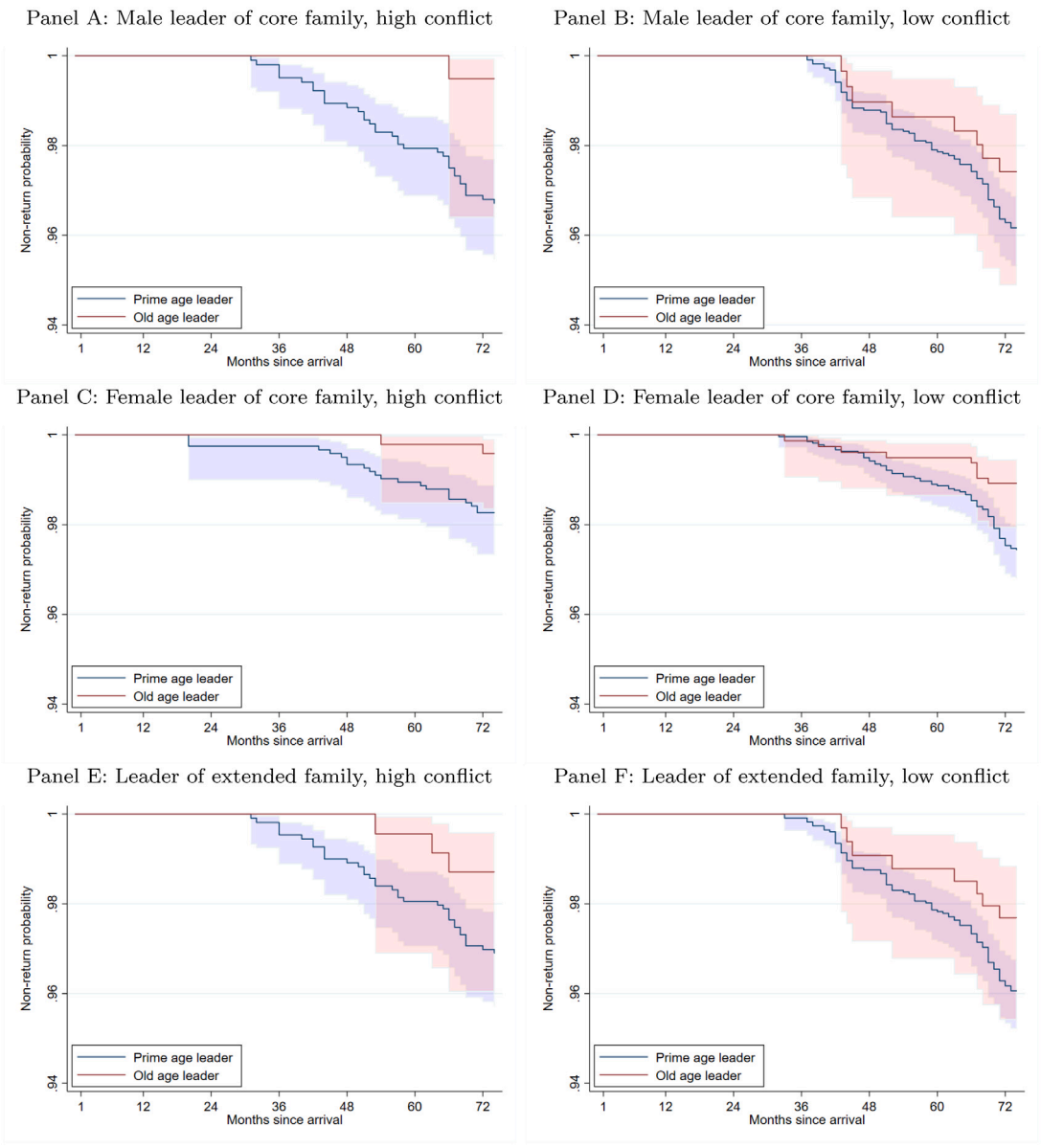

Fig. 5. Estimated Kaplan–Meier survivor function, by gender, family relationship and conflict. Notes. The graph displays the Kaplan–Meier product-limit estimate of the survivor function for all refugees registered with UNHCR in Lebanon, Jordan and Iraq. We use survival analysis to estimate the transition probabilities, since this approach is able to address the right-censored nature of our panel (i.e., the majority of individuals did not return by the end of our records), which is problematic when, for example, using OLS or a binary dependent variable model, such as logit or probit. The survivor functions are reported separately for different conflict intensities in the home district, where a "high conflict" district is in the top 10th percentile of the Conflict Events Index (CEI) averaged over the time period January 2012-March 2018. The shaded area represents the 95% confidence interval.

C). Overall, these observations are aligned with the qualitative evidence on complex refugee return patterns, and they suggest that the nature of early returns can shed light on the pace of subsequent returns.

5. Empirical strategy

Our primary analysis uses a panel data set, where the unit of analysis is the pair of: sub-district level within Syria and district in the country asylum level.[24] In order to understand the relationship between returns and security and access to utilities in Syria, we estimate the following specification:

$$\begin{aligned} ln(returns_{scmt}) = {} & \alpha + \beta_1 \Delta CEI_{dmt} + \beta_2 \Delta LuminosityIndex_{smt} \\ & + \beta_3 share_{scmt} + \beta_4 winter_{mt} + \beta_5 AoC_{smt} + \delta_{dc} \\ & + \tau_{gt} + \varepsilon_{scmt}, \end{aligned} \tag{1}$$

where $returns_{scmt}$ is the number of refugees who returned to Syria in month m of year t originally from sub-district s in Syria who registered with UNHCR in district c within their country of asylum (CoA). d represents the district within Syria, which is one administrative level larger than s. Since refugees will make the decision to return home based on past conflict events and recent changes in standards of living like electricity reliability, we look at a lag of both the Conflict Events Index (CEI) and the Luminosity Index. In particular, we construct ΔCEI_{dmt} as the change in the CEI for district d between the quarter immediately prior to month m in year t from the previous quarter, with

[24] Syria is a unitary state, but for administrative purposes it is divided into 14 governorates, which are further divided into 65 districts and 281 sub-districts.

a lag embedded by construction.[25] $\Delta LuminosityIndex_{smt}$ is analogously constructed using the same lagged time periods. $share_{scmt}$ measures the share of refugees from sub-district s that are located in district c within their country of asylum in month m of year t. We will discuss this control in Section 6.1. $winter_{mt}$ is a dummy for the winter months (December–March), during which returns may slow due to inclement weather. This aims to control for any seasonal variation in refugee flows. As an additional control, we include AoC_{smt}, which is a series of Area of Control dummy variables to capture who is in control of sub-district s in month m in year t. These include a dummy for who controls the area (i.e., only government forces, only non-government forces, or contested). The omitted category is sole control by the Government of Syria, which held about half of the sub-districts during our sample period.

We show results using different sets of fixed effects, including ones at the home district d*CoA district c level, δ_{dc}, to control for any time-invariant characteristics that might determine why refugees from district d choose to locate in the CoA district c. We also use different time controls, including governorate in Syria g by year fixed effects, τ_{gt}, to weaken the common trends assumption we are making for conflict dynamics within Syria. Standard errors are clustered at the district d level. Finally, we exclude refugees originating from Daraa, Syria, which carried a special status during the conflict, for the sake of a more general representation of the conflict-return relationship.[26]

An important challenge we face in this analysis is the low incidence of actual returns. With less than 4% return records in our registration data, 67% of sub-districts-month pairs in Syria have zero returns. Of all 257 sub-districts,[27] 29 sub-districts (11%) have no returns during the entire sample period. The mean return in the entire sample period is three individuals in a given month for a given sub-district, with a variance of 281. To address this problem, we estimate a Poisson Quasi Maximum Likelihood (PQML) count model, with robust standard errors clustered by district d (Wooldridge, 1999). Specifically, we estimate by Maximum Likelihood estimator equations such that

$$E(returns_{scmt}) = \delta_{dc} exp(\beta_1 \Delta CEI_{dmt} + \beta_2 \Delta LuminosityIndex_{smt} + \beta_3 share_{scmt} + \beta_4 winter_{mt} + \beta_5 AoC_{smt} + \delta_{dc} + \tau_{gt}), \quad (2)$$

where the variables are defined as in Eq. (1). Note that the β coefficients in Eq. (2) represent the semi-elasticity of returns with respect to changes in conflict intensity or luminosity in the origin district. The PQML count model is particularly suitable because it is robust to arbitrary distributional assumptions so long as the conditional mean is specified by Eq. (2).

In the second part of the empirical analysis we use the same panel and econometric specifications to look at the relationship between returns and conditions in the country of asylum. We restrict this analysis to the Syrian refugees based in Jordan and Lebanon as their host country conditions are proxied by the geographical aggregates computed from the VAF and VASyr surveys. Despite this restriction, more than 85 percent of all refugees in our data set is included in the analysis. For our preferred specification we, thus, estimate by Maximum Likelihood estimator equations such that

$$E(returns_{scmt}) = \delta_{dc} exp(\beta_1 \Delta CEI_{dmt} + \beta_2 \Delta LuminosityIndex_{smt} + \beta_4 FoodIndex_{ct} + \beta_5 HousingIndex_{ct} + \beta_5 share_{scmt} + \beta_6 winter_{mt} + \beta_7 AoC_{smt} + \delta_{dc} + \tau_{gt}), \quad (3)$$

where the variables are defined as in Eq. (1). In addition, we include the Food Security Index ($FoodIndex_{ct}$) and Housing Quality Index ($HousingIndex_{ct}$) for each district c in the country of asylum and year t to proxy for host country conditions. The Food Security Index is computed using PCA of normalized food consumption variables at the case level, including the average number of meals per day, and the average number of days a week a case did not have to borrow food, restrict portion sizes, limit the number of meals or restrict consumption of adults. The Food Security Index is then aggregated for all non-returnee cases to the smallest possible geographic unit available (district level for Lebanon and governorate level for Jordan). The Housing Quality Index is computed analogously using dummies for whether the case has an acceptable roof and windows, and access to a (private) latrine for the PCA. It is important to note that the conditions in the country of asylum may be the result of a refugee's anticipated length of stay in the host country. As such, we discuss these results in Section 6.2 as correlations and not necessarily causal relationships.

6. Results

6.1. Conditions in Syria

The evolution of the security situation and overall quality of life are likely important factors for refugees to consider returning home. We therefore start this analysis by looking at how the return decision varies as a function of our composite measure of security (the CEI) and the luminosity measure using nightlights.

Table 2 shows an overall robust relationship between security and returns. In column (2), we find that a one standard deviation decrease in the quarterly Δ CEI – that is a standard deviation improvement in conflict conditions relative to the previous quarter – increases returns by 6.2 percent. A challenge in estimating the effect of conflict in Syria is that there are also push factors from the country of asylum that is also affecting return, as we discuss in Section 6.2. We at least partially address this concern in multiple ways. First, the $share_{smct}$ variable is intended to capture push factors from host countries, if refugees from a particular sub-district within Syria s make up a large share of the overall refugees in a locality c within a country of asylum.[28] Columns (2)–(7) show that the results are robust to a variety of different controls and fixed effects.

In our preferred specification, column (8), we also include home sub-district-by-district in the CoA fixed effects to address the possibility that refugees from certain places within Syria resettle in the same areas which may also have a differential propensity to return (say those who settle near the border). The results are very similar: we find that a one standard deviation decrease in the quarterly Δ CEI – that is a standard deviation improvement in conflict conditions relative to the previous quarter – increases returns by 5.6 percent.[29] These effects are small but precisely estimated in OLS.

Access to electricity, and by extension utilities (measured at the sub-district level), also encourages returns. In particular, our preferred specification column (8) demonstrates that a one standard deviation

[25] Note that the conflict data is only available at the district level.

[26] Daraa was part of a deescalation zone established by the governments of the United States, Jordan and the Russian Federation. Unlike the other three deescalation zones created as part of the Astana Process, Daraa shared a contiguous border between origin and asylum locations for a large group of Syrian refugees for the duration of our analysis. This made movements between the two regions, including the delivery of international aid, less complicated as crossing government-controlled areas was not necessary (Bojicic-Dzelilovic and Turkmani, 2018). The patterns of asylum seeking and returns are therefore distinct in this area, and less correlated with the time series conflict pattern.

[27] Not all sub-districts are included in the analysis, as there is missing conflict data at the district level (Dreikish, Safita and Tartous districts). Moreover, while the ProGres database represents the universe of registered refugees, refugees do not necessarily hail from all districts in Syria with two district missing (Salkhad and Shaba).

[28] This variable is correlated with returns but the estimates of ΔCEI and $\Delta LuminosityIndex$ change very little with or without the variable. These results are available from the authors upon request.

[29] To address possible endogeneity concerns between return and the proxies of conflict intensity, we also compute the CEI without casualties. Our results are robust to the inclusion of this revised CEI.

Table 2
Security and quality of life in Syria and aggregate returns.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
Δ Conflict Events Index$_{dmt}$	−0.002	***	−0.062	***	−0.001	***	−0.057	**	−0.001	***	−0.056	**	−0.001	***	−0.056	**
	(0.000)		(0.023)		(0.000)		(0.025)		(0.000)		(0.024)		(0.000)		(0.022)	
Δ Luminosity Index$_{smt}$	0.004	***	0.043	***	0.004	***	0.049	***	0.004	***	0.049	***	0.003	**	0.022	**
	(0.001)		(0.014)		(0.001)		(0.014)		(0.001)		(0.013)		(0.001)		(0.011)	
Share$_{scmt}$	0.767	***	6.548	***	0.767	***	6.694	***	0.584	***	5.097	***	0.585	***	5.101	***
	(0.126)		(0.808)		(0.126)		(0.766)		(0.121)		(1.088)		(0.121)		(1.099)	
Winter													−0.008	***	−0.435	***
													(0.001)		(0.035)	
Constant	0.006		−1.845	***	0.006		−1.817	***	0.010	**	−1.174	***	0.014	***	−1.033	***
	(0.004)		(0.216)		(0.004)		(0.205)		(0.004)		(0.236)		(0.004)		(0.232)	
Observations	518,615		518,615		518,615		514,865		518,615		326,738		518,615		326,738	
Clusters	52		52		52		52		52		52		52		52	
Model	OLS		PQML		OLS		PQML		OLS		PQML		OLS		PQML	
Area of Control	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Year fixed effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
Governorate*Year fixed effects	No		No		Yes		Yes		Yes		Yes		Yes		Yes	
District SYR fixed effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
District CoA fixed effects	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
District SYR*District CoA fixed effects	No		No		No		No		Yes		Yes		Yes		Yes	

Notes. The dependent variable is a monthly panel of aggregate returns from a district (administrative level 2) in Lebanon, Jordan, or Iraq c to a sub-district (administrative level 3) in Syria s, which uses ProGres data. OLS specifications use the logs of returns. The Conflict Events Index (CEI) is computed for each district in Syria d in month m in year t using principal components analysis of normalized conflict events including chemical attacks and airstrikes. The Luminosity Index is computed for each sub-district s using average nightlights for month m in year t scaled by the sub-district's area. The Δ Conflict Events Index and Δ Luminosity Index reflect changes in their respective levels between the quarter prior to the month of return and the quarter preceding it. For each refugee we assign the CEI based on their home district, as we are assuming that refugees return to their home sub-district (in Syria). The $Share$ variable captures the percent of refugees from sub-district s located in district c in the country of asylum in month m in year t. All specifications include dummy variables capturing whether sub-district s is contested or under the sole control of non-government forces (omitted category is "Sole control by the Government of Syria") in month m in year t. The $Winter$ dummy captures the months when it snows in Syria (December-March). Daraa, which carried a special status during the conflict, is excluded. The decline in the sample size for the PQML regressions is mechanical as the maximum likelihood estimator (different to OLS) automatically drops observations with insufficient variation, which occurs in this case due to the large number of fixed effects and many zeros in the dependent variable. Robust standard errors are clustered at district d level and reported in parentheses. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

increase in luminosity, relative to the level of luminosity two quarters ago, increases returns by 2.2 percent. These results are robust to less saturated econometric specifications, as shown in columns (1)–(6). The net takeaway from this analysis is that an omnibus measure of quality of life, proxied by nightlights, is a factor in refugees' decisions to return home even in the presence of ongoing conflict at the country level.

6.1.1. Vignette analysis

We complement the study of returns that have already happened with data from hypothetical vignettes. There are advantages and disadvantages of both of these data sources. The returns observed to date are still a very small percentage of the overall refugee population, and therefore the factors affecting their return decision may not be representative of the larger Syrian refugee population. The hypothetical vignettes, while clearly weaker in that they represent hypothetical scenarios, seek to provide insights into the factors which will affect the return decisions of refugees going forward.

The data is analyzed using a straightforward regression specification:

$$Pr(y_i = 1|X) = \quad \alpha + \beta_1 VignetteScenario_i + \gamma Jordan_i + \varepsilon_i \tag{4}$$

where y_i is an indicator variable = 1 if the respondent i reported that the family depicted in the vignette was Very likely or Likely to return to Syria in the next two months. The variable Jordan = 1 if the respondent resides in Jordan and = 0 if the respondent i currently resides in Lebanon. The different scenarios are captured by either a dummy variable $VignetteScenario_i$ or a series of dummy variables. β_1 captures how changes in a refugee family's conditions – either in their country of asylum or back in Syria – affect the perception that the refugee will return to Syria. We also show the results separately for Jordan and Lebanon.

In Table 3, we look at whether the condition of the family's house back in Syria affects the return decision. The reference group in this table is the scenario where the family hears from their former neighbors that their house is still intact. Column (1) shows that in this scenario, 38% of respondents say that the hypothetical household is likely or very likely to return to Syria in the next 2 months. However, finding out that their home was destroyed has a substantial negative effect, reducing the probability of expected return by 22 to 23 percentage points. This constitutes a 60% reduction in respondents' stated perception of the family's likelihood to return. The source of the information, neighbors or extended family still in the village, did not matter. The results are similar for both refugees residing in Jordan (column 2) and Lebanon (column 3).

Table 4 looks at another aspect of life in Syria: whether schools are open and functioning well. In column (1), we see that 43% of respondents who are told about a scenario in which a refugee whose

Table 3
Hypothetical return decisions, by source of information on condition of family home (vignette 3).

	(1)		(2)		(3)	
Family says house intact	0.020		0.015		0.023	
	(0.028)		(0.039)		(0.040)	
Neighbor house destroyed	−0.224	***	−0.229	***	−0.219	***
	(0.028)		(0.040)		(0.039)	
Family house destroyed	−0.229	***	−0.254	***	−0.204	***
	(0.028)		(0.039)		(0.039)	
Observations	1900		950		950	
Sample	All		Jordan		Lebanon	
Mean: Wife is working and Schools in	0.381		0.385		0.377	
SD	0.486		0.488		0.486	

Notes. The hypothetical return decisions are investigated in a survey where each respondent was presented with three hypothetical scenarios or vignettes and was asked about the likelihood of return for the family depicted in the vignette using a 5-point Likert Scale (the complete set of vignettes as worded in the survey can be found in the Appendix.). For all regressions, the dependent variable is a dummy capturing whether a respondent reported that the family depicted in vignette 3 was "Very likely" or "Likely" to return to Syria in the next two months. All regressions include dummy variables capturing the housing condition and source of information (the omitted category is "Neighbor says house is intact") and whether the refugee resides in Jordan (the omitted category is "Lebanon"). Standard errors are reported in parentheses. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

Table 4
Hypothetical return decisions by access to education in Syria (vignette 2).

	(1)		(2)		(3)	
Wife not working	0.012		0.025		0.000	
	(0.022)		(0.030)		(0.031)	
Schools poor resources	−0.187	***	−0.128	***	−0.245	***
	(0.021)		(0.030)		(0.031)	
Observations	1900		950		950	
Sample	All		Jordan		Lebanon	
Mean: Wife is working & Schools in Syria Open	0.433		0.372		0.498	
SD	0.496		0.484		0.501	

Notes. See notes to Table 3 for details on specification and variable construction.

wife is working in their country of asylum and who hears that schools in Syria are open say that the depicted family are likely or very likely to return in the next two months. Whether the wife is working or not in the country of asylum does not have a significant effect on the reported likelihood of return. However, the vignette highlights how schools in Syria affect the return decision: respondents are 19 percentage points less likely to expect the hypothetical household to return when the schools are under-resourced. This is more than a 40% reduction in the likelihood of expected return. We see schools could be an important factor in the return decision among refugees in both Jordan (column 2) and Lebanon (column 3).

Overall, the two vignettes signal that specific conditions back home in Syria have a large and economically meaningful impact on the return decisions.[30] Camarena and Hägerdal (2020) also find that economic opportunities in the home community encouraged return among Christian Lebanese.

6.2. Conditions in host countries

A refugee's livelihood opportunities and housing conditions in the host country may also be important determinants of her return decision. In fact, this point often appears in the popular media in different forms like "good conditions make refugees stay" and, by extension, "bad conditions make refugees return".[31] In this section, we provide suggestive evidence, from both the analysis of actual returns and vignette analysis, that does not necessarily support this view.

We analyze how conditions in the host countries affect aggregate returns by using the same panel as Table 2. However, because the country of asylum indicators are reported for only 2 years for Jordan and 3 years for Lebanon, the time variation in this panel is severely constrained. Nevertheless, we pursue the panel estimates since they remove time invariant characteristics which we do not observe and are correlated with the return decision. This includes reporting biases, since UNHCR provides assistance to the respondents, that are constant over time. Results are shown in Table 5. Better livelihood opportunities - proxied for by the food security PCA index - is positively associated with the number of returnees. Housing conditions also show a positive association with the aggregate return numbers, as shown in column (4), though the finding is only strong using the PQML specification. However, since there remains concern that changes over time could reflect changes in refugee characteristics in that locality, we interpret these findings with caution.

Table 5
Quality of life in country of asylum and aggregate returns.

	(1)		(2)		(3)		(4)	
Food security index	0.018	*	0.784	***				
	(0.009)		(0.273)					
Housing quality index					0.013		1.298	***
					(0.012)		(0.428)	
Constant	0.046	***	−0.475	**	0.046	***	−0.295	
	(0.010)		(0.202)		(0.010)		(0.204)	
Observations	133,450		90,903		133,450		90,903	
Cluster	53		50		53		50	
Model	OLS		PQML		OLS		PQML	
District SYR*District CoA fixed effects	Yes		Yes		Yes		Yes	

Notes. The dependent variable is a monthly panel of aggregate returns from a district in Lebanon or Jordan c to a sub-district in Syria s, using ProGres data; see notes of Table 2 for more details. The Food security index is computed by using principal components analysis of normalized food consumption variables, including the number of meals for an average day. The Housing quality index is computed analogously using housing condition variables, including dummies for acceptable windows and roof. All regressions control for the Δ Conflict Events Index, Δ Luminosity Index, the share of refugees from the sub-district s located in district c, dummy variables capturing whether sub-district s is contested or under the sole control of non-government forces, a dummy variable for the winter months, as well as governorate-by-year, and home district-by-country of asylum district fixed effects. Robust standard errors are clustered at district d level and reported in parentheses. *** significant at 0.01 level, ** significant at 0.05 level, * significant at 0.1 level.

Table 6
Hypothetical return decisions: Occupation and security.

	(1)		(2)		(3)
Miner in Syria now working with permit	0.053	*	0.088	**	0.021
	(0.032)		(0.044)		(0.046)
9 months of security	0.033		0.052		0.016
	(0.027)		(0.039)		(0.039)
12 months of security	0.008		0.058		−0.045
	(0.028)		(0.039)		(0.040)
Observations	1900		950		950
Mean: Miner in Syria now working no permit	0.401		0.316		0.490
SD	0.491		0.466		0.501
Sample	All		Jordan		Lebanon

Notes. See notes to Table 3 for details on specification and variable construction. Also included are indicators for the physician vignettes.

The vignettes also present evidence that improvements in conditions in the country of asylum may not discourage return. The first scenario to consider is one where the household head of a refugee family was previously a miner in Syria and was either working with or without a permit in the country of asylum. Table 6 suggests on average, about 33% of respondents think this household is likely or very likely to return to Syria when the household head has no permit. This percentage is quite a bit higher in Lebanon (column 3, 42%) than in Jordan (column 2, 25%). Column 1 suggests that access to a work permit is positively associated with returning. This estimate is only significant at the 10% level, because there is a strong effect (8.8 percentage points, standard error of .044) in Jordan but no effect (positive coefficient of .02 with standard error of .046) in Lebanon. This is additional evidence that (relatively) good conditions in the country of asylum may not be strong barriers to return.

These results on net do not support the view that poor living conditions in host countries push refugees to spontaneously return to their origin country.

Finally, the vignettes also point to a future role of the international community in facilitating the return process. Following the first vignette, we informed respondents that an international organization is

[30] Table 6 looks at vignette 1, where we varied how conditions at home were described in very vague terms "The family has been monitoring the situation in their home town and heard that the security situation has dramatically improved in the last ... *randomized between* six months, nine months and 12 months." We do not observe significant differences on anticipated returns across these different scenarios. It may be because the information we gave them about the conditions were too vague, our preferred interpretation, or that the survey simple did not pick up the relevant time frames. I.e. there is no difference between six months and 12 months in the minds of refugees but instead the relevant margin is three months or 18 months, for example.

[31] See Berry et al. (2015) for a review of press coverage about Syrian refugees until 2015.

Table 7
Hypothetical return decisions: financial assistance.

	(1)		(2)	(3)	
2100 USD cash assistance per returnee	0.082	***	0.006	0.158	***
	(0.022)		(0.031)	(0.032)	
Observations	1900		950	950	
Mean: 1000 USD cash assistance per returnee	0.356		0.326	0.388	
SD	0.479		0.469	0.488	
Sample	All		Jordan	Lebanon	

Notes. See notes to Table 3 for details on specification and variable construction.

offering relocation cash assistance. Table 7 shows that offering $2100 per family member compared to $1000 per family member led to an increase by 8 percentage points, or 23 percent, in the likelihood that the respondent thought the family would return to Syria (under conditions of good security in Syria of at least 6 months, which is obviously very different than conditions in 2018 when the survey was conducted). The result is driven by respondents in Lebanon, where 55% of respondents think the family would be likely to return to Syria with a $2100 cash relocation assistance — compared to only 39% of respondents who heard the family would receive $1000 per family member. There is no difference among the respondents in Jordan to the $1000 or $2100 assistance, and the overall rate of anticipated return is lower, at 33%.

6.2.1. Discussion of the role of host country conditions in return decision

The results have so far shown that the effects of conditions at the origin, such as security, on returns are as expected. Other things being equal, an increase in risk-adjusted payoffs from return (delivered by better security and living conditions) tends to increase the probability of return. The opposite, however, may not be true for conditions in countries of asylum, where improvements in payoffs from staying is associated with more returns. While our results must be interpreted with caution given concerns about endogeneity, this section provides a simple framework for thinking about why this could be true.

The core idea is that, for refugees with incomes at the low end of the distribution, the costs associated with return by themselves can generate the result we observed in the data. While it is not possible to detect or estimate such costs and incomes by using our data set, a descriptive literature on the rules and regulations governing the return of Syrian refugees provides ample facts in support of the transaction cost argument. For instance, about 70 percent of Syrian refugees in Lebanon and Jordan are reported to lack basic civil documentation, which is required for returns (Norwegian Refugee Council, 2017). Moreover, the cost of accessing those documents could be prohibitively high: with a $325 price tag, the Syrian passport is one of the costliest passports to acquire in the world ($825 if expedited). To put that into perspective, the average cash transfer received by refugees is $27 per person per month. Thus, such costs alone can impede returns in a trivial manner: as refugees are credit constrained, they may not be able to afford the return. However, there is also a non-trivial effect: given the trade-offs they face; refugees may not be willing to return even when transaction costs are affordable.

To see this last point, let us consider a simple dynamic environment with 2 periods, where the second period has a variable length $l > 0$, reflecting differences in the planning horizon (i.e., the age of the refugee). Each refugee i is endowed with an income w_i in the first period. If a refugee stays in the host country, she is endowed with the same endowment w_i in the second period as well.

Alternatively, refugees may choose to return to their country of origin at the end of the first period. Once attempted, this return may succeed with a probability $\pi_i \in [0,1]$. Successful returnees then receive an endowment v_i in the second period of their lives in the country of origin. Ruling out access to credit and transfers of resources across periods or countries, we can now define the lifetime utilities in different situations as follows

$$W_i = u(w_i) + \begin{cases} lu\left(w_i\right), & if\ no\ return \\ lu\left(v_i\right), & if\ successful\ return \\ 0, & if\ failed\ return \end{cases} \tag{5}$$

where $u(.)$ is the period utility function with $u'(.) > 0$ and $u''(.) < 0$.

In this simple framework, net utility gains from returning $\Delta_i = l\left[\pi_i u\left(v_i\right) - u\left(w_i\right)\right]$ increases in π_i, v_i and l and decreases in w_i provided that $v_i > w_i$. That is, improving conditions in the country of origin should increase returns and improving conditions in the country of asylum should decrease returns. Similarly, a longer life expectancy and the likelihood of a successful (safe) return also increase returns.

Let us now introduce a mobility cost τ, e.g., which can represent the actual logistical costs of the return trip or other factors, e.g., psychological burden, in monetary terms. The ex-ante life-time utility of return is then given as:

$$W_i = u\left(w_i - \tau\right) + \pi_i l u\left(v_i\right) \tag{6}$$

and the ex-ante net gain from return is:

$$\Delta_i = \left[u\left(w_i - \tau\right) - u\left(w_i\right)\right] + l\left[\pi_i u\left(v_i\right) - u\left(w_i\right)\right] \tag{7}$$

where the first term on the right hand side shows the welfare cost of attempting the return in the first period and the second term shows the expected increase in second period welfare. Like the case without mobility costs, the expected net gain from return increases monotonically in l and pull factors v_i and π_i. An increase in the mobility cost decreases Δ_i; however, w_i now has ambiguous effects:

$$\frac{\partial \Delta_i}{\partial w_i} = \left[u'\left(w_i - \tau\right) - u'\left(w_i\right)\right] - lu'\left(w_i\right) \gtrless 0. \tag{8}$$

Intuitively, a small improvement in the asylum conditions has two effects on the comparison of expected lifetime utilities across options. First, it reduces the welfare cost of the return decision in the first period (the term in brackets on the right hand side). That is, the welfare effect of reducing own consumption by a given amount (τ) is smaller when the first-period income is greater. Second, it makes the return option less attractive (the term outside the brackets on the right hand side) as the gap between the second period utilities in return and no-return scenarios decreases with a higher w_i. When l is sufficiently small and $u(.)$ satisfies the Inada conditions, the following properties are observed for a given τ:

$$\lim_{w_i \to \infty} \frac{\partial \Delta_i}{\partial w_i} < 0 \tag{9}$$

$$\lim_{w_i \to \tau^+} \frac{\partial \Delta_i}{\partial w_i} > 0 \tag{10}$$

When the initial income is high, i.e., $w_i \to \infty$, the second effect discussed above dominates the first one following a small increase in w_i and return becomes less attractive. In contrast, starting from an income that is too close to τ, i.e., $w_i \to \tau$, a small increase in income relaxes the welfare cost of returning in the first period drastically, which dominates the second effect. Thus, for those at the lower end of income spectrum, a higher income in exile can make return more attractive.

Fig. 6 shows the effect of explicitly recognizing the mobility cost for low income levels. The first panel shows the payoffs associated with not returning (W^{nr}_{NC}), which are identical regardless of mobility costs, and the payoffs for a returning refugee with (W^{r}_{C}) and without (W^{r}_{NC}) mobility costs. Inclusion of a mobility cost makes the return payoff steeper than the staying payoff for low income levels. As a result, as shown in the second panel, whereas the net gain from return is a monotonically decreasing function of the host country income in the case with no mobility costs (Δ^{r}_{NC}), it becomes non-monotonous in the case with mobility costs (Δ^{r}_{C}). As a result, for low income levels, a small increase in host country income can make return more desirable. In fact, for some ranges of v_i and τ, a double crossing of the stay

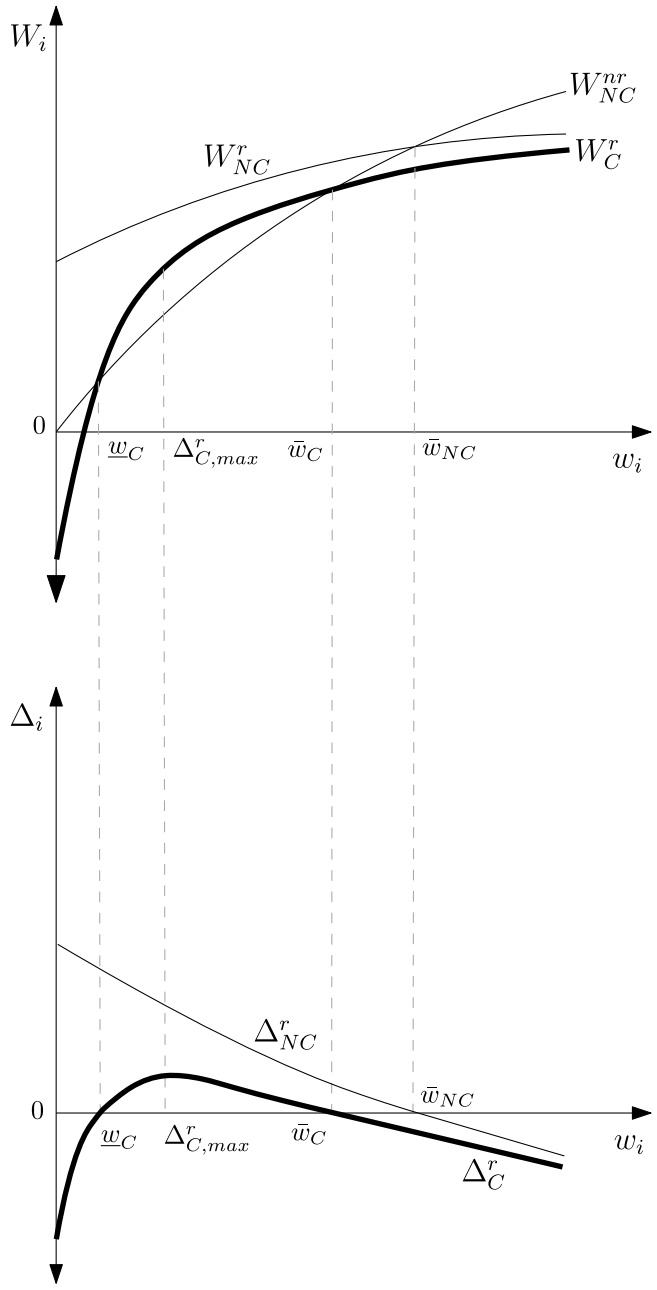

Fig. 6. Gains from return with and without mobility costs. Notes. The top panel shows lifetime payoffs (W) associated with return decisions (superscripts r for return and nr for not return) with and without mobility costs (subscripts C and NC, respectively), and the bottom panel shows the gains (Δ) from return migration. Accordingly, only refugees with initial period incomes between $\underline{w}$ and $\bar{w}$ find return incentive compatible.

payoff and return payoff streams is also possible. In that case, only refugees with an intermediate income range would return as Bijwaard and Wahba (2014) find for economic migrants.

We also observe similar suggestions in our vignette surveys. Table 6 showed that a former miner who is now working in construction with a work permit may be more likely to return to Syria than the analogous refugee who does not have a work permit. This may reflect the fact that getting a work permit in the country of asylum – this pattern was more evident in Jordan in particular – likely means higher wages and a relaxation of the credit constraint.

In Table 8 we show the results of the vignette which asks respondents to consider the returns decisions of a refugee who was a doctor in Syria and now in the country of asylum is either working as a janitor or

Table 8
Hypothetical return decisions: refugees with high skilled occupation.

	(1)		(2)	(3)	
Physician working as a physician	−0.055	*	−0.024	−0.090	*
	(0.032)		(0.046)	(0.046)	
Observations	938		465	473	
Mean: Physician working as a janitor	0.509		0.408	0.596	
SD	0.500		0.493	0.492	
Sample	All		Jordan	Lebanon	

Notes. See notes to Table 3 for details on specification and variable construction. Also included are indicators for the security scenarios. The sample is limited to respondents who were given the version of the vignette about a former physician in Syria.

has been given temporary authorization to treat Syrians in a hospital. We see that respondents think the physician who is able to practice medicine is less likely to return (5.5 percentage points; standard error of .032) than the physician working as a janitor. Without over-interpreting this finding, we would like to point out that the result is consistent with the framework above in that physicians not practicing their profession in exile may have greater gains from returning to Syria and that they may be less likely to be constrained in paying for the return trip home.

7. Conclusion

In this paper, we analyzed the factors that influenced the early, voluntary, and unassisted returns of Syrian refugees from Lebanon, Jordan, and Iraq to Syria during an active period of conflict, from January 2011 to March 2018. Our analysis with a novel monthly panel of returnees from districts in countries of asylum to sub-districts in Syria shows that better security and improved access to utilities, as proxied by nightlight luminosity, at the location of origin increase the likelihood of return. This result is aligned with the risk-adjusted expected payoff-based explanations of standard models of migration. However, we show that worse conditions in exile are not always associated with more returns to the origin country. In certain aspects (e.g., access to food), poor conditions in exile are negatively associated with the likelihood of return. In discussing these results, we propose a simple framework where an increase in income in exile can trigger return for those with low income in the presence of mobility costs.

Our study has both analytical and practical implications. From an analytical point of view, our findings complement an exhaustive legal literature on the conditions under which refugees may return, or stay in exile, by analyzing the conditions under which they may choose to do so. On the practical side, our results suggest that host country policies targeting the return of refugees by means of limitations on economic opportunities may be self-defeating. Overall, a formal consideration of the endogenous return behavior can help design humanitarian interventions more effectively in countries of asylum.

The study of refugee returns is a data-demanding process in a data-poor field. This limits the generation of systematic empirical evidence and, thus, our knowledge. Any future research that expands the respective data space, especially longitudinally, and examines the validity of our results in different refugee situations will help improve our understanding significantly.

Appendix

Vignettes of the household survey

A survey of 1900 Syrian refugees was conducted in Jordan and Lebanon in July and August 2018. The survey included three vignettes, which aimed to elicit the refugees' opinions on how likely a hypothetical refugee family would be to return to Syria. Each vignette explored a combination of push and pull factors, including employment opportunities in the country of asylum (differentiated by skill or household member), the availability and level of financial return assistance, and

the security situation, employment opportunities as well as condition of the family's home in Syria.

The complete set of vignettes as worded in the survey can be found below, which was translated into Arabic for the survey. The country was matched to the country of asylum where the respondent was residing at the time of the survey. Key elements of each vignette were also varied randomly across respondents to reduce the likelihood of introducing systematical biases into the findings.

Vignette 1: How does the ability to work in the country of asylum (differentiated by skill), the security situation in Syria and the availability of financial assistance affect the perceived likelihood of return?

Now, I am going to tell you the story of a family that lives in this region. Adnan and Sara have three children: Tarek (12), Mohammed (10) and Haya (7). These are not their real names. After fleeing from their small town in Syria, they have been living in [COUNTRY] for 3 years. In their home town in Syria, Adnan was a ... *randomized between:*

- physician at the local hospital. He is currently working as a janitor in the hospital in [COUNTRY] since he does not have work authorization to work as a physician.
- physician at the local hospital. Adnan has been able to get a temporary authorization to provide medical care to Syrians in a hospital in [COUNTRY].
- miner near his home town. He is currently working as a construction worker, which gives him irregular employment since he does not have a work permit.
- miner near his home town. He is currently working as a construction worker, which provides a low but steady income since he was able to procure a temporary work permit.]

The family has been monitoring the situation in their home town and heard that the security situation has dramatically improved in the last ... *randomized between*:

- six months.
- nine months.
- twelve months.

How likely is this family to return to Syria in the next 2 months?

- Very likely
- Likely
- Neither likely nor unlikely
- Unlikely
- Very unlikely

Now Adnan and Sara have heard that an international organization is offering ... *randomized between:*

- USD 1000 per family member or USD 5000 total (converted into local currency for Jordan and Lebanon as appropriate)
- USD 2100 per family member, USD 10,500 total (also converted to local currency)

if everyone in the family returns to Syria.
How likely is this family to return to Syria in the next 2 months?

- Very likely
- Likely
- Neither likely nor unlikely
- Unlikely
- Very unlikely

Vignette 2: How do employment opportunities for the husband and wife in the country of asylum and Syria affect the perceived likelihood of return?

I am now going to tell you a story about another family in this region. Fathi and Amena fled their village in Syria in 2017 and have been living in [COUNTRY] ever since. They have two children: Rasha (16) and Yara (9). These are not their real names. Fathi was a teacher in Syria but is unable to get a job teaching in [COUNTRY]. *The next sentenced was randomized between...*

- Amena was a housewife in Syria but now works as a housekeeper in [COUNTRY]. This earns a steady but small income.
- Amena is a housewife and takes care of Fathi and the children as they adjust to life and school in [COUNTRY]

The next sentence was randomized between:

- Fathi finds work where he can. Yara and Rasha are both attending secondary school.
- Fathi hears that schools in his village have re-opened and they are offering attractive salaries to teachers.
- Fathi hears that schools in his village still do not have the resources to pay teachers their full salaries.

How likely is this family to return to Syria in the next 2 months?

- Very likely
- Likely
- Neither likely nor unlikely
- Unlikely
- Very unlikely

If the entire family does not return together, how likely is that each of the family members returns to Syria in the next two months:

Fathi:	Rasha:
∘ Very likely	∘ Very likely
∘ Likely	∘ Likely
∘ Neither likely nor unlikely	∘ Neither likely nor unlikely
∘ Unlikely	∘ Unlikely
∘ Very unlikely	∘ Very unlikely
Amena:	Yara:
∘ Very likely	∘ Very likely
∘ Likely	∘ Likely
∘ Neither likely nor unlikely	∘ Neither likely nor unlikely
∘ Unlikely	∘ Unlikely
∘ Very unlikely	∘ Very unlikely

Vignette 3: How does the status of the family home in Syria affect the perceived likelihood of return and does it matter who provides the information?

I am now going to tell you a story about a third family in the region. Yousef and Lena got married 5 years ago and they have two children: Marwa (4) and Heba (2). These are not their real names. *The next sentenced was randomized between...*

- A resident of their village in Syria told Yousef that his house back home is intact and unoccupied.
- Yousef's two brothers and their family remain in their village in Syria and informed Yousef that his house back home is intact and unoccupied.
- A resident of Yousef's village told Yousef that his house back home has been destroyed.
- Yousef's two brothers and their family remain in their village in Syria and informed Yousef that his house back home has been destroyed.

How likely is this family to return to Syria in the next 2 months?

- Very likely
- Likely
- Neither likely nor unlikely
- Unlikely
- Very unlikely

Table A.1
Refugee characteristics by county of asylum: vignette survey data.

		Lebanon		Jordan	
			Observations		Observations
Panel A: All Household Members					
Male	Mean	0.516	3593	0.491	4448
	SD	0.500		0.500	
Age 0–14	Mean	0.319	3593	0.355	4448
	SD	0.466		0.478	
Age 15–44	Mean	0.569	3593	0.503	4448
	SD	0.495		0.500	
Age 45–59	Mean	0.081	3593	0.099	4448
	SD	0.272		0.299	
Age Over 60	Mean	0.032	3593	0.044	4448
	SD	0.175		0.204	
Education no/informal Age 25+	Mean	0.421	1567	0.182	1785
	SD	0.494		0.386	
Education primary Age 25+	Mean	0.467	1567	0.631	1785
	SD	0.499		0.483	
Education secondary/vocational Age 25+	Mean	0.083	1567	0.137	1785
	SD	0.276		0.344	
Education university Age 25+	Mean	0.029	1567	0.050	1785
	SD	0.167		0.218	
Family relationship core Age 15+	Mean	0.913	2444	0.952	2869
	SD	0.282		0.215	
Panel B: Respondents of vignette questions					
Male	Mean	0.759	950	0.750	947
	SD	0.428		0.433	
Age	Mean	35.384	950	41.955	947
	SD	11.532		12.930	
Marital status married	Mean	0.758	947	0.851	946
	SD	0.428		0.356	
Marital status single	Mean	0.167	947	0.047	946
	SD	0.373		0.211	
Marital status widowed	Mean	0.055	947	0.081	946
	SD	0.228		0.274	
Marital status other	Mean	0.020	947	0.021	946
	SD	0.140		0.144	

Notes. This data comes from the survey described in Section 3.4. To present characteristics of cases in the survey in a comparable way as in subsequent tables using the ProGres data, we define age in 5-year brackets. We report education levels only for adults aged 25 years or older. Individuals are grouped together as "cases" who are usually blood relatives and who are traveling and staying together while in asylum. We define individuals as "core family" if they are spouses and children, as well as parents and siblings. The omitted category is "extended family", such as in-laws and aunts.

Table A.2
Refugee characteristics upon entry by country of asylum and return status: ProGRES data.

		(1) Lebanon	(2)	(3)	(4)	(5) Jordan	(6)	(7)	(8)	(9) Iraq	(10)	(11)	(12)
		Return	No return	Observations	p-value	Return	No return	Observations	p-value	Return	No return	Observations	*p*-value
Male	Mean	0.406	0.477	1,214,243	0.000	0.447	0.492	623,091	0.000	0.573	0.526	166,210	0.000
	SD	0.491	0.499			0.497	0.500			0.495	0.499		
Age 0–14	Mean	0.134	0.480	1,214,243	0.000	0.306	0.492	623,091	0.000	0.290	0.427	166,210	0.000
	SD	0.340	0.500			0.461	0.500			0.454	0.495		
Age 15–44	Mean	0.546	0.442	1,214,243	0.000	0.519	0.415	623,091	0.000	0.589	0.494	166,210	0.000
	SD	0.498	0.497			0.500	0.493			0.492	0.500		
Age 45–59	Mean	0.139	0.052	1,214,243	0.000	0.103	0.053	623,091	0.000	0.076	0.050	166,210	0.000
	SD	0.346	0.221			0.304	0.224			0.266	0.218		
Age 60 or above	Mean	0.182	0.027	1,214,243	0.000	0.072	0.040	623,091	0.000	0.045	0.028	166,210	0.000
	SD	0.385	0.161			0.258	0.196			0.207	0.165		
Marital status married Age 15+	Mean	0.433	0.743	637,952	0.000	0.590	0.713	323,610	0.000	0.522	0.714	97,598	0.000
	SD	0.495	0.437			0.492	0.452			0.500	0.452		
Marital status single Age 15+	Mean	0.438	0.205	637,952	0.000	0.333	0.215	323,610	0.000	0.442	0.255	97,598	0.000
	SD	0.496	0.404			0.471	0.411			0.497	0.436		
Marital status widowed Age 15+	Mean	0.004	0.006	637,952	0.004	0.001	0.003	323,610	0.000	0.001	0.001	97,598	0.589
	SD	0.063	0.075			0.038	0.057			0.035	0.037		
Marital status other Age 15+	Mean	0.126	0.046	637,952	0.000	0.075	0.069	323,610	0.000	0.035	0.030	97,598	0.001
	SD	0.332	0.210			0.264	0.254			0.183	0.170		
Education no/informal Age 25+	Mean	0.400	0.182	468,623	0.000	0.203	0.128	244,804	0.000	0.257	0.192	68,385	0.000
	SD	0.490	0.386			0.402	0.335			0.437	0.394		
Education primary Age 25+	Mean	0.489	0.669	468,623	0.000	0.621	0.635	244,804	0.000	0.543	0.536	68,385	0.216
	SD	0.500	0.471			0.485	0.481			0.498	0.499		
Education secondary/vocational Age 25+	Mean	0.071	0.103	468,623	0.000	0.133	0.167	244,804	0.000	0.119	0.161	68,385	0.000
	SD	0.257	0.304			0.339	0.373			0.324	0.367		
Education university Age 25+	Mean	0.040	0.046	468,623	0.003	0.044	0.069	244,804	0.000	0.081	0.112	68,385	0.000
	SD	0.196	0.210			0.204	0.254			0.273	0.315		
Family relationship core Age 15+	Mean	0.898	0.991	637,952	0.000	0.994	0.999	323,610	0.000	0.983	0.996	97,598	0.000
	SD	0.302	0.096			0.078	0.032			0.128	0.062		

Notes. Data from UNHCR's ProGres database for Lebanon, Jordan, and Iraq. Since we do not know if a refugee receives an education while in exile, we restrict the sample for adults even further to 25 years or older to measure education levels upon entry. The database records age in 5-year brackets, which is updated over the sample period, and groups individuals together as "cases" who are usually blood relatives and who are traveling and staying together while in asylum. Thus, we can identify the relationship of individuals within each case relative to the principal applicant using a dummy variable for "core family", such as spouses and children, as well as parents, siblings, nephews and nieces (omitted category is "extended family", such as in-laws and aunts).

Table A.3
Refugees characteristics upon entry by conflict intensity: ProGRES data.

		(1) High conflict	(2) Low conflict	(3) Observations	(4) p-value
Male	Mean	0.486	0.484	2,003,544	0.047
	SD	0.500	0.500		
Age 0–14	Mean	0.468	0.473	2,003,544	0.000
	SD	0.499	0.499		
Age 15–44	Mean	0.437	0.444	2,003,544	0.000
	SD	0.496	0.497		
Age 45–59	Mean	0.058	0.052	2,003,544	0.000
	SD	0.234	0.222		
Age 60 or above	Mean	0.037	0.031	2,003,544	0.000
	SD	0.190	0.174		
Marital status married Age 15+	Mean	0.717	0.723	1,059,160	0.000
	SD	0.450	0.448		
Marital status single Age 15+	Mean	0.215	0.224	1,059,160	0.000
	SD	0.411	0.417		
Marital status widowed Age 15+	Mean	0.005	0.004	1,059,160	0.000
	SD	0.073	0.064		
Marital status other Age 15+	Mean	0.062	0.049	1,059,160	0.000
	SD	0.241	0.216		
Education no/informal Age 25+	Mean	0.157	0.178	781,812	0.027
	SD	0.364	0.382		
Education primary Age 25+	Mean	0.663	0.635	781,812	0.000
	SD	0.473	0.481		
Education secondary/vocational Age 25+	Mean	0.124	0.127	781,812	0.000
	SD	0.330	0.333		
Education university Age 25+	Mean	0.056	0.059	781,812	0.000
	SD	0.230	0.236		
Family relationship core Age 15+	Mean	0.992	0.992	1,059,160	0.000
	SD	0.092	0.086		

Notes. See notes to Table A.2 for details on test and variable construction. Districts that were in the top 10th percentile of the conflict events index (CEI) averaged over the time period January 2012-March 2018 are designated as "high conflict".

References

Alix-Garcia, J., Walker, S., Bartlett, A., Onder, H., Sanghi, A., 2018. Do refugee camps help or hurt hosts? The case of Kakuma, Kenya. J. Dev. Econ. 130, 66–83.

Bauer, T.K., Sinning, M.G., 2011. The savings behavior of temporary and permanent migrants in Germany. J. Popul. Econ. 24 (2), 421–449.

Becker, S.O., Ferrara, A., 2019. Consequences of forced migration: A survey of recent findings. Labour Econ. 59, 1–16.

Bell, B., Fasani, F., Machin, S., 2013. Crime and immigration: Evidence from large immigrant waves. Rev. Econ. Stat. 21 (3), 1278–1290.

Berry, M., Garcia-Blanco, I., Moore, K., 2015. Press Coverage of the Refugee and Migrant Crisis in the EU: A Content Analysis of Five European Countries. UNHCR, United Kingdom.

Bertrand, M., Mullainathan, S., 2001. Do people mean what they say? Implications for subjective survey data. Amer. Econ. Rev. 91 (2), 67–72.

Bijwaard, G.E., Wahba, J., 2014. Do high-income or low-income immigrants leave faster? J. Dev. Econ. 108, 54–68.

Bojicic-Dzelilovic, V., Turkmani, R., 2018. War economy, governance and security in Syria's opposition-controlled areas. Stab.: Int. J. Secur. Dev. 7 (1), 1–17.

Borjas, G.J., 1987. Self-selection and the earnings of immigrants.. Amer. Econ. Rev. 77 (4), 531–553.

Borjas, G.J., Bratsberg, B., 1996. Who leaves? The outmigration of the foreign-born. Rev. Econ. Stat. 78 (1), 165–176.

Borjas, G.J., Monras, J., 2017. The labour market consequences of refugee supply shocks. Econ. Policy 32 (91), 361–413.

Braun, S., Kvasnicka, M., 2014. Immigration and structural change: Evidence from post-war Germany. J. Int. Econ. 93 (2), 253–269.

Camarena, K.R., Hägerdal, N., 2020. When do displaced persons return? Postwar migration among Christians in Mount Lebanon. Am. J. Political Sci. 64 (2), 223–239.

Cortes, K.E., 2004. Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. Rev. Econ. Stat. 86 (2), 465–480.

Del Carpio, X.V., Wagner, M.C., 2015. The Impact of Syrian Refugees on the Turkish Labor Market. World Bank Policy Research Working Paper, (7402).

Dustmann, C., 1997. Differences in the labor market behavior between temporary and permanent migrant women. Labour Econ. 4 (1), 29–46.

Dustmann, C., Görlach, J.-S., 2016. The economics of temporary migrations. J. Econ. Lit. 54 (1), 98–136.

Dustmann, C., Vasiljeva, K., Piil Damm, A., 2019. Refugee migration and electoral outcomes. Rev. Econom. Stud. 86 (5), 2035–2091.

Galor, O., Stark, O., 1990. Migrants' savings, the probability of return migration and migrants' performance. Internat. Econom. Rev. 463–467.

Galor, O., Stark, O., 1991. The probability of return migration, migrants' work effort, and migrants' performance. J. Dev. Econ. 35 (2), 399–405.

Ghosn, F., Chu, T.S., Simon, M., Braithwaite, A., Frith, M., Jandali, J., 2021. The journey home: Violence, anchoring, and refugee decisions to return. Am. Political Sci. Rev. 1–17.

Gibson, J., McKenzie, D., 2011. The microeconomic determinants of emigration and return migration of the best and brightest: Evidence from the Pacific. J. Dev. Econ. 95 (1), 18–29.

Goodwin-Gill, G.S., McAdam, J., 2021. The Refugee in International Law. Oxford University Press.

Hangartner, D., Dinas, E., Marbach, M., Matakos, K., Xefteris, D., 2019. Does exposure to the refugee crisis make natives more hostile? Am. Political Sci. Rev. 113 (2), 442–455.

Harild, N.V., Christensen, A., Zetter, R.W., 2015. Sustainable refugee return: Triggers, constraints, and lessons on addressing the development challenges of forced displacement.

Moser, P., Voena, A., Waldinger, F., 2014. German Jewish émigrés and US invention. Amer. Econ. Rev. 104 (10), 3222–3255.

Ruiz, I., Vargas-Silva, C., 2015. The labor market impacts of forced migration. Amer. Econ. Rev. 105 (5), 581–586.

Schwartz, S., 2019. Home, again: Refugee return and post-conflict violence in burundi. Int. Secur. 44 (2), 110–145.

Sjaastad, L.A., 1962. The costs and returns of human migration. J. Polit. Econ. 70 (5, Part 2), 80–93.

Tumen, S., 2016. The economic impact of Syrian refugees on host countries: Quasi-experimental evidence from Turkey. Amer. Econ. Rev. 106 (5), 456–460.

Verme, P., Schuettler, K., 2021. The impact of forced displacement on host communities a review of the empirical literature in economics. Journal of Development Economics 102606.

Wooldridge, J.M., 1999. Distribution-free estimation of some nonlinear panel data models. J. Econometrics 90 (1), 77–97.

World Bank, 2017. The Toll of War: The Economic and Social Consequences of the Conflict in Syria. The World Bank Group.

World Bank, 2019. The Mobility of Displaced Syrians: An Economic and Social Analysis. The World Bank Group.

Yahya, M., Kassir, J., El-Hariri, K., 2018. Unheard Voices: What Syrian Refugees Need to Return Home. Carnegie Endowment for International Peace, Washington, DC.