The Dynamics of Refugee Return: Syrian Refugees and Their Migration Intentions *

Ala' Alrababa'h[1,2], Daniel Masterson[2,3], Marine Casalis[2], Dominik Hangartner[1,2,5], Jeremy Weinstein[2,4]

Word Count: 11,979

[1]Center for International and Comparative Studies, ETH Zurich

[2]Immigration Policy Lab, Stanford University and ETH Zurich

[3]Department of Political Science, University of California, Santa Barbara

[4]Department of Political Science, Stanford University

[5]Department of Government, London School of Economics and Political Science

February 14, 2022

*The authors would like to thank Claire Adida, Stathis Kalyvas, Kelsey Norman, Stephanie Schwartz, Alexandra Siegel and workshop participants at the American University in Beirut, APSA 2019 and 2020, CaliWEPS III, EGAP 2019, and MESA 2019. We thank Nasser Yassin and AUB's Issam Fares Institute for hosting a research design workshop and providing input on several aspects of the research. We are grateful to Oxfam and Amel for their support with several parts of the project and to CARE for sharing some of their data on the return intentions of Syrians in Jordan with us. We also thank staff at Basmeh & Zeitooneh, SMEX, IRC, DSP, and Human Rights Watch. Funding for this project was provided by the UK Foreign, Commonwealth & Development Office awarded through Innovation for Poverty Action's Peace & Recovery Program, the Swiss Network for International Studies, the Leverhulme Trust, a consortium of the Wellcome Trust, the Riksbankens Jubileumsfond, the Volkswagen Stiftung, and the Swiss Federal Department of Foreign Affairs. All analyses, unless otherwise noted, were pre-specified in EGAP registration 20190914AB. This project was reviewed and approved by the Stanford Human Subjects Committee under IRB protocol 49387 and by the Ethics Committee of ETH Zurich.

Abstract

In order to understand how refugee crises end, we must examine when and why refugees choose to return home. We study the drivers of refugees' decision-making using original observational and experimental data from a representative sample of 3,003 Syrian refugees in Lebanon. We find that conditions in a refugee's home country are the main drivers of return intentions. Refugees' decisions are influenced primarily by safety and security in their place of origin, their economic prospects upon return, and the availability of public services. Personal networks and confidence in information are also important. By contrast, the conditions in refugee-hosting countries—so-called "push" factors—play a much smaller role. Even in the face of hostility and poor living conditions, refugees are unlikely to return unless the situation at home improves significantly. We explore the generality of our findings using a second original survey of Syrian refugees in Jordan.

1 Introduction

Mass forced displacement has proven to be an enduring challenge in contemporary international politics. More than 26 million people live as refugees worldwide and the consequences of these persistent refugee crises are profound. Forcibly displaced people face joblessness and food insecurity, lack legal status, and experience hostility and violence in host countries (Khawaja, 2003; Onoma, 2013; Lehmann and Masterson, 2020). The governments of many hosting countries also struggle to meet the additional demands that refugees place on public services and infrastructure (The World Bank, 2017). The consequences of forced migration are particularly acute in developing countries, where more than 85% of refugees reside, because of constrained government budgets, weak state capacity, and limited public infrastructure (UNHCR, 2019*a*). As of 2018–2019, roughly 16 million people were in protracted refugee situations, accounting for approximately 78% of all refugees worldwide. The largest protracted refugee populations are Afghans, Syrians, and South Sudanese; the vast majority of these people reside in neighboring countries in the Middle East and North Africa and south Asia.[1] Despite the significant challenges that refugee crises pose to refugees themselves, hosting countries, and international donors, effective responses are lacking. Each year over the last decade, less than 1% of refugees worldwide received citizenship in a hosting country and only 1-2% were resettled (UNHCR, 2019*a*, pp. 28-33). The vast majority of refugees remain in a state of limbo, neither able to integrate locally nor find a new home through resettlement.

How then do refugee crises come to an end? This is a critical issue for politicians in hosting countries as well as policymakers in the humanitarian sector who must raise resources to support these populations. To answer this question, we need an understanding of whether, when, and why refugees return home. But this proves to be a challenging issue to explore empirically. Existing administrative data on refugee return is incomplete: in the past, many returns went unrecorded and the definition of return varied across organizations and across countries, making systematic analysis difficult. Moreover, data collection is especially challenging with mobile populations. The unpredictable timing of return means that it has been difficult to capture household return migration through surveys, especially

[1] UNHCR: https://tinyurl.com/2eyh6ee6. Accessed January 17, 2022. Note that these UNHCR statistics do not include the 5.7 Palestinian refugees who are under UNRWA's mandate. If Palestinians were included the statistics would shift to 21.7 million people in protracted refugee situations, accounting for 83% of refugees worldwide. See UNHCR: https://tinyurl.com/2p868pnj. Accessed January 26, 2022.

in contexts of ongoing violence, which compound challenges related to data collection and sample attrition.

We tackle the challenges of studying refugee return with original cross-sectional survey data from a nationally representative sample of 3,003 Syrian refugee households in Lebanon. We use this data to examine predictors of return intentions and preparations, to explore the role of information, and to identify differences in the drivers of short- and long-term return intentions. We supplement this analysis of observational data with a conjoint experiment in order to isolate the causal effect of conditions in Syria and Lebanon on return intentions. Finally, we explore the generality of our findings using a second original survey of Syrian refugees in Jordan.

The Syrian refugee crisis in Lebanon provides a useful setting in which to examine the dynamics of refugee return. When we launched our study in October 2019, active conflict in Syria was diminishing and many governmental and humanitarian organizations had begun discussing and even facilitating returns. Conditions across Syria varied widely—many areas remained insecure, and overall prospects for safety, economic recovery, and service provision were uncertain. At the same time, Syrian refugees in Lebanon experienced highly differentiated living conditions, local government policies, and levels of community hostility. In some municipalities, local governments actively targeted refugees for harsh treatment and prominent politicians called for accelerating their return, while in others refugees were integrating both economically and socially. We leverage this variation in prospects in the country of origin and well-being in the host country to learn about the drivers of return intentions.

Our data yield three important findings regarding the drivers of refugee return. First, there is strong evidence that pull factors play a more important role in shaping choices about return than push factors. Perceptions of individual-level safety in Syria are highly predictive of return intentions, as are economic conditions and the availability of public services. Personal networks in Syria also play an important role. By contrast, conditions in Lebanon do not significantly shape return intentions, even though some Syrians confront extremely challenging living situations. Second, the confidence that refugees have in their information about the situation at home is important for translating underlying preferences into actual plans to return. We find evidence that a number of drivers of return—regime control, economic prospects, access to services, and networks in Syria—are moderated by whether people have high confidence in information about conditions in Syria. Third, the results reaffirm the fundamental humanitarian nature of refugee crises. Despite having been

displaced for nearly a decade and having little hope to return in the near future, people who have fled the violence and societal devastation of civil war generally want to return home when those threats dissipate.

This article contributes to an emerging body of work on the drivers of return among displaced populations. Much of the existing work focuses on internally displaced persons (IDPs), who face the question of whether to return to their place of origin after flight within their country (Arias, Ibáñez and Querubin, 2014; Stefanovic, Loizides and Parsons, 2015; Camarena and Hagerdal, 2020; Weber and Hartman, 2022). A small but growing body of research has begun to shed light on the return choices of refugees, people who flee their home country during conflict and as a result face choices and constraints distinct from those that IDPs confront (Beber, Roessler and Scacco, 2021; Ghosn et al., 2021; Beaman, Onder and Onder, 2021). This is important as refugees generally have fewer prospects for long-term settlement in their place of refuge than IDPs do as citizens in their home country, and refugees face distinct legal, political, and economic challenges in a foreign country from those that IDPs face. We also contribute to this work by providing evidence on the role of a comprehensive set of theoretically motivated drivers of return decisions. Building on recent studies such as Ghosn et al. (2021), which focuses on refugees' wartime experiences and psychological anchoring, we extend the scope and examine a broad range of factors including material well-being and future prospects. Finally, our detailed measurement of return intentions and preparations allows us to study the interrelated roles of timing and aspirations, building on recent work that focuses on observed migration behavior (see, e..g, Beaman, Onder and Onder, 2021; Camarena and Hagerdal, 2020). In light of the many constraints that refugees face, evidence on return intentions is important because focusing solely on migration behavior does not allow us to distinguish if someone stays in exile because they were unable to return, preferred to not do so, or both (Schewel, 2020).

Our research is also relevant to the literature on host country politics. Traditionally, research on immigrant–native dynamics focuses on host populations in the receiving country, examining the effect of immigration and refugees on local labor markets (e.g., Scheve and Slaughter, 2001), political attitudes and behavior (e.g., Hainmueller and Hopkins, 2014), and tensions, discrimination, and conflict (e.g., Adida, 2014). We explore the possibility that these host country dynamics may in turn shape choices about return. Our results show that refugees may be willing to live with extreme hardship in the absence of a viable opportunity to return to their home country.

2 When Do Refugees Return?

In considering the question of how refugee crises end, a natural place to begin is the cessation of violence. Forced migration is often driven by violence (Davenport, Moore and Poe, 2003; Adhikari, 2013; Schon, 2019; Holland and Peters, 2020), and the end of hostilities in the home country may be a major driver of refugee return. In fact, when conflicts come to an end, refugees often find themselves pushed to return by aid agencies in neighboring countries that see their funding dry up and by governments that feel overstretched by hosting displaced persons (Crisp, 2019). At the same time, many governments in post-conflict countries actively seek the return of refugee populations as they work to catalyze an economic recovery (Jacobsen, 2005, p. 8)

How strong is the relationship between the end of conflict and refugee return? Figure 1 presents mobility patterns from four of the largest refugee crises in recent decades, focusing on Afghanistan, Iraq, Somalia, and Vietnam. The figure displays the number of refugees over time (from UNHCR data) and when each country was experiencing active conflict (from UCDP/PRIO Armed Conflict Dataset, defined as at least 25 battle-related deaths in a given year) (UNHCR, 2019*b*; Gleditsch et al., 2002; Pettersson, 2019). Declines in refugee numbers are largely attributable to refugee return, as only a small share of refugees are resettled and informal migration to third countries (e.g., in Europe) does not in itself remove a registered refugee from UNHCR's database (UNHCR, 2019*a*).[2]

The observed pattern is consistent with violence driving displacement and the end of active fighting spurring return. We see that the number of refugees often increases sharply when a conflict begins and tends to decrease in the years surrounding the end of conflict. However, we also see significant variation in when refugees return during both conflict and non-conflict years. The number of refugees sometimes increases and other times decreases during conflict. Moreover, the number of refugees often decreases very slowly after conflicts come to an end.

[2]UNHCR reports that in 2019 less than one percent of refugees were resettled to third countries (*UNHCR Resettlement*, 2021).

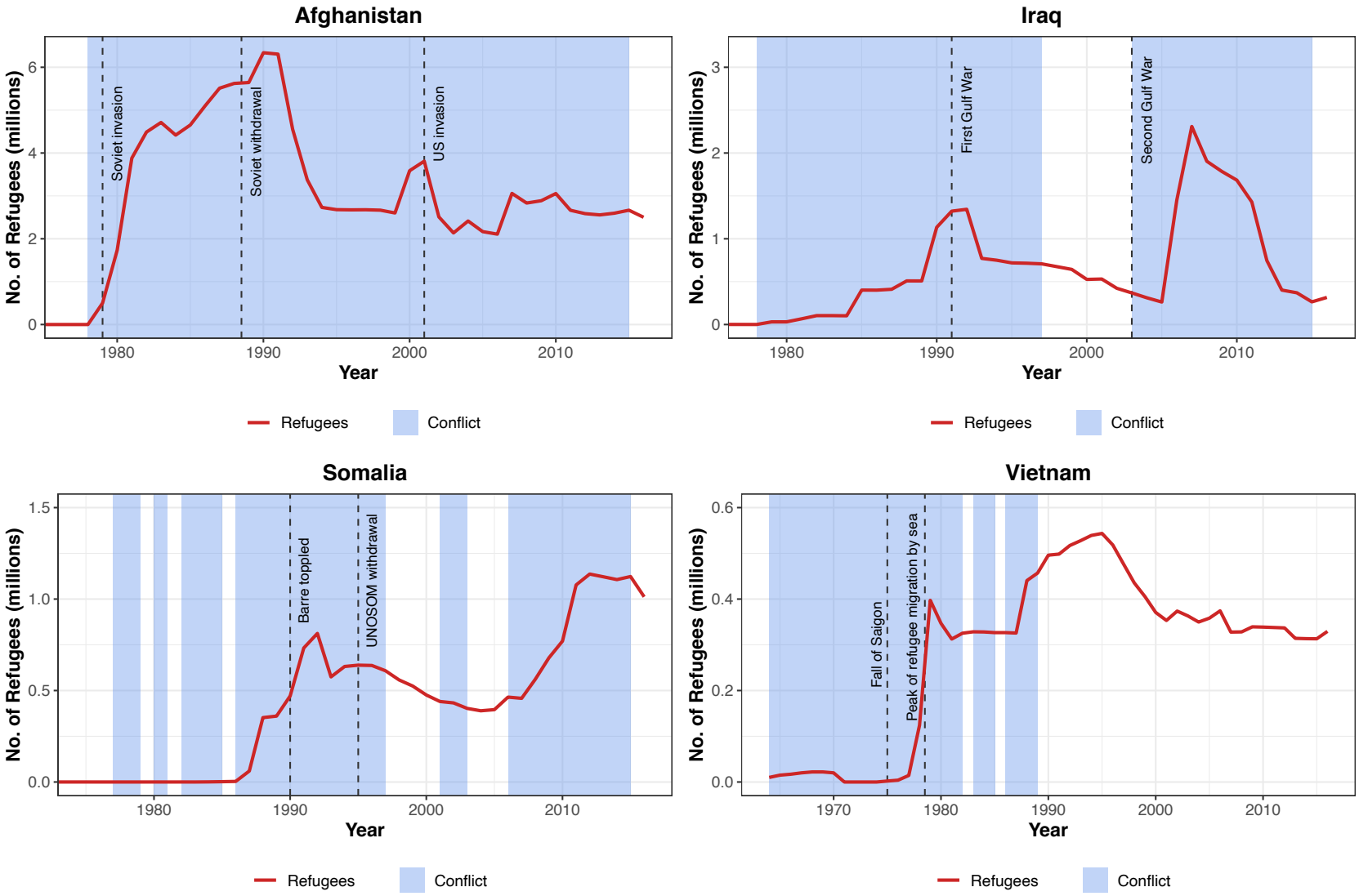

Figure 1: Conflict and Number of Refugees Over Time

These patterns suggest that macro-level measures of conflict may be insufficient on their own to explain refugee return. Given the fact that violence is rarely distributed uniformly across space or time, it is not surprising that refugee numbers wax and wane during conflict and do not automatically drop when a war notionally comes to an end. It may also be the case that a particular party to the conflict is victorious while another one is defeated, leaving some refugees concerned about their safety if they were to return. Further, constraints other than war and violence may be impediments to refugee return even if households desire to go home at the conflict's end. And it is possible that, after an extended time away from home, refugees have adjusted to a new context and wish to remain there, even in the absence of citizenship.

Given the variability of macro-level return depicted in Figure 1, we pivot to exploring the dynamics of return through the lens of household decision-making. We consider people's preferences, the environment in which they live, the context to which they might return, as well as other factors including the costs of moving and people's access to information. A focus on household decision-making enables us to consider the impact of macro-level changes in a home country or a host country, sub-national processes including localized violence and anti-refugee sentiment, and micro-level measures of household experiences, beliefs, and resources.

We begin by defining *return* as moving from a *host* country to one's *home* country with no immediate plans to depart again. Our focus is on the binary choice of whether to return to the home country, thereby setting aside other migration-related choices that refugees face such as internal migration within a host country, location choice within their home country after return, and formal or informal migration to a third country. We offer this definition with an awareness that during war and in its aftermath, the process of return may not be straightforward. Some people may return only to find that the situation in their home country necessitates migrating again in search of safety and a livelihood.[3]

In identifying the factors that influence return, our theoretical starting point is neoclassical economic theories of migration (Borjas, 1987). Though developed to explain patterns of labor migration, these models provide a useful framework for understanding individual decisions to migrate based on the costs and benefits of living in different countries. In this framework, potential migrants consider their long-term expected well-being in a home coun-

[3]Some people living as refugees may consider migration to third countries when deciding whether to return home or not. However, our fieldwork and survey data suggest that staying in Lebanon or returning to Syria are by far the predominant options that Syrian refugees in Lebanon consider. Thus, a two-country setup allows us to focus on the essential aspects of the decision to return or not.

try against conditions in possible destination countries, while accounting for the costs of travel, the challenges of adapting to a new labor market and culture, and any non-monetary costs or benefits of migration.

The new economics of migration extends these models to incorporate household considerations (Stark and Bloom, 1985). Recognizing that individuals often make decisions in coordination with other household members, this framework envisions migration as one strategy that households use to diversify and thereby minimize risk. These perspectives are complementary, as individuals may seek to maximize income while households aim to minimize risk (Massey et al., 1993). Together, they underscore the value of considering both individual- and household-level factors in modeling migration decisions (Borjas and Bratsberg, 1996; Constant and Massey, 2003).

Although one might question the value of rationalist models of migration in contexts of forced displacement, recent research suggests that these frameworks provide important insights even in environments where decision-making is influenced by violence. For example, one study of Lebanese Christians who were internally displaced during the country's civil war in the 1980s points out that, in the absence of attractive economic opportunities, people may not return to their home areas even if they have strong aspirations to do so (Camarena and Hagerdal, 2020). In the Colombian context, researchers found that IDPs were more likely to return home if they had access to land and meaningful economic opportunities. By contrast, the most vulnerable households—and those most directly affected by violence—exhibited the lowest willingness to return (Arias, Ibáñez and Querubin, 2014).

While this new work on IDPs suggests the relevance of rationalist models of migration, there are a number of unique challenges confronted by refugee households. The decision of refugees to return is often significantly more costly and difficult to reverse, as it involves crossing borders and giving up one's refugee status. At the same time, the feasibility of remaining in place is greater for IDPs who generally have greater access to the labor market and services in their new environments. By contrast, refugees often lack residency status making it difficult to find employment and provide health and education for their families.

Building on the existing literature, we focus on four factors that might shape refugees' choices about return: (a) push factors, or the situation in the host country (b) pull factors, or the dynamics in the country of origin, (c) the costs of mobility, and (d) the role that information plays in how households evaluate the costs and benefits of return.

Push factors: Existing research has highlighted the situation in the hosting country

as a primary determinant of people's choices about return migration (Dustmann and Weiss, 2007). In particular, previous work on labor migration and IDPs has found that economic and social factors in people's place of residence affect their choices about return. Constant and Massey (2003) find that a lack of stable full-time employment roughly doubles the odds of return migration for foreign workers in Germany. Arias, Ibáñez and Querubin (2014) find that IDPs in Colombia are roughly four percentage points more likely to intend to return if the household head is unemployed. Stefanovic, Loizides and Parsons (2015) find that integration into a new environment in western Turkey, measured by economic advancement and knowledge of Turkish, decreased return intentions among displaced Kurds from eastern Turkey.

Social networks may be a critical feature of people's migration decision during and after civil war. Civil war reconfigures societies, changing the roles of existing social networks while also creating new ones (Harpviken, 2009; Wood, 2008). Existing evidence validates that social networks play a key role in people's migration and return choices. Constant and Massey (2003) find that the range and nature of social attachments to Germany have a large negative impact on return migration. Masterson (2020) finds that Syrian refugees in Lebanon build and leverage network ties to access services and resources. Stefanovic and Loizides (2011) find that social capital among IDPs in Bosnia and Cyprus—manifested through refugee associations—was important in the coordination of mass returns, even in the face of resistance from opposition groups.

Historically, host governments often apply intense pressures for refugees to return en masse (Schwartz, 2019). Many refugee-hosting countries—even those that are initially receptive to refugee migration—gradually ramp up anti-refugee rhetoric and undercut refugees' legal residency and right to work. Often, the rationale behind such restrictions is that harsher living conditions will incentivize refugees to return home. We expect that a range of push factors in the host country will shape decision-making about whether to return including a household's economic situation, access to humanitarian aid, availability of public services, extent of social integration and acceptance, and legal status.

Pull factors: Household decision-making also depends on the environment in the home country. Refugees must consider the current conditions in their country of origin as well as their expectations about how the situation will evolve. For example, will violence pick up again and would it affect their particular region? Moreover, might they be at risk of targeted persecution or arrest upon return? The threats that people confront come not only

from armed conflict, but also from potential retribution. As households assess their safety if they were to return, they may consider current levels of violence in their hometown, their connections or proximity to existing political divisions, and expectations about continued violence and persecution by the government or armed groups.

Previous research documents a robust relationship between the intensity of conflict and forced migration (e.g., Fearon and Shaver, 2020), as well as people's exposure to violence and their decisions about displacement (e.g., Steele, 2009; Adhikari, 2013). People who have been forcibly displaced may be less likely to return when they face greater risk of violence in the place of origin. This is consistent with Arias, Ibáñez and Querubin (2014), who find that IDPs in Colombia who directly suffered violence before displacement have lower return intentions than IDPs who did not experience violence.

Expectations of safety must include strategic considerations about group identity. Armed groups often displace civilians strategically based on ethnic group, sect, tribe, or political affiliation (Steele, 2009; Balcells and Steele, 2016; Lichtenheld, 2020). If a government or armed group that engaged in collective targeting retains territorial control, refugees may fear that such communal punishment could be used again in the future, particularly if postwar political divisions map onto past migration choices (see, e.g., Schwartz, 2019). Past experiences with violence and forced displacement may shape people's political attitudes (see, e.g., Fabbe, Hazlett and Sinmazdemir, 2019), which may in turn affect the conditions under which they would return.

As people consider return, they need to evaluate the conditions of the post-war economy and whether they believe they will be able to meet their family's basic needs. Arias, Ibáñez and Querubin (2014) find that people who own land or have prospects for employment in their place of origin have higher return intentions to return. Beber, Roessler and Scacco (2021) find that the South Sudanese residents of North Sudan who were least likely to return were the middle class for whom employment opportunities were scarce in the South. War may also impact people's economic prospects by generating disputes over public policy or property rights (Schwartz, 2019). In one example of such dynamics, Weber and Hartman (2022) find that displaced Iraqis were more likely to return if they had property at the place of origin only when they possessed written documentation of their rights and the property was not disputed or badly damaged. Such concerns may have broad political salience if governments claim and redistribute property for the sake of demographic engineering or rewarding allies (McNamee, 2018).

Lastly, networks in the home country may help people survive if they were to return. Refugees with more friends and family in the home country—whether they are returned refugees themselves or simply never fled the country—may be able to rely on those people for support and community if they were to return. In line with this expectation, Arias, Ibáñez and Querubin (2014) find that social networks in people's place of origin (measured through membership in a peasant organization or collective land ownership) increase intentions to return.

Mobility costs: Households considering migration must weigh the financial costs and physical risks associated with moving (Hunt and Mueller, 2004). Long-distance travel for refugees, in particular, may be expensive and require passage through unsafe territory. Return migrants might also face the prospect of being stopped at military checkpoints run by armed groups that charge tolls or taxes, steal possessions, or detain, interrogate, or abuse travelers (e.g., Stork and Abrahams, 2004). Given these concerns, households facing higher mobility costs may be less likely to return. Existing work presents mixed evidence about the impact of a country's travel infrastructure and geography on refugee migration. While Schmeidl (1997) and Moore and Shellman (2006) find no evidence that distance and terrain affect refugee migration, Adhikari (2013) shows that the presence of accessible roads increases the probability of displacement.

Information: Finally, decisions about whether to return are influenced by a fourth factor: information. As social media, smart phones, and internet connections are now widely available, this factor arguably deserves greater attention than it has received in earlier work. After months or years away from home, people need to seek out and piece together information—often incomplete, often contradictory—in order to form expectations about what life would be like if they were to return (Batista and Cestari, 2016). The confidence that people have in their beliefs about the quality of life back home is likely to shape how they evaluate the costs and benefits of return.

Existing evidence shows that networks facilitate refugee migration by disseminating information about travel routes and destinations (Davenport, Moore and Poe, 2003; Moore and Shellman, 2004, 2007; Schmeidl, 1997; Schon, 2018). Further, violence and poverty motivate people to acquire information about conditions and policies in potential migration destinations (Holland and Peters, 2020). Recent work provides evidence on the role of information in refugee return. Ghosn et al. (2021) find that Syrian refugees who have prior experience of violence in Syria are more willing to leave Lebanon and return home, which the authors

argue is due to a greater ability to understand and assess their risk tolerance for violence.

3 The Syrian Refugee Crisis in Lebanon

To test the drivers of refugee return intentions, we focus on the Syrian refugee crisis in Lebanon. Given the diversity of localities in which Syrians have settled and the heterogeneity in conditions in Syria, this is a helpful case for examining the role of push and pull factors, mobility costs, and information in shaping return intentions. Lebanon, in particular, provides a critical test of the importance of push factors, given the documented hostility, discrimination, and violence that many Syrians have faced in Lebanon (Lehmann and Masterson, 2020). In addition, the context provides meaningful variation in prospects in Syria, mobility costs, and access to information. Syrians in Lebanon vary widely in their characteristics and backgrounds, originating from all of Syria's regions and spanning the country's pre-war socioeconomic spectrum.

What began in Syria in 2011 with street demonstrations and calls for political reforms collapsed into a devastating civil war, which caused an enormous refugee crisis, with millions of people fleeing to Lebanon, Jordan, Turkey, Iraq, Egypt and beyond. As of late 2019, when our study was conducted, more than five million Syrians had fled to neighboring countries and more than six million were displaced inside Syria. Approximately 930,000 Syrians lived in Lebanon, alongside 4.5 million native residents, in a small country with a land area three-quarters the size of the US state of Connecticut.[4]

One driving assumption behind Lebanon's national policy agenda for Syrian refugees is that exploitation, vulnerability, and material hardship will force Syrians to leave the country (Janmyr, 2016). Syrians in Lebanon face widespread hostility, confront significant restrictions on the right to work, and have only limited legal status in the host country. Most Syrians in Lebanon lack reliable access to education, healthcare, stable housing, and safe transportation (see, e.g., Mourad, 2017; Lehmann and Masterson, 2020). They live primarily in urban and peri-urban settings, with 15% in camps, informally managed by NGOs, as the UN did not establish official refugee camps in the country. The situation of Syrians in Lebanon is similar in many respects to the hardship that refugees worldwide

[4]We use UN registration numbers, which provide a conservative estimate of displacement. Refugee population: UNHCR Operational Portal. https://data2.unhcr.org/en/situations/syria. IDP data: UNHCR Refugee Data Finder. https://www.unhcr.org/refugee-statistics/. And IDMC. https://www.internal-displacement.org/countries/syria. Accessed November 15, 2019.

face; notably, many governments restrict refugee rights in order to accelerate return and, worldwide, less than one third of the world's 26 million refugees live in camps.

As the Syrian government regains control of much of the country, tens of thousands of Syrians have begun returning home, even as violence continues to displace more people. State and non-state actors in Lebanon have begun taking steps to facilitate and push for the return of refugees, tensions between Lebanese and Syrians remain high, and discussions about the return of refugees are increasingly prominent in journalistic and policy circles. Looking to Syria, the war has devastated the country's infrastructure and public services, including water supply, electricity, schools, and healthcare. Many people fear the persecution and violence that may result from government retribution and collective punishment in the postwar period. People who escaped Syria during the conflict may be especially prone to retaliation by the regime upon return. Men aged 18-42 are subject to military conscription in Syria, and serving in the Syrian military is likely to put conscripts in dangerous situations for years to come, where they may have to kill innocent people or be killed in combat or insurgent attacks. Even if the ultimate victor in the war is no longer in question, the specter of future violence remains.

4 Research Design and Data

4.1 Survey Design

We use original survey data from interviews with a nationally representative sample of 3,003 Syrian refugee households living in Lebanon.[5] The survey measured a wide range of household characteristics, predictors of return, and migration intentions, and also included a conjoint experiment to identify drivers of return intentions. The research team contracted a Lebanese survey firm to conduct data collection, and participated in all stages of research including enumerator training, survey piloting, and oversight of data collection. Data collection for the survey took place from August to October 2019.

To obtain a representative sample of Syrian households in Lebanon, we used stratified random sampling to ensure variation in Syrian and Lebanese demographics in localities and households sampled. A household head (either gender) served as survey respondent.

[5] Conducting research with Syrian refugees in Lebanon requires particular attention to the sensitive situation in which they live. See Appendix Section 8 for a full discussion of our study's ethical considerations and precautions.

Appendix Section 1 provides a detailed discussion of sampling protocols.

4.2 Measuring Return Intentions

Measuring return intentions is challenging, and survey instruments must account for the different time horizons across which households consider decisions in addition to the uncertainty that people face. Capturing intentions is also difficult in the absence of concrete behaviors consistent with stated intentions. As a result, we also measure preparations to return, as a self-reported but behavioral manifestation of return intentions. We asked respondents about their return intentions in three ways:

- Return intentions: "Do you plan to return to Syria in the next 12 months?"
- Return preparations: We asked a battery of binary questions about legal, financial, and logistical steps to prepare for return. Metrics of preparation include saving resources, collecting paperwork such as birth certificates or marriage documents, reaching out to Lebanese authorities and UNHCR, and taking scoping trips. We use these questions to calculate a preparations index with polychoric principal component analysis (PCA).
- Long-term return intentions: "Do you hope to move back to Syria and live there one day?"

The return intentions measures are binary variables, coded as 1 if the head of household plans to return within the specified time frame and 0 otherwise.[6] It is worth noting that our key outcomes are stated intentions and self-reported preparations to return, not a retrospective measure of return behavior. Such forward-looking outcomes are an important quantity of interest as people consider whether to return and policymakers design and implement programs to address refugee situations. A foundational principle of return policy is ensuring its *voluntary* nature, which requires placing people's *intentions* to return at the center of planning (see, e.g., Mixed Migration Centre, 2019, p. 93).

4.3 Measuring Drivers of Return

We measure four key concepts that we hypothesize will drive return decisions: (1) well-being in Lebanon, (2) prospective well-being in Syria, (3) information, and (4) mobility costs. To

[6] Robustness tests using alternative coding for the outcome are included in Appendix Section 6.4.

measure concepts 1-3, we draw on data from multiple survey questions and use PCA to construct indices to capture aspects of respondents' living situation in Lebanon, prospects in Syria, and access to information. We present the full set of PCA inputs in Appendix Sections 2 and 3.[7] In both Syria and Lebanon, we measure economic well-being, using data on assets and earning potential in each country, and current employment, earnings, and aid in Lebanon. We also examine the availability of services, including education, healthcare, water, and electricity, in Lebanon and Syria. We analyze the size of social networks and the number of friends and family in Lebanon and Syria. We examine people's ability to move freely and safely around Lebanon, and their integration in the country using the measures from the IPL-12 integration scale (Harder et al., 2018). To construct an index for the security situation in Syria, we focus on both general factors, such as whether there is still fighting, and personal factors, such as whether a family has any draft-aged men and whether the respondent personally experienced violence. The index on safety also includes an input about safety expectations in one year. The economic conditions in Syria and the services in Syria indices also include inputs that measure expectations about the future situation in Syria. We also construct an index for regime control, including detailed questions on which parties currently and formerly controlled a respondent's hometown.[8] The index for information includes whether the respondent speaks regularly with family or friends in Syria about the situation as well as measures of people's confidence in the information they have about safety, jobs, services, and conscription in their hometown.

In addition to measuring people's confidence in information directly, we ask questions about the size of refugees' networks in the host and the home country. Family and friends may serve as important sources of information about the conditions in one's hometown. Networks of family and friends may also directly impact people's return choices independent of the information they provide, in the sense that many people want to live in the same place as others in their close network.

We study mobility costs using two metrics: travel distance to one's hometown and household size. We calculate travel distance from each survey respondent's town of residence in Lebanon to their hometown in Syria, via the Beirut–Damascus highway and border crossing, using the Google Maps API. Our fieldwork revealed that this was the only legal border

[7] Although PCA inputs were pre-specified, some survey questions were listed in the PAP for inclusion in two indices. We departed from the PAP in these cases in order to maintain mutually exclusive index inputs. Appendix Section 5 documents these changes.

[8] We deviated from the PAP to separately study the role of regime control and safety conditions in Syria.

crossing open at the time of research and that a majority of Syrians moving back travel via official routes.

One potential concern with our survey is that affective biases may shape both people's reports of their situation and their return intentions, possibly leading to spurious correlations. To mitigate this threat, our metrics of well-being in Lebanon and prospects in Syria aim to measure objective facts rather than people's affect about their potential future in Lebanon and Syria. For instance, we ask respondents questions including "Did you work outside the home for money in the past 30 days?" and "Do you receive humanitarian aid through the UN cash assistance program?" (rather than about affect such as "How would you rate the quality of your life in Lebanon?").

4.4 Conjoint Experiment

We also present a conjoint analysis that experimentally manipulates potential drivers of return intentions.[9] This allows us to isolate the effects of conditions in Lebanon and Syria, individual circumstances, and social networks in shaping respondents' thinking about return. In the conjoint, the enumerator informed respondents: "I will now present you with five conditional scenarios. Please listen to these scenarios carefully and answer the questions about them." Respondents were then read a sequence of five separate vignettes, and after each one, they were asked the following question: "Under these conditions, would you return to Syria?"

In the vignettes, each of the numbered attributes below was randomly given one of the lettered values, and the order of the attributes was randomized across respondents. The vignettes were presented as follows:

> *Imagine that one year from now, regarding the security situation in Syria, [INSERT FROM (1) BELOW]. It appears that in [INSERT HOMETOWN], [INSERT FROM (2)]. As for conscription, [INSERT FROM (3)]. In Lebanon, [INSERT FROM (4)]. Finally, regarding your friends and relatives, are [INSERT FROM (5)].*

1. **Safety in Syria:** (a) Your hometown is quite safe; (b) Your hometown remains insecure; (c) All of Syria is quite safe

[9] See also, Ghosn et al. (2021) and Beaman, Onder and Onder (2021).

2. **Economic conditions in Syria:** (a) There are many job opportunities; (b) Public services, such as health centers and schools, are relatively easy to attain; (c) There are few job opportunities; (d) Public services, such as health centers and schools, are difficult to attain

3. **Personal safety:** (a) Military conscription has stopped; (b) Military conscription is still in place

4. **Conditions in Lebanon:** (a) You have a good job in Lebanon; (b) You do not have a good job in Lebanon; (c) Health centers and schools in Lebanon are available and affordable; (d) Health centers and schools in Lebanon are unavailable and unaffordable

5. **Network effects:** (a) Most of your friends and relatives are in Lebanon; (b) Most of your friends and relatives are in Syria; (c) Most of your friends and relatives are in Jordan, Turkey, and Iraq

5 Results: Observational Data on Return Intentions

We begin by describing our sample. Around 50% of our 3,003 respondents reside in urban areas in Lebanon and 33% live in informal settlements. The median year of arrival for respondents was 2013. The majority, 80%, are registered or recorded with UNHCR. In terms of education levels, 49% had an education level less than completing primary school, 39% completed primary school, and 12% had a secondary education or higher. As for aid, 48% of respondents received cash transfers, 62% received food vouchers, and 32% received both. Discrimination toward refugees in Lebanon is quite high but far from universal. 37% of respondents reported living in towns that had curfews in the past two years (which usually target refugees) and 40% reported facing discrimination when searching for houses. Finally, when it comes to conditions in Syria, 67% of respondents reported that protests occurred in their hometown during the revolution and 96% said that there was heavy fighting in their hometown at some point during the war. By the time the survey was conducted, 66% of respondents said that their hometowns were controlled by the government.

We examine the distribution of return intentions in Figure 2. We find that return intentions are increasing with the time horizon. Only 5% of Syrians plan to return in the next 12 months, that is, before approximately September 2020, and about a quarter of Syrians antic-

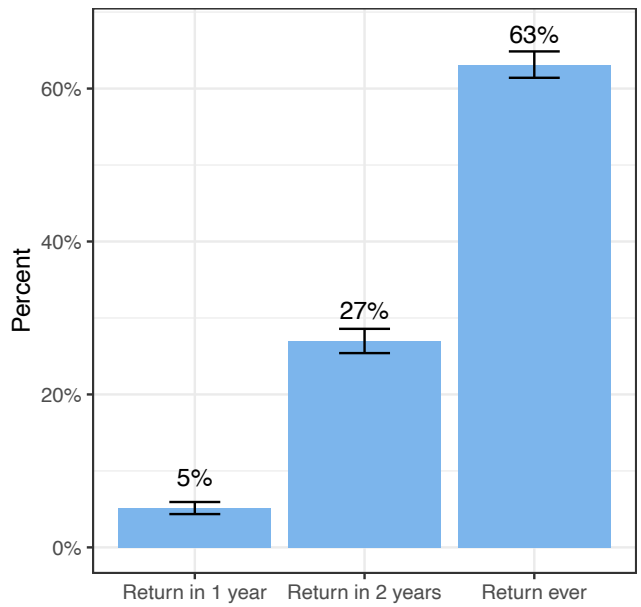

Figure 2: Return Intentions (Short, medium, and long-term)

ipate returning before September 2021. 63% plan to return at some point in the future.[10] To put these numbers in context, the median year of arrival for respondents was 2013, meaning that the median respondent had been displaced for more than six years at the time of data collection.

To study how cross-sectional differences shape return intentions, we examine the predictive power of a range of potential drivers of refugee return described above. We estimate the following regression model:

$$Y_i = \alpha + \beta T_i + \gamma X_i + \epsilon_i \ , \tag{1}$$

for each outcome Y and a vector of indices T. Each index is the first principal component from a PCA analysis of the measures detailed in Section 4.3. We also adjust for a range of control variables, X, including household-level covariates and locality-level fixed effects.[11] In regressions including travel distance on the right-hand side, we drop controls for location

[10] We impute missing values in our data using multivariate imputation by chained equations, discussed in Appendix Section 3.

[11] As defined in the pre-analysis plan, the control set includes the following: indicators for being from an urban area in Syria, living in an informal (tented) settlement, a household member being seriously ill, the head of household having finished secondary school or higher, a toddler household member, an elderly

in Lebanon and hometown in Syria, since travel distance is a deterministic function of these variables. Finally, ϵ is a mean-zero error term. We also run a series of regression models similar to Equation 1, but where the vector of indices T is replaced with each respective index in one model.[12]

5.1 Drivers of Return Intentions

We present results for the drivers of return intentions in Figure 3. Each dot represents the point estimate for the relationship between a given index, labeled on the y axis, and a metric of return, labeled at the top of each panel. Circles represent point estimates drawn from our main model in Equation 1, and triangles represent point estimates drawn from models with each respective index in one model. The independent variables are grouped into four categories: people's prospective situation in Syria, people's living situation in Lebanon, mobility costs to return to Syria, and people's confidence in the information they possess about Syria. The horizontal line around each point estimate shows the 90% and 95% confidence intervals (dark and light, respectively). Standard errors are clustered at the locality level, following from the sampling strategy. Indices are normalized to have mean zero and standard deviation one, and the point estimates present the change in the probability of return intentions that corresponds to a one standard deviation shift in an index. As shown in the Appendix Section 6.5, results are robust to using additive indices rather than PCA indices and using alternative control sets.

Figure 3 provides strong evidence for a relationship between conditions in Syria and intentions to return within 12 months (first panel). We see that safety in Syria, economic prospects in Syria, the availability of public services in one's hometown, and respondents' family and friend networks in Syria are positively and significantly associated with return. For each of these indices, we see that a one standard deviation shift in the index corresponds with about a 2 percentage point increase in return intentions. In light of the small fraction of refugees (only 5%) who plan to return in the next year, this constitutes a large increase

household member, female-headed household, whether an area is a Hizbullah-controlled area, and Syrian governorate of origin and Lebanese district of residence. All variables come from the survey data except the final covariate. Data for Hizbullah control is coded based on whether a Hizbullah-aligned candidate won a seat in the 2018 parliamentary elections, using Lebanese Ministry of Interior data released by the Data Liberation Project.

[12] Figure 3 involves two deviations from the PAP due to multicollinearity, discussed in detail in Appendix Section 5.

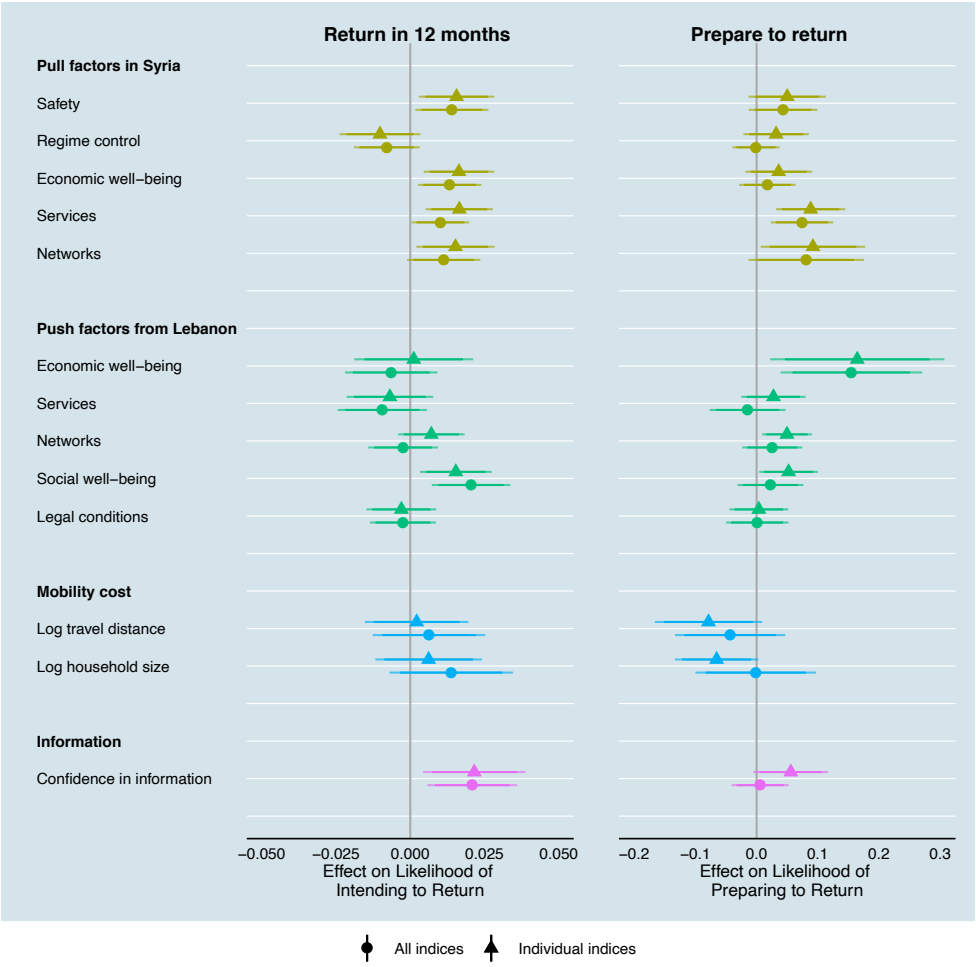

Figure 3: Index Results—Return Intentions and Preparations

in return intentions in percentage terms (roughly 40%). Control by the Syrian government correlates negatively with intentions to return, although we cannot rule out a null relationship at either the 90% or 95% level.

The relationship between conditions in Syria and preparations for return (second panel) is less clear but points in the same direction. Point estimates are consistently positive, but only the availability of services and the size of social networks are statistically significant. Security in Syria and economic prospects predict preparations to return but the results are not statistically significant. Regime control has no detectable relationship with preparations to return.

The results on push factors in Lebanon are quite different. First, looking at the left panel, we do not find a clear correlation between well-being in Lebanon and return intentions. We

cannot rule out a zero association for most of the indices. The one index that demonstrates a statistically significant association with return intentions is social well-being. In contrast to the lack of evidence for a role of push factors in shaping return *intentions*, the second panel reveals evidence for an association between conditions in Lebanon and return *preparations*. We find that higher levels of economic well-being, networks, and social well-being in Lebanon exhibit a detectable *positive* correlation with having taken steps to prepare to return to Syria in at least one specification. The direction of the relationship is not what we expected ex ante, based on a theory of preparations being driven by a simple utility comparison between conditions in Lebanon and prospects in Syria. The finding highlights that the theory's focus on migration costs and incentives may have overlooked migration capacities. Indeed, return is a complex and daunting process, and people with more financial and social resources may be better able to undertake a safe voluntary return.

Looking at the next group of drivers, we see in the first panel that the results do not provide evidence of a relationship between mobility costs and return intentions. In the second panel, we find a negative association between mobility costs and preparations for return, significant at the 10% level, when we consider indices separately. Looking at the bottom row of Figure 3, we see that confidence in information about one's hometown is positively associated with both intentions and preparations. Information access may both have a direct effect on return intentions and play a moderating role. We examine this possibility in the next section, testing whether conditions in Syria have a larger effect on people's intentions when they have high levels of confidence in their information about the situation in Syria.

We conduct two additional tests to probe whether pull factors are more powerful predictors than push factors in shaping return intentions. First, we fit predictive models based on push factors and pull factors using 10-fold cross validation. We present the results in Appendix Section 6.7, and find that models based on pull factors consistently demonstrate higher predictive power than models based on push factors.

Second, we test whether Syrians' conditions in Lebanon exhibit sufficient variation at both the national and local levels. If Syrians' conditions in the country were fairly homogeneous, then a null relationship between push factors in Lebanon and return intentions would be substantively trivial. Reassuringly, the data are not consistent with this concern. In Appendix Section 3.1, the descriptive statistics demonstrate wide variation in the living conditions of Syrians in Lebanon. In Appendix Section 6.3, we re-run all models that controlled for Lebanese locality level fixed effects, but without adjusting for locality. Our findings are

robust to this alternative specification, suggesting that our null findings for the role of push factors in Lebanon are not driven by a lack of variation in living conditions within localities. In Appendix Section 6.8, we examine this more systematically by producing a map of Lebanon that shows the variation in respondents' conditions by district using an index of all the components used to measure push factors in Lebanon. The map shows significant variation across districts—the difference in the push factor index between the district with the, on average, worst and the best conditions for respondents is about 2.4 standard deviations.

5.2 Information and Return Intentions

To further explore the relationship between information and return, we examine whether information moderates the role of perceived conditions in Syria. Specifically, we examine whether conditions in Syria have a larger effect on people's intentions when they have high levels of confidence in their information about the situation in Syria.

$$Y_i = \alpha + \beta_1 T_i + \beta_2 \left(T_i \times \mathbb{1}(I_i > 0)\right) + \gamma X_i + \epsilon_i \tag{2}$$

Equation 2 is similar to the "individual indices" specification of Equation 1, but includes a multiplicative interaction term between each index T and confidence in information. The indicator function, $\mathbb{1}(I_i > 0)$, denotes whether a respondent i had an index value for information confidence above the mean. X denotes the same vector of covariates as in Equation 1. Figure 4 presents regression results, displaying the estimated marginal effect of a one standard deviation change in each index for people with low (below-average) confidence in information compared to those with high (above-average) confidence in information.

The results in Figure 4 suggest that the relationship between conditions in Syria and return intentions and preparations is shaped by respondents' confidence in their information sources for some key factors. Specifically, we find evidence that information confidence is a significant moderator for the effects of regime control, economic prospect, and availability of services in people's hometown on intentions to return within 12 months. Information confidence also moderates the relationship between availability of services and social networks, respectively, with preparations to return.

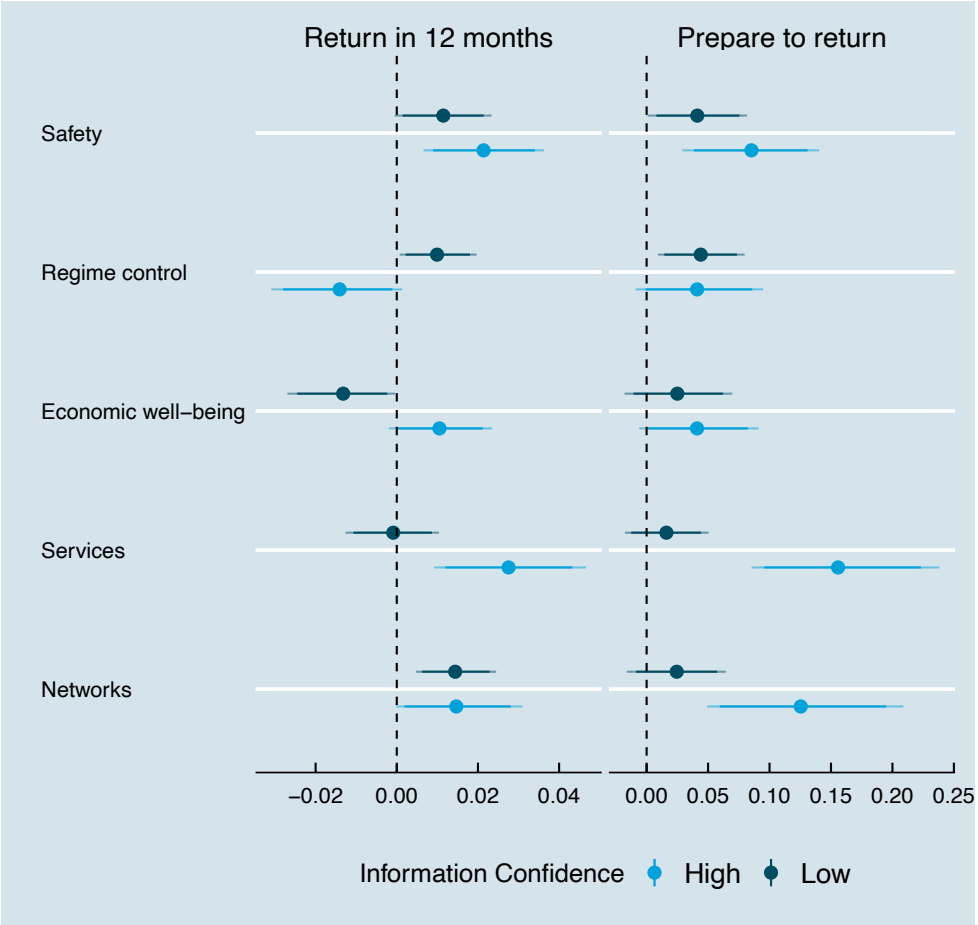

Figure 4: Interactive Effects of Information Confidence on Intentions and Preparations. Estimated marginal effects are presented with the corresponding 95% confidence intervals, with standard errors clustered by locality in Lebanon.

6 Results: Conjoint Experiment

The analysis of observational data strongly suggests that pull factors are more predictive of return intentions than push factors. Yet, despite our extensive set of control variables, our correlational estimates might be affected by other factors not included in the model. In this section, we present the results of a conjoint experiment designed to provide greater leverage on the causal effects of these drivers on return intentions. We follow a standard approach for analyzing conjoint experiments, using OLS regressions to estimate the Average Marginal Component Effect (AMCE) for each attribute (Hainmueller, Hopkins and Yamamoto, 2014). Figure 4 displays the effects for respondents' answers to the question: "Under these conditions, would you be willing to return to Syria?"

The main findings from the conjoint experiment are consistent with our analysis of the observational data. Conditions in Syria play a more important role in shaping people's return intentions than conditions in Lebanon. Results suggest that safety is the most powerful driver of return, with security in one's hometown increasing return intentions by 35 percentage points and nationwide security increasing return intentions by 42 percentage points. The fact that safety in one's hometown has nearly as large of an effect as nationwide safety, suggests that the majority of variation in people's consideration of security is driven by conditions in their hometown, highlighting the local nature of security concerns in postwar environments. The availability of jobs and public services in Syria increases return intentions by 8 percentage points. An end to military conscription also plays an important role in shaping people's return intentions, and increases the likelihood of return by 18 percentage points.

Both access to a good job and public services in Lebanon play a small but statistically significant role in people's return intentions. Someone with a good job in Lebanon is 2 percentage points less likely to return, and if someone has access to public services they are 3 percentage points less likely to return. Despite the statistical significance of these results, the differences in magnitudes between push and pull factors is substantial.

In the bottom of the figure, we see the effect of networks on people's responses. People were nearly 5 percentage points more likely to say that they would return to Syria if they have family and friends there (compared to having people outside of Syria and Lebanon). In contrast, we see a precisely estimated null effect for having family and friends in Lebanon on people's return intentions. These network results align with our earlier findings about the relative importance of the conditions in the home country compared to the hosting country.

7 Beyond Lebanon: Return Intentions in Jordan

Given the magnitude of the Syrian refugee crises, Syrians migrated to numerous countries, including to three primary hosting countries: Lebanon, Jordan, and Turkey. In order to ascertain whether our results are driven by unique circumstances among Syrians in Lebanon, we ran a separate survey with 1,286 Syrian refugees living in Jordan. These data offer a test of the external validity of our findings to the broader population of Syrian refugees. Our sampling strategy selected individuals from the four metropolitan areas in Jordan with the largest refugee populations: Amman, Irbid, Mafraq and Zarqa (including Azraq town).

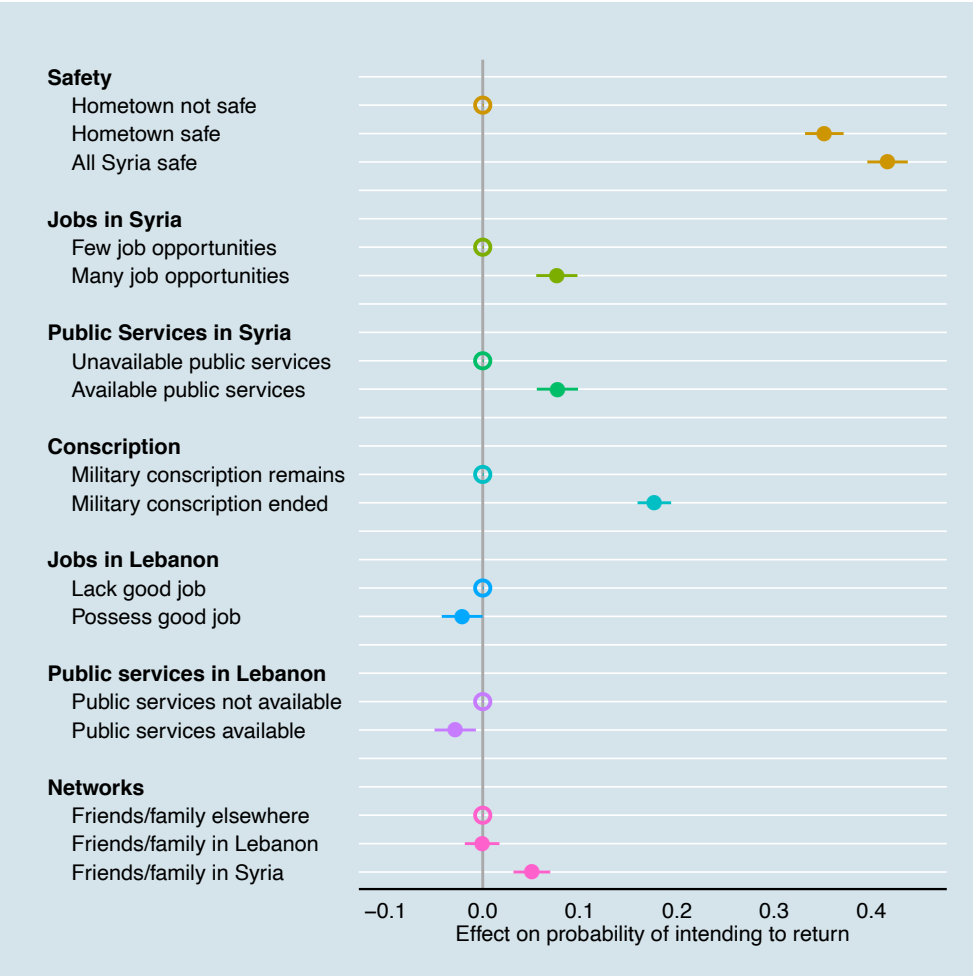

Figure 5: Conjoint Experiment Results. Each dot represents the effect on the probability that respondents would return to Syria in a given hypothetical situation, presented with its corresponding 95% confidence interval. The empty circles at $x = 0$ are reference categories. We cluster standard errors at the respondent level.

In the summer of 2019, enumerators interviewed a random sample of Syrians who received services from the NGO CARE during the study period. The participants were recruited from Syrian refugees living outside of camps, as do more than 90% of Syrians in Jordan (The World Bank, 2017, p. 93).

The two cases make for a valuable comparison given some key similarities and critical differences. Similar to Lebanon, Jordan hosts a large number of Syrian refugees relative to its population, and public discourse in the country widely frames refugees as having large negative economic and fiscal impacts. In contrast, the baseline rate of return intentions for

Syrians in Jordan is very low. When we asked Syrian refugees in Jordan if they plan to ever return to Syria, we find that a large majority of respondents (around 75%) reported that they never want to return to Syria. Further, unlike Lebanon, national political discourse in Jordan at the time of the survey was not pushing aggressively for Syrians to return. Therefore, the data enable us to examine whether our results from Lebanon pertain only to a context with major political pressure to return and where a large share of people hope to return home someday.

The difference in baseline return intentions between our samples in Lebanon and Jordan is likely driven by different selection into displacement to the countries. Similar to Lichtenheld (2020)'s theory of assortative displacement, we can imagine that refugees sort into host countries based on a range of personal characteristics, including their relationship to a war's armed groups in the country of origin, and these views may influence the return decision. First, we see a difference between the two samples in self-reported level of security in respondents' hometowns. As of summer 2019, 51% of the sample in Jordan said that their place of origin continues to be very dangerous. In contrast, only 28% of respondents in Lebanon said so, when we conducted our survey there a few months later in August–October 2019. Second, our fieldwork suggests that the political attitudes of Syrians living in Jordan tend to be more anti-regime whereas the Syrian population in Lebanon is more divided in its views toward the Syrian government, which aligns with public opinion surveys on the topic (Corstange, 2018).

Using our data from Jordan, we construct indices for dimensions of people's well-being in Jordan and prospective well-being in Syria. We then regress return intentions on the indices, as defined in the "individual indices" specification of Equation 1, to estimate the impact of each factor on peoples' stated plans to ever return to Syria.[13]

Figure 6 presents results from our analysis of the Jordan data. Despite the sizeable difference in baseline return intentions and the political climate, the drivers of return intentions in Jordan are strikingly similar to Lebanon. First, prospective conditions in Syria play an important role. We see that conditions in respondents' place of origin in Syria—specifically safety, economic prospects, and public services—are positively correlated with return intentions. Also, having family and friend networks in Syria is positively correlated with return intentions.

Second, in line with results from Lebanon, we do not find strong evidence that conditions

[13]The list of questions used in each index are included in Section 3 of the Appendix.

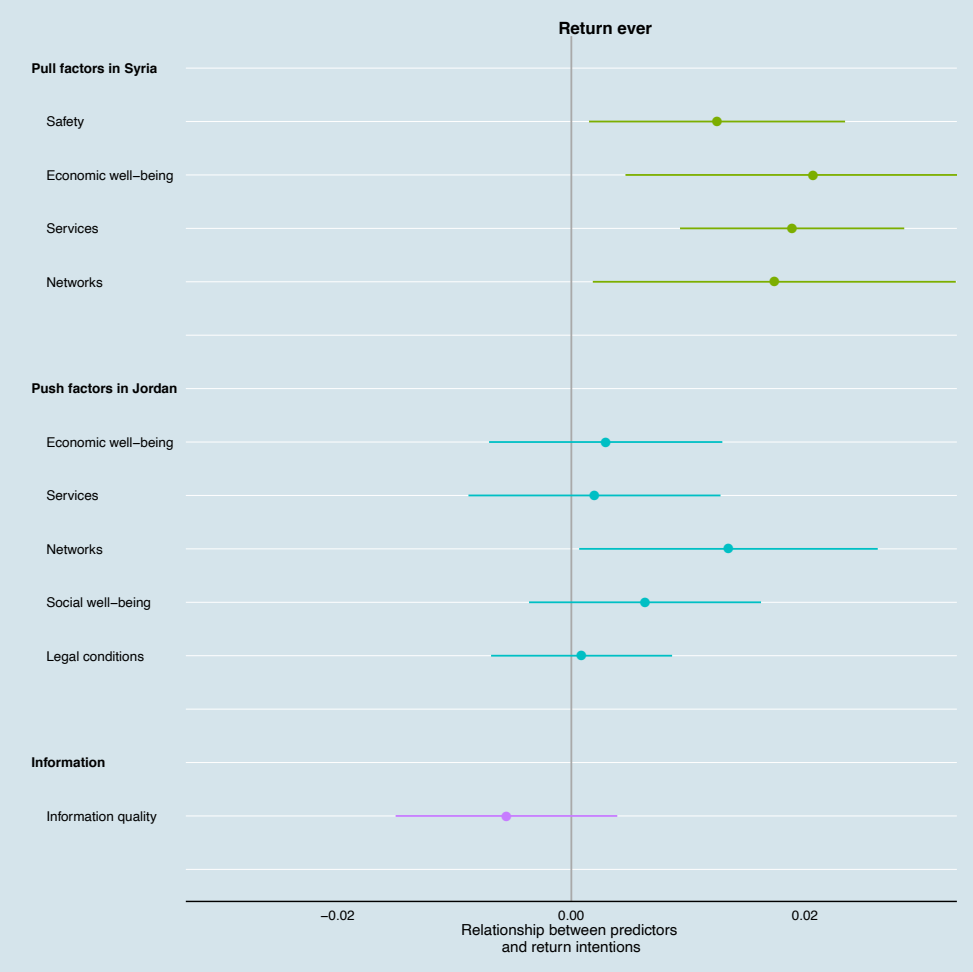

Figure 6: Index Results in Jordan

in Jordan drive return intentions. First, we see in Figure 6 that economic conditions, access to public services, social well-being, and legal conditions are not strongly associated with return intentions. Networks is the one dimension of conditions in Jordan where we find a relationship with return intentions. This contrasts with results from Lebanon, where social well-being is the only push factor that consistently predicts return intentions.

Finally, looking at the impact of information, we do not find evidence of a relationship between information quality and return intentions in Jordan. This contrasts with the evidence we found in Lebanon for the importance of information for Syrians' decision making about return.[14]

[14] We were not able to ask respondents in Jordan for the name of their hometown or district in Syria, preventing analysis of mobility cost.

8 Conclusion

In this paper, we further understanding of the dynamics of refugee return by examining four major drivers of return at the household level: push factors in the host country, pull factors in the home country, the cost of mobility, and the role of information. We test our hypotheses in the context of the Syrian refugee crisis in Lebanon using observational and experimental survey data from a representative sample of refugees and explore the external validity of our results with a second survey in Jordan.

Looking at the drivers of refugee return, we find strong evidence that the return intentions of Syrian refugees are shaped primarily by the situation in the home country (pull factors), rather than the dynamics in the hosting country (push factors). The most important pull factors are perceptions of safety, economic opportunity, the availability of public services, and the size of personal networks. Strikingly, we do not find evidence of a strong relationship between conditions in the hosting country and people's return intentions. The findings complement Ghosn et al. (2021), who find that Syrian refugees' pre-migration wartime experiences and sociopsychological factors related to their home country play large roles in shaping intentions to ever return. We add to this work by shedding light on the comparative importance of push and pull factors, including material well-being in the host country and country of origin, social networks, and access to trusted information about conditions in the home country.

In considering the broader import of these findings, it is useful to reflect on the benefits and limits of a study conducted on Syrian refugees in Lebanon (and Jordan). On the one hand, this context offers a number of important features including wide variation in push and pull factors. Further, the consistency of the results in Jordan suggests that the findings travel to neighboring countries. This is important as Syrian refugees in Turkey, Lebanon, and Jordan alone account for approximately 25% of the world's refugees[15] and approximately 85% of Syrian refugees live in neighbouring states[16] However, Lebanon differs from other contexts in three important ways. First, the median respondent in our data has been in Lebanon for six years at the time data collection, but many refugees worldwide have been displaced even longer. As displacement extends across generations, people who grew up in a host country may develop deeper economic and social ties, while holding fairly limited ties to the country that their parents or grandparents fled. Second, the dynamics of return

[15] UNHCR: https://tinyurl.com/2p868pnj. Accessed January 26, 2022.
[16] UNHCR: https://tinyurl.com/5n6t4saf. Accessed January 26, 2022.

may differ for refugees living in Western countries, where about 15% of refugees worldwide reside. These refugees are likely different in important ways from those who stay in the region of their home country, for example encountering a more welcoming legal environment, and facing much higher transport costs to return. Finally, our findings are driven both by conditions in Lebanon and Syria. Our interpretation of the relative importance of push and pull factors may only generalize to contexts in which high levels of ongoing violence impede return to the country of origin.

The paper also raises a number of important questions for a growing research agenda on refugee crises and the dynamics of return. First, further work is needed to understand the role of well-being in the host country in refugee decision-making. This concern is particularly relevant given the ongoing scholarly debate about the possibility that the presence of refugees may lead to deteriorating security conditions in hosting conditions (Salehyan and Gleditsch, 2006, Lischer, 2006, but see Zhou and Shaver, 2021). The weak relationship we find between host-country conditions and return intentions could emerge if well-being in the host country has countervailing negative and positive effects on people's return aspirations and preparations. People with greater resources may be both better off in exile but also more able to afford the costs of moving back home. Our results provide evidence for this possibility, where we see that better conditions have very small or null effects on people's intentions to return, but a positive association with concrete steps to prepare to return.

Second, although previous research explores *why* refugees seek out information about potential destinations (Holland and Peters, 2020), little is understood about *how* refugees acquire and assess information about the situation in potential destinations. It is intuitive that high-quality information will condition migration choices, especially given the potential negative consequences of returning prematurely to a dangerous context. But theories accounting for risk aversion would predict that the uncertainty that refugees have about the situation at home may lead them to underweight outcomes in the home country relative to those in the host country (see, e.g., Kahneman and Tversky, 1979). Given the complexity of our findings on information and the absence of a well-identified causal effect, further research is needed on how information quality influences return decisions.

Third, future research should explore the degree to which return intentions predict people's subsequent migration choices. Recent studies of labor migration with direct measurement of both migration intentions and behavior find that intentions were strong predictors of future emigration (see, e.g., Van Dalen and Henkens, 2013; Docquier, Peri and Ruyssen,

2014; Tjaden, Auer and Laczko, 2019). That being said, the possibility of large-scale long-term displacement in the Syria crisis may have important ramifications on lessons about the drivers of refugee return. One-third of our sample in Lebanon intends to not return to Syria, and previous studies focused on return migration find that labor migration and wartime displacement can lead to high rates of long-term residence in destinations (see, e.g., Constant and Massey, 2003; Camarena and Hagerdal, 2020).

Finally, refugees' decision-making about return is a product of both individual-level factors and aggregate shocks (e.g., economic crisis or civil war in the hosting country). In considering the role of push factors, it will be important for future work to examine how aggregate shocks condition household decision-making, such as studying cases with large changes over time in the security or economic conditions of refugees in a hosting country. It will be valuable to explore how shocks and household characteristics interact, something that we are not in a good position to explore with cross-sectional data.

We conclude with two key takeaways for policymakers and humanitarian organizations. First, the results reaffirm the fundamental humanitarian mandate of the refugee protection regime. Refugees are not economic migrants by another name. Many intend to return to their home country when threats to their physical, economic, and social well-being have decreased, and when they feel that they possess credible information. Even after years in a host country, people's migration choices do not appear to be driven by opportunities in that country. We find that more than two thirds of Syrians in Lebanon want to return home, and prospects for a good job and access to public services in Lebanon do not influence people's likelihood of staying. The evidence is clear that policies that deny rights to refugees or broader anti-refugee hostility are unlikely to drive people to return.

The findings also suggest that humanitarian agencies can support refugee well-being without undermining the goal of safe voluntary return. In light of our findings, efforts to deliver humanitarian assistance and provide economic opportunities may be unlikely to incentivize refugees to remain in the host country. Traditional development programs may be effectively applied with refugee populations, supporting refugees' economic integration, benefiting both refugees and host societies, and freeing refugees from a reliance on aid. This would, in turn, allow humanitarian agencies to focus their attention and resources on emergencies, rather than struggling to provide ongoing assistance in protracted displacement situations.

References

Adhikari, Prakash. 2013. "Conflict-Induced Displacement, Understanding the Causes of Flight." *American Journal of Political Science* 57(1):82–89.

Adida, Claire L. 2014. *Immigrant exclusion and insecurity in Africa*. Cambridge University Press.

Arias, María Alejandra, Ana María Ibáñez and Pablo Querubin. 2014. "The desire to return during civil war: Evidence for internally displaced populations in Colombia." *Peace Economics, Peace Science and Public Policy* 20(1):209–233.

Balcells, Laia and Abbey Steele. 2016. "Warfare, political identities, and displacement in Spain and Colombia." *Political Geography* 51:15–29.

Batista, Cátia and Francesco Cestari. 2016. "Migrant Intentions to Return: The Role of Migrant Social Networks." *Working paper* .

Beaman, Lori, Harun Onder and Stefanie Onder. 2021. "When Do Refugees Return Home? Evidence from Syrian Displacement in Mashreq." *Policy Research working paper; no. WPS 9688* .

Beber, Bernd, Philip Roessler and Alexandra Scacco. 2021. "Coping with Partition: Wealth, Security, and Migration in Post-Separation Sudan." *Working paper* .

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." *American Economic Review* 4(77):531–55.

Borjas, George J and Bernt Bratsberg. 1996. "Who leaves? The outmigration of the foreign-born." *The Review of Economics and Statistics* 78(1):165–76.

Camarena, Kara Ross and Nils Hagerdal. 2020. "When Do Displaced Persons Return? Postwar Migration among Christians in Mount Lebanon." *American Journal of Political Science* 64(2):223–239.

Constant, Amelie and Douglas S Massey. 2003. "Self-selection, earnings, and out-migration: A longitudinal study of immigrants to Germany." *Journal of Population Economics* 16(4):631–653.

Corstange, Daniel. 2018. "The Syrian Conflict and Public Opinion Among Syrians in Lebanon." *British Journal of Middle Eastern Studies* 46(1):178–200.

Crisp, Jeff. 2019. "Repatriation principles under pressure." *Forced Migration Review* (62):19–23.

Davenport, Christian, Will Moore and Steven Poe. 2003. "Sometimes you just have to leave: Domestic threats and forced migration, 1964-1989." *International Interactions* 29(1):27–55.

Docquier, Frédéric, Giovanni Peri and Ilse Ruyssen. 2014. "The cross-country determinants of potential and actual migration." *International Migration Review* 48(1):37–99.

Dustmann, Christian and Yoram Weiss. 2007. "Return migration: theory and empirical evidence from the UK." *British Journal of Industrial Relations* 45(2):236–256.

Fabbe, Kristin, Chad Hazlett and Tolga Sinmazdemir. 2019. "A persuasive peace: Syrian refugees' attitudes towards compromise and civil war termination." *Journal of Peace Research* 56(1):103–117.

Fearon, James D. and Andrew Shaver. 2020. "Civil war violence and refugee outflows." *Working paper* .

Ghosn, Faten, Tiffany S Chu, Miranda Simon, Alex Braithwaite, Michael Frith and Joanna Jandali. 2021. "The Journey Home: Violence, Anchoring, and Refugee Decisions to Return." *American Political Science Review* .

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed conflict 1946–2001: A new dataset." *Journal of peace research* 39(5):615–637.

Hainmueller, J., D.J. Hopkins and T. Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." *Political Analysis* 22(1):1–30.

Hainmueller, Jens and Daniel J. Hopkins. 2014. "Public Attitudes Toward Immigration." *Annual Review of Political Science* 17(1):225–249.

Harder, Niklas, Lucila Figueroa, Rachel M Gillum, Dominik Hangartner, David D Laitin and Jens Hainmueller. 2018. "Multidimensional measure of immigrant integration." *Proceedings of the National Academy of Sciences* 115(45):11483–11488.

Harpviken, Kristian. 2009. *Social networks and migration in wartime Afghanistan*. Springer.

Holland, Alisha C and Margaret E. Peters. 2020. "Explaining Migration Timing: Political Information and Opportunities." *International Organization* 74(3):560–583.

Hunt, Gary L and Richard E Mueller. 2004. "North American migration: returns to skill, border effects, and mobility costs." *Review of Economics and Statistics* 86(4):988–1007.

Jacobsen, Karen. 2005. *The economic life of refugees*. Kumarian Press.

Janmyr, Maja. 2016. "Precarity in Exile: The Legal Status of Syrian Refugees in Lebanon." *Refugee Survey Quarterly* 35(4):58–78.

Kahneman, D and A Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica* 47(2):263–291.

Khawaja, Marwan. 2003. "Migration and the reproduction of poverty: the refugee camps in Jordan." *International Migration* 41(2):27–57.

Lehmann, M Christian and Daniel Masterson. 2020. "Do Cash Transfers Reduce Anti-Refugee Violence? Evidence from Syrian Refugees in Lebanon." *American Political Science Review* 114(4):1335–1342.

Lichtenheld, Adam G. 2020. "Explaining Population Displacement Strategies in Civil Wars: A Cross-National Analysis." *International Organization* 74(2):253–294.

Lischer, Sarah Kenyon. 2006. "Dangerous Sanctuaries: Refugee Camps, Civil War, and the Dilemmas of Humanitarian Aid.".

Massey, Douglas S, Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino and J Edward Taylor. 1993. "Theories of International Migration: A Review and Appraisal." *Population and Development Review* 19(3):431–466.

Masterson, Daniel. 2020. "Refugee Networks, Cooperation, and Resource Access." *Working paper* .

McNamee, Lachlan. 2018. "Mass Resettlement and Political Violence: Evidence from Rwanda." *World Politics* 70(4):595–644.

Mixed Migration Centre, The. 2019. "Mixed Migration Review 2019.".

Moore, Will H and Stephen M Shellman. 2004. "Fear of persecution: Forced migration, 1952-1995." *Journal of Conflict Resolution* 48(5):723–745.

Moore, Will H and Stephen M Shellman. 2006. "Refugee or internally displaced person? To where should one flee?" *Comparative Political Studies* 39(5):599–622.

Moore, Will H and Stephen M Shellman. 2007. "Whither will they go? A global study of refugees' destinations, 1965–1995." *International Studies Quarterly* 51(4):811–834.

Mourad, Lama. 2017. ""Standoffish" Policy-making: Inaction and Change in the Lebanese Response to the Syrian Displacement Crisis." *Middle East Law and Governance* 9(3):249–266.

Onoma, Ato Kwamena. 2013. *Anti-refugee violence and African politics.* Cambridge University Press.

Pettersson, Therese. 2019. "UCDP/PRIO Armed Conflict Dataset Codebook v 19.1.".

Salehyan, Idean and Kristian Skrede Gleditsch. 2006. "Refugees and the Spread of Civil War." *International Organization* 60(02):335–366.

Scheve, Kenneth F. and Matthew J. Slaughter. 2001. "Labor Market Competition and Individual Preferences Over Immigration Policy." *Review of Economics and Statistics* 83(1):133–145.

Schewel, Kerilyn. 2020. "Understanding immobility: Moving beyond the mobility bias in migration studies." *International Migration Review* 54(2):328–355.

Schmeidl, Susanne. 1997. "Exploring the causes of forced migration: A pooled time-series analysis, 1971-1990." *Social Science Quarterly* 78(2):284–308.

Schon, Justin. 2018. Bullets, Bombs, and Conversations: How Civilians Make Their Migration Decisions During Conflict PhD thesis Indiana University.

Schon, Justin. 2019. "Motivation and opportunity for conflict-induced migration: An analysis of Syrian migration timing." *Journal of Peace Research* 56(1):12–27.

Schwartz, Stephanie. 2019. "Home, Again: Refugee Return and Post-Conflict Violence in Burundi." *International Security* 44(2):110–145.

Stark, Oded and David E Bloom. 1985. "The new economics of labor migration." *The American Economic Review* 75(2):173–178.

Steele, Abbey. 2009. "Seeking safety: Avoiding displacement and choosing destinations in civil wars." *Journal of Peace Research* 46(3):419–429.

Stefanovic, Djordje and Neophytos Loizides. 2011. "The way home: Peaceful return of victims of ethnic cleansing." *Human Rights Quarterly* 33(2):408–430.

Stefanovic, Djordje, Neophytos Loizides and Samantha Parsons. 2015. "Home is where the heart is? Forced migration and voluntary return in Turkey's Kurdish regions." *Journal of Refugee Studies* 28(2):276–296.

Stork, Joe and Fred Abrahams. 2004. "Sidelined: Human Rights in Postwar Iraq." *Human Rights Watch World Report 2004* .

The World Bank. 2017. "Syrian Refugees and Their Hosts in Jordan, Lebanon, and the Kurdistan Region of Iraq: Lives, Livelihoods, and Local Impacts." *Unpublished Report* .

Tjaden, Jasper, Daniel Auer and Frank Laczko. 2019. "Linking migration intentions with flows: Evidence and potential use." *International Migration* 57(1):36–57.

UNHCR. 2019*a*. "Global Trends Forced Displacement in 2018.".

UNHCR. 2019*b*. "UNHCR Refugee Population Statistics Database.".

UNHCR Resettlement. 2021.
URL: *https://www.unhcr.org/resettlement.html*

Van Dalen, Hendrik P and Kène Henkens. 2013. "Explaining emigration intentions and behaviour in the Netherlands, 2005–10." *Population Studies* 67(2):225–241.

Weber, Sigrid and Alexandra Hartman. 2022. "Property Rights and Post-Conflict Recovery: Theory and Evidence from IDP Return Movements in Iraq." *Working paper* .

Wood, Elisabeth Jean. 2008. "The social processes of civil war: The wartime transformation of social networks." *Annu. Rev. Polit. Sci.* 11:539–561.

Zhou, Yang-Yang and Andrew Shaver. 2021. "Reexamining the effect of refugees on civil conflict: a global subnational analysis." *American Political Science Review* 115(4):1175–1196.

Appendix

The Dynamics of Refugee Return:

Syrian Refugees and Their Migration Intentions*

February 13, 2022

Contents

*All analyses, unless otherwise noted, were pre-specified in [XXXX] registration [XXXXX]. This project was reviewed and approved by the [University XXXXX] Human Subjects Committee under IRB protocol [XXXXX].

1 Sampling Strategy

We conducted stratified multistage sampling. In the first stage, we selected localities based on two dimensions: the prevalence of Syrians and the majority sect. We drew Syrian population data from UNHCR registration records and Lebanese population data from voter registration records, since there is no current census available. In the second stage, we used a random walk procedure to select households within each locality. In the third stage, we selected a head of household—defined as a person regardless of gender who plays a large role in household decision-making.

For this study, we wanted to draw inferences about all individuals from Syria in Lebanon as of the study's start date irrespective of their legal status and access to services and support. Thus, this study includes Syrian citizens regardless of whether they registered with UNHCR, and also Palestinian refugee from Syria (PRS), regardless of whether they registered with UNRWA.

First sampling stage: Locality selection

The sampling frame for the first stage is the list of localities published by the Lebanese Council for Development and Reconstruction (CDR) in 2018 and the 2018 UNHCR data on the number of registered Syrians by locality. Each locality is identified by way of its administrative affiliation—Kaza and Mohafza.

The localities were sorted into nine strata depending on their prevalence of Syrian population and the majority Lebanese sectarian group, as follows:

- Dimension 1: Prevalence of Syrian refugees
  - Low prevalence: where the Syrian population accounted for less than 20% of the total population.
  - Medium prevalence: where the Syrian population is between 20% and 50% of the total population.
  - High prevalence: where the Syrian population accounted for over 50% of the total population.
- Dimension 2: Sectarian component of the non-Syrian population
  - Sunni non-Syria majority: More than 50% of the documented non-Syrian population of the village is Sunni.
  - Non-Sunni non-Syrian majority: More than 50% of the documented non-Syrian population of the village is of a single non-Sunni sectarian group.

- Mixed: No single sectarian group makes up more than 50% of the village population.

In the first stage sample we selected 150 Lebanese localities with replacement, leading to 93 unique localities. We then randomly sampled people in each locality. Our goal is to say something about the drivers of return in the overall population. Hence, in our main analysis, we cluster standard errors by locality, since there are localities in the population of interest beyond those captured in the sample (Abadie et al., 2017).

The distribution of the sample of 150 localities into strata faced the classical dilemma of whether doing it in proportion to the population of the strata or rather selecting the same number of localities in each stratum. Since both are important considerations for our study, and we want to study subnational variation while also making nationwide claims, we followed a first-stage sampling distribution in accordance to Markward's rule (also known as the '50/50 equal/proportional allocation' rule), which is generally considered a good compromise between the two extremes. Given the small number of mixed localities in the pure PPS sample, the 50/50 equal/proportional allocation has an additional desirable feature of avoiding any bins with a very small number of localities.

Second-stage sampling

Enumerators interviewed 20 households for each sample. In order to capture all Syrians and not only registered refugees, we opted for a random walk strategy rather than sampling from UNHCR's registration database. To do so, team leaders initially met with local key informants (such as the head of the municipal government) and had a map print out of the locality. The local key informant was asked to draw boundaries around the neighborhoods that include Syrians and the rough proportion of Syrians in each part of the town. The team leader then assigned enumerators in proportion to Syrians across the towns and provided them with a starting direction (N, NE, E, SE, S, SW, W, NW) for the day. Given that streets limit the direction of movement, the enumerator were asked to proceed along the street that is closest to the starting direction indicated in the enumerator schedule. The enumerators were told to interview a household in every third building.

Third-stage sampling: Respondent Selection

Enumerators interviewed a head of household—a person with primary decision-making responsibility in the household. An eligible 'head of household' should know the details of the household, its economic situation, and important household decisions. If a household had multiple adults who share decision-making responsibility, this classifies as there being multiple heads of household, in which case enumerators would interview whichever head of household was at home.

Response Rate

The team leaders recorded the total number of doors they knocked on each day and kept track of refusal to participate as well as non-answers. Out of 3,882 interview attempts, enumerators conducted 3,003 interviews. This resulted in a response rate of 77%.

Replacement

We needed to drop one research site in the Baalbek area from our sample because a shawish refused us entry to the informal settlement. We returned to our original sampling methodology and drew another town in order to replace this site.

2 Survey questions

2.1 Dependent variables

Our outcome of interest is intention to return. The survey included multiple questions about stated short-term and long-term return intentions as well as return preparations. Table A1 shows the list of questions used to measure these concepts.

Short term return intentions	**Long term return intentions**	**Return preparations**
Do you (the respondent) plan to return to Syria in the next 12 months?	Two years from now, where do you expect to actually be living?	Have you or your immediate family been saving resources in order to prepare for your return to Syria?
Are other members of your household planning to return in the next 12 months?	Do you hope to move back to Syria and live there one day?	Have you or your immediate family prepared any legal paperwork, such as marriage documents, birth certificates, or proofs of property to prepare for your return to Syria?
	Would you say it is correct that you don't want to return no matter what happens?	Have you or your immediate family reached out to Lebanese authorities to discuss returning to Syria?
		Have you or your immediate family reached out to UNHCR to discuss returning to Syria?
		Have you or anyone in your immediate family made a scoping trip back to Syria to learn about the situation there?
		Have there been times in the past 12 months when you were planning to return and aborted those plans?

Table A1: Questions about return intentions and preparations

2.2 Independent variables

2.2.1 Push factors from Lebanon

We use several measures of push factors in Lebanon. Table A2 shows questions used to build the economic well-being and access to services in Lebanon indices. Table A3 shows the indices that measure social well-being, legal conditions, and networks in Lebanon.

Economic well-being in Lebanon	**Services in Lebanon**
Do you possess the status that allows you to work legally in Lebanon (do you currently possess a work permit or Lebanese residency)?	Have you been sick in the past 6 months to the point of requiring medical treatment? (To enumerator: this does not mean that they actually saw a doctor, just that they needed to be treated. This includes chronic illnesses but not common illnesses like cold)
During the past 4 weeks, how many days did you work outside home to make money? (If they did not work, enter 0.)	(If yes) Were you able to see a doctor? (To Enumerator: Going to a pharmacy does not count as seeing a doctor)
(If worked) On the days that you worked during the past 4 weeks, how many hours did you usually work per day?	(For each household member) Has (household member) has been sick in the past 6 months to the point of needing medical treatment? (To enumerator: this does not mean that they actually saw a doctor, just that they needed to be treated. This includes chronic illnesses but not common illnesses like cold)
(If worked) During the past 4 weeks, how much money did you make in total?	(If yes) Was (household member) able to see a doctor? (To Enumerator: Going to a pharmacy does not count as seeing a doctor)
Can you please tell me which sources your household normally receives income from? By household we mean people who are family members or close relatives and who live under the same roof and share meals with you: Aid or assistance from other organizations (such as the United Nations, other international NGOs, local Lebanese charities, etc.)	In Lebanon, how difficult or easy would it be for you to do each of the following? See a doctor (aside from the cost)
In the last month, how much money did your household withdraw in total using cards from humanitarian organizations? To clarify, I do not mean using a card to buy from certain shops. I mean using a card to go to and ATM and withdraw cash. (This includes the 260 and other cash aid.) Enumerator: enter the sum of all cash aid used by household members in the last month.	Do you think it is hard for your or your family to access healthcare in Lebanon because you are Syrian? (To enumerator: if hard but not because they are Syrians, select "Not hard because we are Syrians")
In the last month, how much money did your household spend in total using WFP (World Food Program) support, that is, using a card to buy food only from certain shops? (We are referring to the taghziye program.) Enumerator: enter the sum of all food card aid used by household members in the last month.	Need school: Whether any children between ages of 6 and 18 never studied OR (did not finish primary school and are above 10 years old) OR (are not currently attending school)
To enumerator: This question is written in colloquial Arabic. Read it loud as it is and do not try to summarize it. If you had to live solely on your remaining savings and assets, without any income or debt, approximately how many months of expenses and spending would your savings and assets support you? Note that we're talking about the assets and savings of your household in Lebanon.	Now we would like you to think about all the areas you have lived since moving to Lebanon since you arrived here in (insert year). How many different towns have you lived in Lebanon (including this place) since you came here in (insert year)? (To measure stability in Lebanon, if always in the same town, the variable gets a value of 3 (most stable). If lived in two towns, the variable gets a value of 2. If lived in more than two towns, the variable gets a value of 1 (least stable))
Does this dwelling have the following items that you are able to use? Refrigerator, Washing Machine, Oven, Desktop or Laptop computer at home, Car, Microwave oven, Television, Internet connection at home (other than through a smartphone, not through a neighbor but owned at home), Indoor toilet, Central heating	Which year did you start living in in this area (neighborhood or town)? (To code stability in current town, we subtracted answer from 2019 to find number of years in this town then we cut the answers into quantiles)
How does the aid that you're currently receiving compare to the amount of aid that you were receiving a year ago? (Enumerator: Here we are referring to cash and all other aid.)	(If age of a child <18 AND child not currently attending school) Why is (child) not attending school? (Do not read options. Let respondent answer and select all that apply).
To enumerator: This question is written in colloquial Arabic. Read it loud as it is and do not try to summarize it. Then read all the answer options. In a typical month, what share of your household's monthly expenses and spending needs you are you able to satisfy from household members' income?	In Lebanon, how difficult or easy would it be for you to do each of the following? Get help with legal problems
What is the approximate total value of assets and cash that you possessed when you first came to Lebanon? Note that we're talking about the assets and savings of your household in Lebanon and not anything you left in Syria.	Does this dwelling have the following items that you are able to use? Running water
Does your household income vary from month to month?	
Did you or anyone in your household work in Lebanon before 2011?	
What was your total household income in the past month? By household we mean people who are family members or close relatives and who lived under the same roof and share meals with you. Enumerator: This does not include income from aid.	

Table A2: Push factors: Economic well-being and access to services in Lebanon

Social well-being in Lebanon	Legal conditions in Lebanon	Networks in Lebanon
How would you describe your relationship with Lebanese people?	Do you possess the status that allows you to work legally in Lebanon (does (respondent) currently possess a work permit or Lebanese residency))?	How many of (original household members from Syria) (excluding yourself) are living in Lebanon now?
Have you been detained by Lebanese authorities since arriving to Lebanon? Please note that we will not share this information.	What is your status with UNHCR? Please note that we will not share this information with anyone (Enumerator: If the respondent says registered or recorded, please ask to see the UNHCR registration file with names of registered individuals).	Please think about the Lebanese people in your phone contacts. With how many of them did you have a conversation–either by phone, messenger chat, face-to-face, or text exchange –in the last week? Note that this does not include service provides such as the United Nations or NGOs
Do you think it is hard for you or your family to get housing in Lebanon because you are Syrian? To enumerator: if hard but not because they are Syrians, select "Not hard because we are Syrians"	(If this person is a Palestinian from Syria) What is your status with UNRWA? Please note that we will not share this information with anyone.	Please think about the Syrians in Lebanon in your phone contacts. With how many of them did you have a conversation–either by phone, messenger chat, face-to-face, or text exchange –in the last week?
In your personal experiences, How friendly or hostile would you describe your personal experiences with Lebanese authorities in this area (town/neighborhood)?		Outside of your household, do you have any close Lebanese relatives in this area or elsewhere in the country?
In your personal experiences, how friendly or hostile would you describe your personal experiences with the Lebanese public in this area (town/neighborhood)? We are asking about the Lebanese general public, not authorities.		In the last 12 months, how often did you share a meal with Lebanese people who are not part of your family? (To enumerator: Those do not need to be friends. They can be people at work or other people.)
How well do you understand the important issues facing Lebanon?		
In the last 12 months, how often did you typically discuss major issues facing Lebanon with others?		
In Lebanon, how difficult or easy would it be for you to do each of the following? Search for a job		
Which year did you move to Lebanon to stay here until now?		
Does this town currently have curfews for Syrians?		
(If no) Has this town had curfews for Syrians in the last two years?		
How often do you feel like an outsider in Lebanon?		
(To enumerator) Did the respondent speak Arabic well?		
What is the highest level of education you have completed?		
How well can you read and write?		
How often are you personally able to travel freely and safely around this area of Lebanon?		
Are all your household members able to move freely in this town?		

Table A3: Push factors: Social well-being, legal conditions, and networks in Lebanon

2.2.2 Pull factors in Syria

For pull factors in Syria, Table A4 shows questions used to build the safety, control, and economic well-being in Syria indices. Table A5 shows the indices that measure services and networks in Syria.

Safety in Syria	**Control in Syria**	**Economic well-being in Syria**
How would you describe the current risk to civilians' physical safety in (place of origin)?	Who mainly controlled (place of origin) in the month before you left? Syrian army; Opposition forces such as the FSA; Jabhat al-Nusra; ISIS; Kurdish forces; Russian forces; Turkish forces; It was contested	How would you describe the current availability of jobs at present in (place of origin) currently?
To measure sympathy with opposition, we examine difference between trusting two anti-regime media (Al-Jazeera and Al-Arabiya) and two pro-regime media (Al-Mayadeen and Al-Manar) using the question: How trustworthy would you say each of the following news sources is?	Who mainly controls (place of origin) currently? [Same options]	What is the total amount of outstanding debts you currently have in Syria? This includes any debts on unpaid electricity, water, or other bills while you were away
Were there anti-regime protests in (place of origin) in 2011 and 2012?	Did ISIS control (place of origin) at all during the conflict?	Did you or your immediate family own (not rent) the following items in Syria? (ask for each): House (not an apartment); Apartment (other than their house, if they owned a house); Land
Have you suffered physical or psychological harm because of violence in Syria?		(If they stated that they own land in previous question) Do you think you would be able to continue as owner and operator of this land if you returned to Syria?
How do you expect the safety situation to be in (place of origin) one year from now?		(If owned house/apartment/land) Do you or your immediate family have property documents that prove you are the owner?
Conscription: To examine if someone in household is of/near conscription age, we saw if household includes any males born between 1977 and 2003.		

Table A4: Pull factors: Safety, control, and economic well-being in Syria

Services in Syria	**Networks in Syria**
As far as you know, how many hours per day is there electricity in (place of origin) currently?	How many of (household members from Syria before leaving) are living in Syria now?
As far as you know, how many hours per day is there running water in (place of origin) currently?	Approximately how many Syrian relatives or friends who have lived in Lebanon have gone back to Syria with the goal of staying there?
As far as you know, are schools operating in (place of origin) during the school year?	Next, think about your Syrian friends and relatives from (place of origin) who have lived in Lebanon. Approximately how many of them have gone back to (place of origin)?
As far as you know, are health centers operating in (place of origin) currently?	How many people who were in your household in Lebanon at some point since 2011 have gone back to Syria, regardless of where they are now?
How good do you think that public service provision in (place of origin) (such as health centers, schools, infrastructure) will be one year from now?	

Table A5: Pull factors: Services and networks in Syria

2.2.3 Confidence in information and mobility

Table A6 describes our measures of confidence in information and mobility costs.

Confidence in information	**Mobility cost**
To enumerator: This question is written in colloquial Arabic. Read it loud as it is and do not try to summarize. Then read all the answer options. When thinking of your knowledge about the safety situation in (place of origin), would you say that... : I know enough to be confident I understand the situation (1); I don't know enough, and I want to know more (0); I don't know much but do not feel the need to know (0)	We calculate travel distance from each survey respondent's town of residence in Lebanon to their hometown in Syria, via the Beirut–Damascus highway and border crossing. Travel routes were calculated using the Google Maps API. We used the R package mapsapi and commands mp_directions() and mp_get_routes().
To enumerator: This question is written in colloquial Arabic. Read it loud as it is and do not try to summarize. Then read all the answer options. When thinking of your knowledge about employment opportunities in (place of origin), would you say that... : (same options as above)	The log of household size
To enumerator: This question is written in colloquial Arabic. Read it loud as it is and do not try to summarize. Then read all the answer options. When thinking of your knowledge about the availability of public services (such as health centers, schools, or water) in (place of origin), would you say that... : (same options as above)	
To enumerator: This question is written in colloquial Arabic. Read it loud as it is and do not try to summarize. Then read all the answer options. When thinking of your knowledge about the conscription requirements by the Syrian military, would you say that... : (same options as above)	
Now, please think about the person you communicate with the most who is currently living in Syria. (If respondent answers: I don't communicate with anyone in Syria, this is coded as 0, otherwise 1)	
In general, how often do you communicate with people in (place of origin)?	

Table A6: Confidence in information and mobility costs

3 Index construction and descriptives

We measured four key concepts with multiple independent variables using the first principal component.

1. Well-being in Lebanon
2. Expected well-being in Syria
3. Confidence in information
4. Preparation to return

The component variables are mentioned in Section 2. For the first three independent variables, we have many input variables. Note that all results for the survey in Lebanon impute missing values using multivariate imputation by chained equations. We specify 10 imputations and use random forest to predict missing values using the `mice()` package in R. We use Rubin's rules (Rubin, 1987) when pooling estimates across imputations.

Our primary analysis was based on regression models with indices constructed using polychoric PCA unless any variable in the index had too many categories (above 8 categories, in which case we used Pearson correlations) (Hainmueller, Hangartner and Pietrantuono, 2017). Scales for individual questions were reversed as necessary to simplify interpretation (to make sure they were all in the same positive direction). In addition to the indices constructed using the first principal components, we ran mean effects indices (aka z-scores) as a robustness check (Kling, Liebman and Katz, 2007). We constructed those indices by standardizing each variable (demeaning and dividing it by the standard deviation). We then summed the standardized variables and then standardized the sum again.

This section shows the component variables of each of the indices as well as descriptive statistics for these variables. We also show descriptive statistics for the control variables we included in the main regressions.

3.1 Push factors from Lebanon

Economic well-being in Lebanon

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Aid: atm card	0.63	0.58	1.39	0.00	0.00	7.00	0.30%
Aid change from last year	1.49	1.48	0.56	1.00	1.00	3.00	9.16%
Aid: wfp card	1.12	1.02	1.68	0.00	0.00	7.00	0.13%
Assets: months left	0.13	0.15	0.62	0.00	0.00	12.00	0.00%
Assets: value upon arrival	0.62	0.70	2.26	0.00	0.00	16.00	0.27%
Ability to cover expenses	2.78	2.88	1.26	1.00	3.00	5.00	0.17%
Income	1.75	1.90	3.07	0.00	0.00	13.00	0.10%
Work days past 4 weeks	5.82	6.25	9.91	0.00	0.00	28.00	0.00%
Work hours past 4 weeks	2.92	3.13	4.74	0.00	0.00	24.00	0.00%
Able to work legally	0.10	0.11	0.31	0.00	0.00	1.00	0.03%
Income source: aid	0.36	0.33	0.47	0.00	0.00	1.00	0.03%
Stable household income	0.81	0.79	0.71	0.00	1.00	2.00	0.20%
Household income	4.36	4.72	3.50	0.00	5.00	17.00	0.93%
HH worked in Leb. before 2011	0.23	0.24	0.43	0.00	0.00	1.00	0.00%
Own refrigerator	0.80	0.81	0.39	0.00	1.00	1.00	0.00%
Have indoor toilet	0.83	0.83	0.37	0.00	1.00	1.00	0.00%
Have central heating	0.05	0.05	0.22	0.00	0.00	1.00	0.13%
Own washing machine	0.68	0.67	0.47	0.00	1.00	1.00	0.00%
Own oven/stove	0.79	0.79	0.41	0.00	1.00	1.00	0.00%
Own computer	0.02	0.02	0.13	0.00	0.00	1.00	0.00%
Own car	0.02	0.02	0.14	0.00	0.00	1.00	0.00%
Own microwave oven	0.04	0.05	0.21	0.00	0.00	1.00	0.00%
Own television	0.78	0.79	0.41	0.00	1.00	1.00	0.00%
Have internet	0.24	0.25	0.44	0.00	0.00	1.00	0.07%

Table A7: Summary statistics of variables included in constructing the economic well-being in Lebanon index

Social well-being in Lebanon

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Never had curfews	0.65	0.63	0.48	0.00	1.00	1.00	0.93%
Authorities discrimination (higher is less)	2.71	2.73	0.74	1.00	3.00	4.00	15.38%
Ease of mobility	3.16	3.16	1.09	1.00	4.00	4.00	0.13%
Ease of mobility for household	1.62	1.61	0.78	1.00	1.00	3.00	0.53%
Public discrimination (higher is less)	2.97	2.98	0.68	1.00	3.00	4.00	1.60%
Arabic speaking ability	2.87	2.87	0.34	1.00	3.00	3.00	1.27%
Relation with Lebanese	3.66	3.68	0.95	1.00	4.00	5.00	0.33%
Discuss Lebanese politics	1.35	1.38	0.94	1.00	1.00	5.00	0.47%
Ease job search	1.59	1.61	1.02	1.00	1.00	5.00	0.83%
Literacy level	2.10	2.09	0.72	1.00	2.00	3.00	0.00%
Feeling outsider (higher is less)	2.95	2.96	1.29	1.00	3.00	5.00	0.03%
Know Lebanese politics	2.33	2.33	1.24	1.00	2.00	5.00	0.83%
Time in Lebanon	5.52	5.46	2.05	0.00	6.00	9.00	0.07%
No curfew now	0.75	0.73	0.45	0.00	1.00	1.00	0.50%
Housing discrimination (higher is less)	2.48	2.49	0.69	1.00	3.00	3.00	0.33%
Never detained	0.94	0.94	0.23	0.00	1.00	1.00	0.20%

Table A8: Summary statistics of variables included in constructing the social well-being in Lebanon index

Services in Lebanon

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Can access legal services	1.66	1.66	1.07	1.00	1.00	5.00	2.73%
No healthcare discrimination	2.37	2.37	0.73	1.00	3.00	3.00	0.47%
Not sick	0.83	0.83	0.38	0.00	1.00	1.00	0.07%
Received treatment (if sick)	0.94	0.95	0.23	0.00	1.00	1.00	0.00%
Can access doctor	2.80	2.87	1.66	1.00	3.00	5.00	0.70%
No Kids need school	0.59	0.59	0.49	0.00	1.00	1.00	0.03%
Have running water	0.80	0.81	0.39	0.00	1.00	1.00	0.00%
No HH member sick	0.78	0.79	0.41	0.00	1.00	1.00	0.17%
HH members treated if sick	0.91	0.91	0.28	0.00	1.00	1.00	0.03%
School not preventive	0.96	0.96	0.20	0.00	1.00	1.00	0.07%
Period in current town	2.39	2.37	1.16	1.00	2.00	4.00	0.10%
Towns lived in Lebanon	2.64	2.65	0.63	1.00	3.00	3.00	0.10%

Table A9: Summary statistics of variables included in constructing the services in Lebanon index

Legal situation in Lebanon

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Registered with UNHCR/UNRWA (or resident)	0.81	0.80	0.40	0.00	1.00	1.00	0.00%
Legal resident in Lebanon	0.04	0.05	0.21	0.00	0.00	1.00	0.00%

Table A10: Summary statistics of variables included in constructing the legal situation in Lebanon index

Networks in Lebanon

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Syria HH members living in Leb. now	4.58	4.45	3.51	0.00	4.00	15.00	0.00%
Lebanese phone contacts	1.71	1.71	1.11	1.00	1.00	5.00	0.53%
Share meals with Lebanese	1.51	1.53	1.08	1.00	1.00	5.00	0.17%
Syrian phone contacts	3.22	3.20	1.37	1.00	3.00	5.00	0.30%
Lebanese relatives	0.22	0.22	0.55	0.00	0.00	2.00	0.10%

Table A11: Summary statistics of variables included in constructing the networks in Lebanon index

3.2 Pull factors in Syria

Safety in Syria

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
HH male at conscription age	0.80	0.81	0.39	0.00	1.00	1.00	0.00%
Exposed to violence	0.28	0.27	0.45	0.00	0.00	1.00	0.03%
Follow anti-regime media more than pro-regime media	0.20	0.21	0.41	0.00	0.00	1.00	0.00%
Hometown had protests	0.67	0.67	0.47	0.00	1.00	1.00	2.40%
Expect hometown to be safe	2.54	2.53	0.86	1.00	3.00	4.00	18.81%
Current safety in hometown	2.05	2.06	0.88	1.00	2.00	4.00	7.13%

Table A12: Summary statistics of variables included in constructing the safety in Syria index. Respondents were coded as following anti-regime media more than pro-regime media if they reported following Al-Jazeera or Al-Arabiya (anti-regime) more than Manar/Mayadeen (pro-regime).

Control in Syria

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Contested Now	0.05	0.05	0.22	0.00	0.00	1.00	5.16%
Contested before leaving	0.15	0.15	0.36	0.00	0.00	1.00	3.06%
Controlled by Kurds now	0.13	0.14	0.35	0.00	0.00	1.00	5.16%
Controlled by oppsn/FSA now	0.09	0.09	0.29	0.00	0.00	1.00	5.16%
Controlled by regime now	0.69	0.67	0.47	0.00	1.00	1.00	5.16%
Controlled by Russia now	0.00	0.01	0.08	0.00	0.00	1.00	5.16%
Controlled by Turkey now	0.01	0.02	0.12	0.00	0.00	1.00	5.16%
Controlled by Kurds before leaving	0.02	0.03	0.17	0.00	0.00	1.00	3.06%
Controlled by oppsn/FSA before leaving	0.28	0.29	0.45	0.00	0.00	1.00	3.06%
Controlled by regime before leaving	0.38	0.36	0.48	0.00	0.00	1.00	3.06%
Controlled by Russia before leaving	0.00	0.00	0.06	0.00	0.00	1.00	3.06%
Controlled by Turkey before leaving	0.00	0.00	0.05	0.00	0.00	1.00	3.06%
Controlled by ISIS	0.37	0.39	0.49	0.00	0.00	1.00	5.13%

Table A13: Summary statistics of variables included in constructing the control in Syria index

Economic well-being in Syria

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Debt in Syria	0.31	0.32	0.66	0.00	0.00	2.00	4.96%
Job situation in origin	1.55	1.54	0.70	1.00	1.00	4.00	9.62%
Home ownership docs (1 for some, 2 for everything)	0.58	0.62	0.91	0.00	0.00	2.00	1.90%
Can operate land in future	0.12	0.13	0.33	0.00	0.00	1.00	2.33%
Own house in Syria	0.62	0.64	0.48	0.00	1.00	1.00	0.30%
Own apt in Syria	0.07	0.07	0.25	0.00	0.00	1.00	0.37%
Own land in Syria	0.21	0.21	0.41	0.00	0.00	1.00	0.63%

Table A14: Summary statistics of variables included in constructing the economic well-being in Syria index

Services in Syria

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Electricity in origin	2.51	2.51	1.27	1.00	3.00	5.00	15.25%
Health services in origin	0.36	0.36	0.48	0.00	0.00	1.00	19.21%
Expect services to improve in 1 year	2.43	2.42	0.84	1.00	3.00	4.00	17.82%
Schools in origin	0.34	0.34	0.47	0.00	0.00	1.00	19.28%
Running water in origin	2.47	2.46	1.23	1.00	3.00	5.00	14.72%

Table A15: Summary statistics of variables included in constructing the services in Syria index

Services in Syria

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Electricity in origin	2.51	2.51	1.27	1.00	3.00	5.00	15.25%
Health services in origin	0.36	0.36	0.48	0.00	0.00	1.00	19.21%
Expect services to improve in 1 year	2.43	2.42	0.84	1.00	3.00	4.00	17.82%
Schools in origin	0.34	0.34	0.47	0.00	0.00	1.00	19.28%
Running water in origin	2.47	2.46	1.23	1.00	3.00	5.00	14.72%

Table A16: Summary statistics of variables included in constructing the services in Syria index

Networks in Syria

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
No. HH members returned to Syria	0.10	0.11	0.65	0.00	0.00	5.00	0.27%
Relatives permanently return to Syria	0.97	1.05	2.99	0.00	0.00	15.00	1.40%
Relatives return to origin	0.53	0.60	2.44	0.00	0.00	15.00	1.17%
Syria HH members living in Syria now	1.86	2.00	3.36	0.00	0.00	15.00	0.10%

Table A17: Summary statistics of variables included in constructing the networks in Syria index

3.3 Confidence in information

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Know Syr. conscription policy	0.28	0.30	0.46	0.00	0.00	1.00	0.23%
Know employment in origin	0.27	0.29	0.46	0.00	0.00	1.00	0.13%
Know safety in origin	0.26	0.27	0.45	0.00	0.00	1.00	0.10%
Know services in origin	0.25	0.27	0.44	0.00	0.00	1.00	0.20%
Communication freq. with origin	2.39	2.45	1.69	1.00	2.00	6.00	0.03%
Communication with someone in Syria	0.61	0.62	0.49	0.00	1.00	1.00	0.03%

Table A18: Summary statistics of variables included in constructing the confidence in information index

3.4 Mobility

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
household size (logged)	1.50	1.48	0.56	0.00	1.61	2.89	0.00%
travel distance (logged)	12.83	12.85	0.52	11.13	12.98	15.63	0.37%

Table A19: Summary statistics of variables to measure mobility

3.5 Preparation to return

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Planned to return but aborted	0.03	0.03	0.17	0.00	0.00	1.00	0.27%
Reached to Leb. authorities about return	0.01	0.01	0.08	0.00	0.00	1.00	0.13%
Prepared docs for return	0.02	0.03	0.17	0.00	0.00	1.00	0.10%
Saved resources for return	0.02	0.02	0.15	0.00	0.00	1.00	0.10%
Conducting scoping trip to Syria	0.03	0.03	0.17	0.00	0.00	1.00	0.10%
Reached to UNHCR about return	0.01	0.01	0.09	0.00	0.00	1.00	0.17%

Table A20: Summary statistics of variables included in constructing the preparation to return index

3.6 Covariates

The following covariates are included in the regressions. When including fixed effects, we also add locality fixed effects in Lebanon and Syria.

Variable	Weighted.Mean	Unweighted.Mean	Std.Dev	Min	Median	Max	Pct.Missing
Household includes elderly	0.04	0.04	0.19	0.00	0.00	1.00	0.00%
Female headed single-parent household	0.11	0.10	0.30	0.00	0.00	1.00	0.03%
High school graduate	0.11	0.12	0.32	0.00	0.00	1.00	0.03%
Hezbollah controlled area	0.16	0.17	0.38	0.00	0.00	1.00	0.00%
Location: Tental settlement	0.35	0.33	0.47	0.00	0.00	1.00	0.00%
Sick required medical treatment	0.30	0.30	0.46	0.00	0.00	1.00	0.23%
Syria origin: urban	0.23	0.23	0.42	0.00	0.00	1.00	0.20%
Household includes toddler	0.45	0.44	0.50	0.00	0.00	1.00	0.00%

Table A21: Summary statistics of variables included as controls in the regressions

4 Scree plots for principal component analysis

The following figures display how eigenvalues change with each additional component for the indices we created using PCA. Throughout the analysis, we used the first principal component.

4.1 Push factors from Lebanon

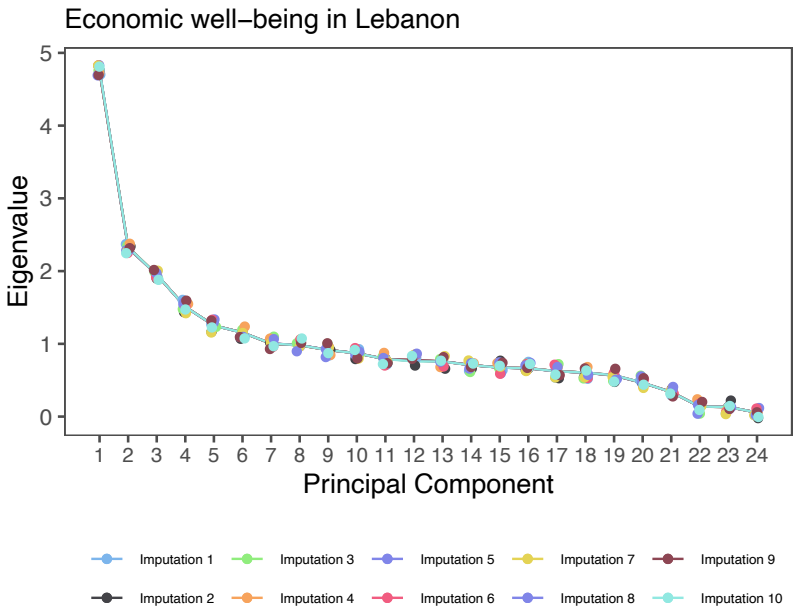

Figure A1: Screeplot for the economic well-being in Lebanon index

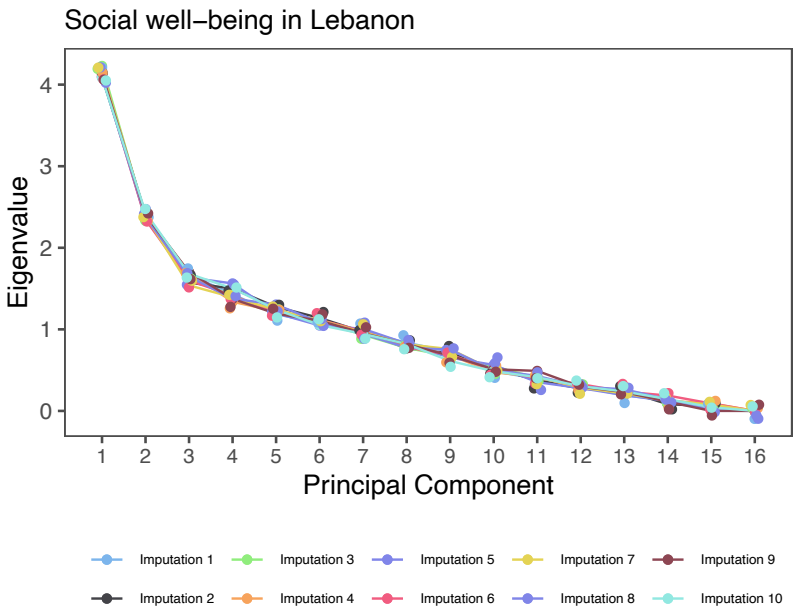

Figure A2: Screeplot for the social well-being in Lebanon index

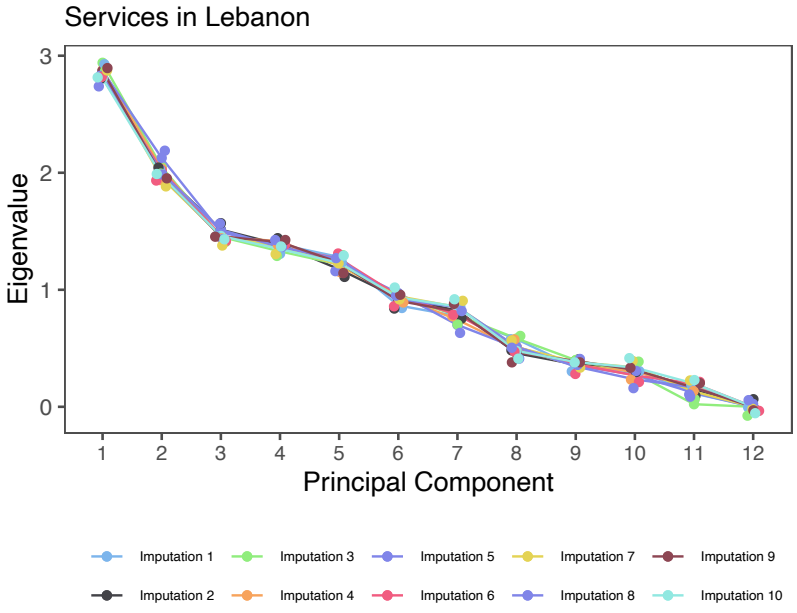

Figure A3: Screeplot for the services in Lebanon index

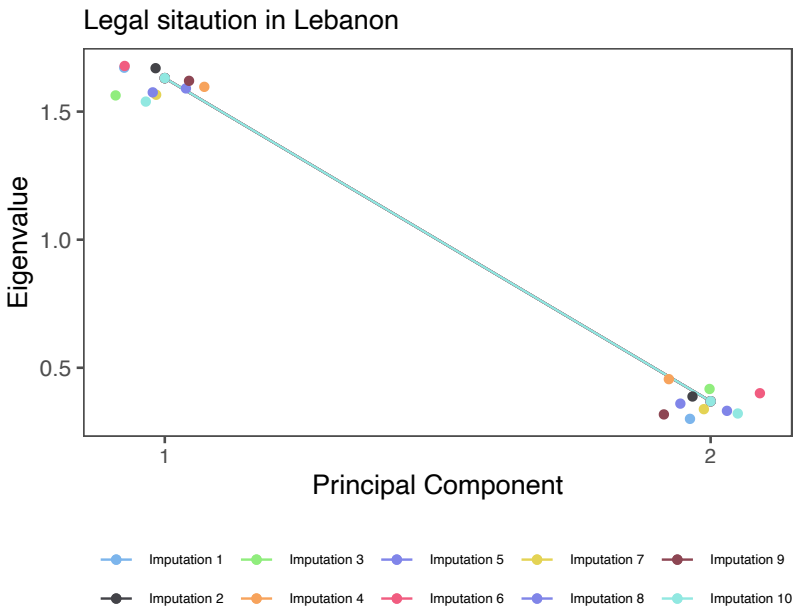

Figure A4: Screeplot for the legal situation in Lebanon index

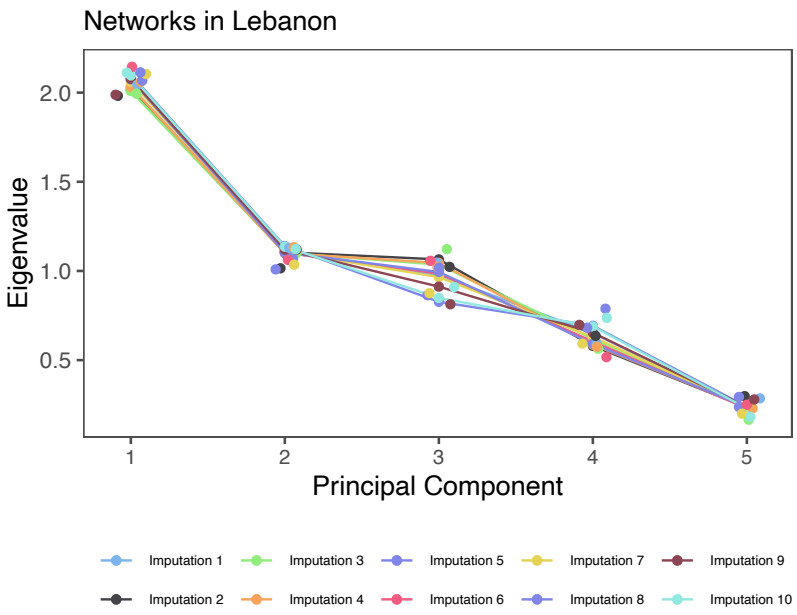

Figure A5: Screeplot for the networks in Lebanon index

4.2 Pull factors in Syria

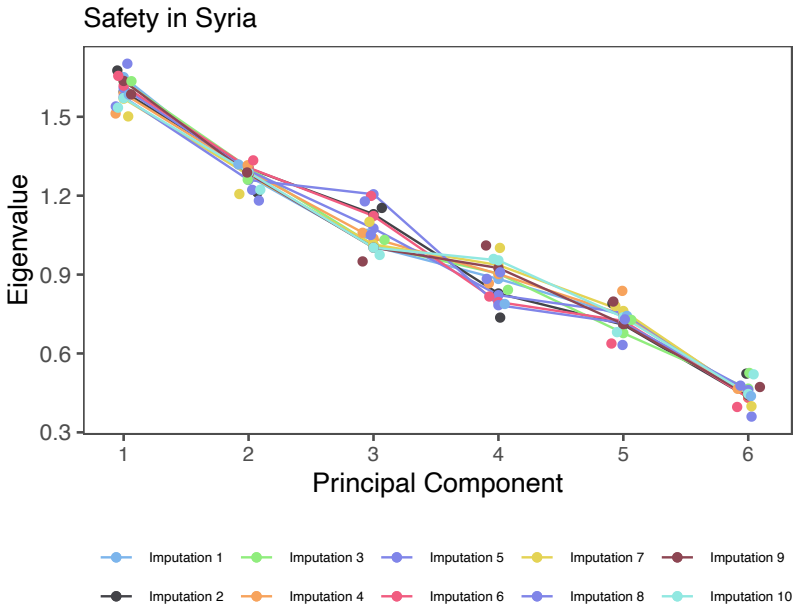

Figure A6: Screeplot for the safety in Syria index

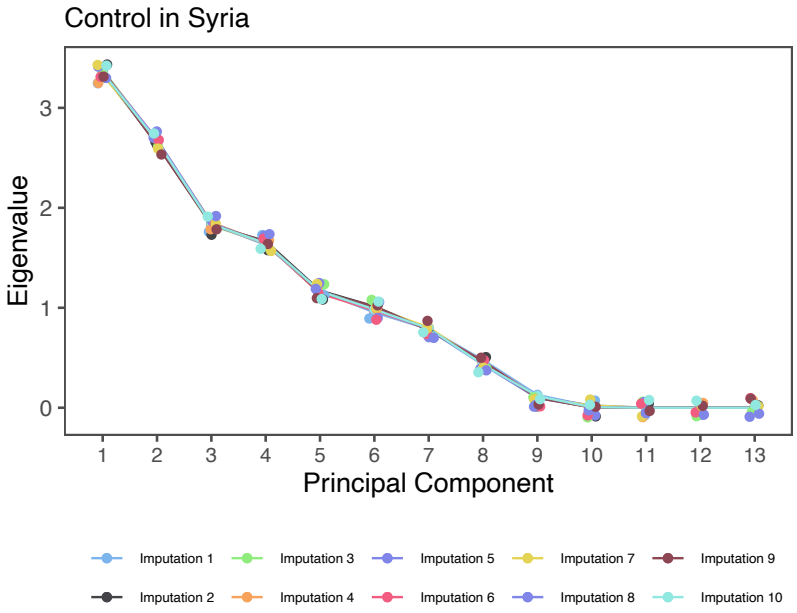

Figure A7: Screeplot for the control in Syria index

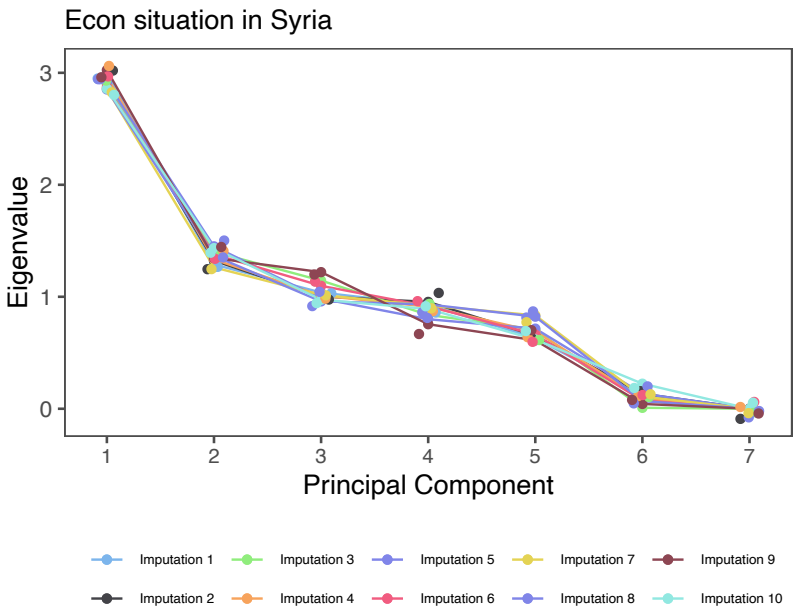

Figure A8: Screeplot for the economic well-being in Syria index

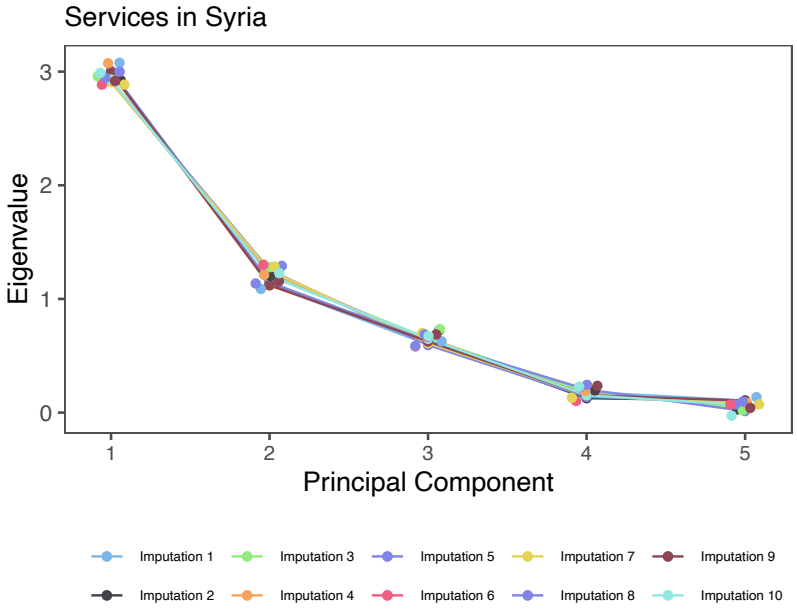

Figure A9: Screeplot for the services in Syria index

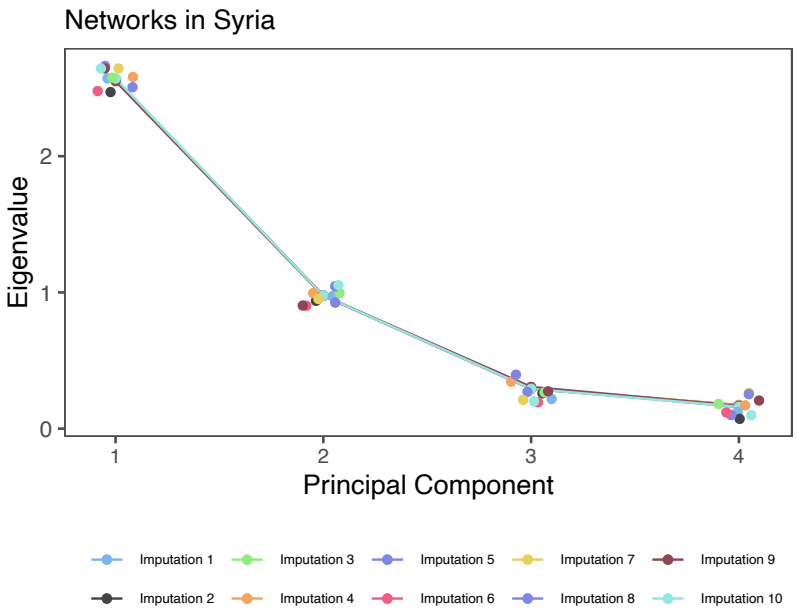

Figure A10: Screeplot for the networks in Syria index

4.3 Confidence in information

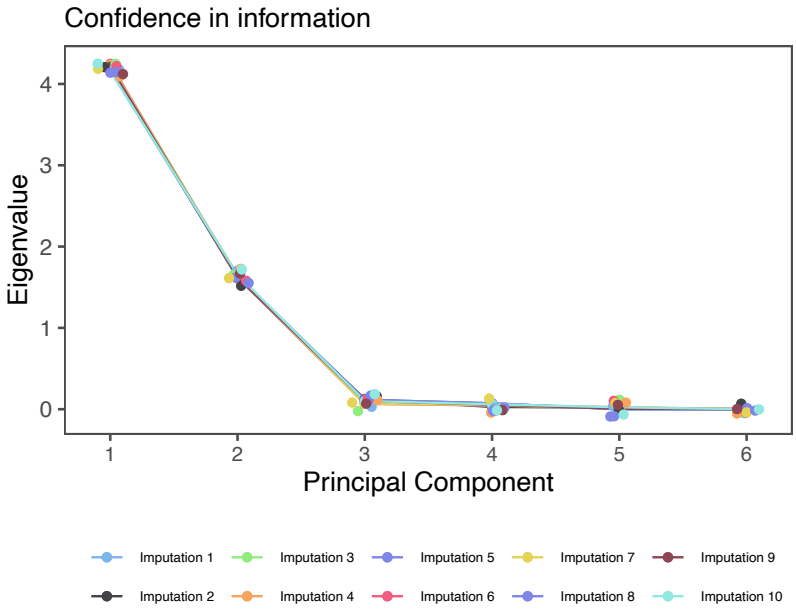

Figure A11: Screeplot for the confidence in information index

4.4 Preparation for return

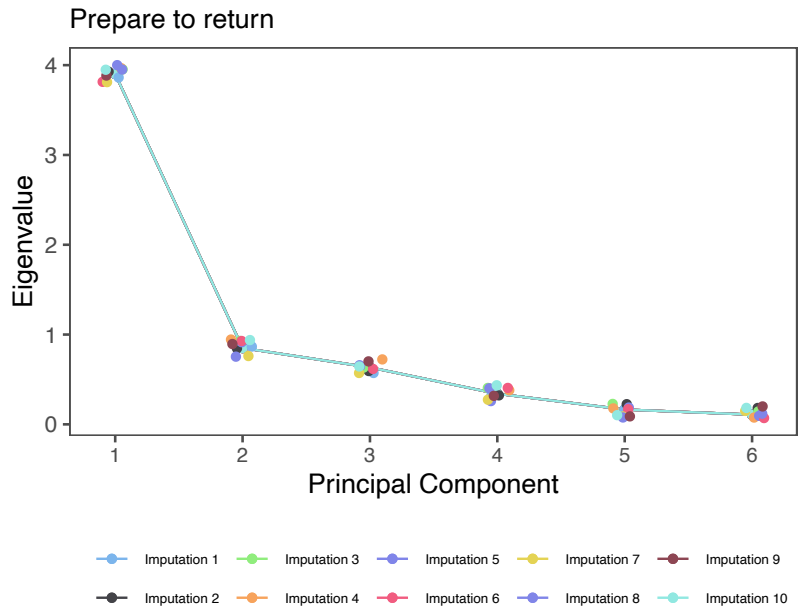

Figure A12: Screeplot for the preparation to return index

5 Deviations from PAP

Multicollinearity: In the PAP, we indicated that we would run one regression with all the indices. Because of multicollinearity, we ran separate regressions for each of the indices and also ran the pre-specified regression with a minor adjustment: we did not include the indices measuring safety and services in Syria in the same regression because of their high correlation. In the paper, we present the results results for all the indices where we control for safety. In Section 6.2, we include the results for all the indices where we control for services in Syria.

We made the choice to present an alternative regression after tests of model performance identified multicollinearity in our regression predictors. The first simple test of model performance involved examining the simple pairwise correlations between our predictors. We find that the pairwise correlation between the security index and services index in Syria is high at 0.57.

Of course, simple correlation is not multicollinearity. Therefore, second, we test for an association between predictors conditional on the other variables in the model. The variance inflation factor is a measure to analyze the magnitude of multicollinearity of model terms. Using the `performance()` package in R, we find that in the pre-specified models we identify very high variance inflation factors for a number of indices, most notably services in Syria and regime control. Evidence of multicollinearity is consistent across multiple tests in the `performance()` package.

Analysis without locality fixed effects: We re-ran our main analysis without Lebanese locality fixed effects as a robustness check. This regression can be seen in Figure A24 in Section 6 of the appendix. Furthermore, after submitting the PAP, we realized that it would be misguided to control for Lebanese locality and Syrian locality when analyzing the role of travel distance from Lebanese locality to Syrian locality. Therefore, in regressions with travel distance on the right-hand side, we do not include locality fixed effects.

Predictive analysis: The predictive analyses included in appendix section 6.7 were not pre-specified.

PCA inputs: PCA inputs were pre-specified but required a number of ex post modifications for reasons explained below.

Index 1—Economic well being in Lebanon: The PAP mistakenly indicated that a question about someone's former job in Syria would be included in the economic well-being *in Lebanon* index. This was a typo, and it was removed since it is not a dimension of economic well-being *in Lebanon*.

Index 2—Social well being in Lebanon: The PAP specifies that we intended to calculate the IPL-12 integration score to use it as an input for PCA. In later revisions to the questionnaire, some questions were modified and no longer matched IPL-12 inputs. We modified this index slightly and now use component questions of IPL-12 as inputs rather than calculating the IPL-12 score. Furthermore, household income in Lebanon was mistakenly included in both index 1 and index 2. We decided this index fits better in index 1 and removed it from index 2.

Index 3—services in Lebanon: A question included in the PAP was subsequently cut from the survey: "Have you been forced to move in the last two years, for instance because you were kicked out of your home or your home was deconstructed/demolished?" So we did not include it in the analysis.

Index 5.1—Safety: First, we separated regime control from safety conditions in Syria. These two concepts were not closely related and we decided to examine territorial control and safety separately. Furthermore, Family deaths in Syria was excluded due to flaw in measurement strategy.

Index 10—Information quality: After submitting the PAP we removed the following question: "How confident are you in your knowledge about conditions in [Piped place of origin]?" so we did not include it in the analysis.

6 Robustness and Additional Tests

6.1 Return in two years and ever

In addition to the return in 12 months and preparation to return outcomes, we have also examined the predictors of expectation to return in two years and intentions to ever return to Syria. In Table A22 the first two models present the regression results for the return ever outcome using one index per regression (Model 1) and using all indices in the same regression (Model 2). Models 3 and 4 show the same analyses for the return in two years outcome.

	Ever (Ind. Indices)	Ever (All Indices)	2 years (Ind. Indices)	2 years (All Indices)
Safety (Syr.)	0.040**	0.042***	0.049**	0.075***
	(0.016)	(0.014)	(0.018)	(0.016)
Regime control (Syr.)	0.004	−0.014	0.006	−0.020
	(0.015)	(0.013)	(0.014)	(0.018)
Economic well-being (Syr.)	0.033**	0.061***	0.026	0.061***
	(0.013)	(0.013)	(0.021)	(0.016)
Services (Syr.)	0.030	0.017	0.023	0.052***
	(0.017)	(0.015)	(0.015)	(0.013)
Networks (Syr.)	0.031***	0.020*	−0.002	−0.003
	(0.009)	(0.010)	(0.009)	(0.009)
Economic well-being (Leb.)	−0.040*	−0.077***	−0.019	−0.053
	(0.020)	(0.020)	(0.015)	(0.034)
Services (Leb.)	−0.038*	−0.008	0.046***	0.033*
	(0.022)	(0.024)	(0.016)	(0.019)
Networks (Leb.)	−0.017	−0.052***	−0.015	−0.007
	(0.012)	(0.014)	(0.018)	(0.012)
Social well-being (Leb.)	0.064***	0.113***	0.004	0.030*
	(0.022)	(0.022)	(0.020)	(0.017)
Legal conditions (Leb.)	0.001	0.011	−0.019*	−0.012
	(0.018)	(0.016)	(0.010)	(0.011)
Log travel distance	−0.078***	−0.055**	0.002	0.022
	(0.024)	(0.022)	(0.038)	(0.029)
Log household size	0.015	0.013	0.009	0.016
	(0.020)	(0.020)	(0.023)	(0.028)
Confidence in information	0.026	0.037**	0.020	−0.015
	(0.016)	(0.017)	(0.015)	(0.016)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A22: Regression results using alternative outcomes. The first two models present the regression results for the return ever outcome using one index per regression (model 1) and using all indices in the same regression (model 2). The last two models present the regression results for the return in 2 years outcome using one index per regression (model 3) and using all indices in the same regression (model 4).

6.2 Safety, services, and additional covariates

Given the high correlation between safety and services in Syria, we ran a regression that included all the predictors except services in Syria, and then ran a separate regression that included all the predictors except safety in Syria. We reported the results of all the coefficients from the regression that included safety (but not services) in Syria. We then added the single coefficient for services from the second regression. In Models 1 and 2 of Table A23, we do the opposite. We include all the coefficients from the regression that included services (but not safety) in Syria and then add the single coefficient for safety in Syria from the first regression.

In models 3 and 4 of Table A23, we present the results using all the indices in the regression (including safety and services in Syria) for the return in 12 months outcome (Model 3) and the preparation to return outcome (Model 4). As can be seen here, the two indices (safety and services) get smaller point estimates and are no longer statistically significant. This difference between the individual-index models and the all-indices model aligns with the evidence of correlation and multicollinearity shown in section 5.

Finally, the main results using individual indices control for a limited set of household covariates because many of the other potential covariates are included in the appendix. Of course, when we include the full list of indices in the same regression (all indices), we control for indices that include a variety of other covariates. In addition to this approach, we also run the main model with the individual indices using a larger set of controls. In addition to the main variables included in the paper, model 5 in Table A23 controls for household income, whether household receives aid, the number of years in Lebanon, the presence of a curfew targeting refugees in the area, whether any household members are school age, and having Lebanese relatives. Note that when one of these variables is included in the index, we do not control for this variable. The full set of variables in each index can be seen in section 3.

6.3 Excluding locality fixed effects

We also test for the robustness of our results by re-running our models, but without controlling for locality fixed effects.[1] Model 1 of Table A24 shows the results for the 12 months outcome while Model 2 shows the results for the preparation to return outcome.

6.4 Return plans for household members and uncertainty about return

The main paper coded intention to return in 12 months as 1 if heads of households indicated their intentions to return. The last two models in Table A24 present alternative codings of this outcome. The first two models re-run the main analysis without fixed effects for the return in 12 months outcome (Model 1) and for the preparation to return outcome (Model 2). In Model 3, the outcome is whether anyone in the household plans to return in 12 months. In Model 4, the outcome at the head of household level is coded as 1 if the heads of household said they planned to return or were uncertain about return and 0 only if they said that they do not plan to return.

[1] This robustness check was not pre-specified.

	12 months (Services)	Prepare (Services)	12 months (Safety + Services)	Prepare (Safety + Services)	Return in 12 months
Economic well-being (Leb.)	−0.007	0.154***	−0.007	0.155***	0.000
	(0.008)	(0.055)	(0.008)	(0.055)	(0.010)
Social well-being (Leb.)	0.021***	0.020	0.020***	0.018	0.017***
	(0.006)	(0.025)	(0.006)	(0.026)	(0.006)
Services (Leb.)	−0.007	−0.011	−0.008	−0.013	−0.006
	(0.008)	(0.029)	(0.008)	(0.030)	(0.007)
Legal conditions (Leb.)	−0.003	0.000	−0.002	0.001	−0.004
	(0.005)	(0.025)	(0.005)	(0.025)	(0.006)
Regime control (Syr.)	−0.007	−0.000	−0.009	−0.003	−0.011
	(0.006)	(0.018)	(0.006)	(0.018)	(0.007)
Economic well-being (Syr.)	0.012**	0.013	0.013**	0.014	0.016**
	(0.005)	(0.022)	(0.005)	(0.023)	(0.006)
Services (Syr.)	0.012**	0.073***	0.006	0.064**	0.018***
	(0.005)	(0.025)	(0.006)	(0.026)	(0.006)
Networks (Syr.)	0.012*	0.077*	0.012*	0.076	0.017***
	(0.006)	(0.045)	(0.006)	(0.045)	(0.006)
Networks (Leb.)	−0.002	0.018			
	(0.006)	(0.024)			
Confidence in information	0.021**	0.004	0.020**	0.003	0.019**
	(0.008)	(0.022)	(0.007)	(0.022)	(0.009)
Log travel distance	0.006	−0.037	0.006	−0.036	0.002
	(0.009)	(0.044)	(0.009)	(0.044)	(0.008)
Log household size	0.013	0.000	0.014	0.002	0.010
	(0.010)	(0.049)	(0.010)	(0.049)	(0.011)
Safety (Syr.)	0.014**	0.050*	0.011	0.017	0.016**
	(0.006)	(0.027)	(0.007)	(0.027)	(0.007)
Networks (Leb.)			−0.003	0.016	0.010*
			(0.006)	(0.024)	(0.006)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A23: Robustness tests. The first two models include the services in Syria index in the all-indices regression (instead of the safety index) for the return in 12 months outcome (model 1) and the preparation to return outcome (model 2). The last two models present the analysis that includes running both the safety and services in Syria indices in the same regression. Note that there is a high correlation between the safety and services in Syria indices. The third model uses the return in 12 months outcome and the fourth model uses the preparation to return outcome. The final model uses the main return in 12 months outcome with individual indices but uses additional covariates. Besides the ones mentioned in the paper, it controls for household income, an indicator for receiving aid, years in Lebanon, an indicator for a curfew in the locality, an indicator for age-school children, and an indicator for Lebanese relatives. Note that these additional controls are included only when the individual index in the regression does not contain any of these variables. The list of variables in each index is included appendix section 3.

6.5 Additive Indices

To build indices in the paper, we use the first principal component for the predictors under push factors, pull factors, and information as well as for the preparation to return outcome. In Table A25, we present results using mean effects indices. The first two models in Table A25 present the results using one index per regression for the 12 months return outcome (Model 1) and the preparation to return outcome (Model 2). The last two models present the same analyses but using all the indices in the same regression.

	12 months (No FEs)	Prepare (No FEs)	Household member return	Head of HH return (counting uncertain)
Safety (Syr.)	0.020***	0.084***	0.018**	0.031
	(0.006)	(0.024)	(0.008)	(0.018)
Regime control (Syr.)	−0.013*	−0.002	−0.010	−0.042*
	(0.007)	(0.021)	(0.007)	(0.021)
Economic well-being (Syr.)	0.018***	0.063**	0.020***	−0.001
	(0.005)	(0.030)	(0.007)	(0.011)
Services (Syr.)	0.021***	0.114***	0.019***	0.046**
	(0.005)	(0.028)	(0.006)	(0.017)
Networks (Syr.)	0.019***	0.110**	0.015**	0.011
	(0.006)	(0.040)	(0.006)	(0.009)
Economic well-being (Leb.)	0.007	0.183***	−0.001	0.044
	(0.007)	(0.054)	(0.008)	(0.026)
Services (Leb.)	−0.003	0.010	−0.012	−0.002
	(0.008)	(0.027)	(0.008)	(0.013)
Networks (Leb.)	0.013***	0.088***		
	(0.004)	(0.021)		
Social well-being (Leb.)	0.021***	0.067***	0.021***	−0.004
	(0.005)	(0.019)	(0.007)	(0.013)
Legal conditions (Leb.)	−0.003	0.002	−0.003	−0.012
	(0.006)	(0.020)	(0.005)	(0.011)
Log travel distance	0.002	−0.079*	0.007	0.019
	(0.008)	(0.043)	(0.012)	(0.015)
Log household size	0.001	−0.103***	0.022	0.012
	(0.009)	(0.034)	(0.013)	(0.029)
Confidence in information	0.022***	0.050*	0.017**	−0.015
	(0.008)	(0.027)	(0.007)	(0.015)
Networks (Leb.)			−0.001	0.024**
			(0.006)	(0.010)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A24: Additional tests. The first two models re-run the main analysis without fixed effects for the return in 12 months outcome (model 1) and for the preparation to return outcome (model 2). The third models uses an alternative dependent variable: household member return intentions. This outcome is coded as 1 if any household member plans to return and 0 otherwise. The final model uses another outcome. The main return intentions in 12 month is coded as 1 if the respondent indicates their plan to return in 12 months and 0 otherwise. The last model uses the return in 12 months outcome but it is coded as 1 if the respondent indicates either that they plan to return in 12 months or they are uncertain about returning in 12 months and 0 otherwise.

6.6 Information and Return Intentions

To further explore the relationship between information and return, we examine whether information moderates the role of perceived conditions in Syria. Specifically, we examine whether conditions in Syria have a larger effect on people's intentions when they have high levels of confidence in their information about the situation in Syria.

$$Y_i = \alpha + \beta_1 T_i + \beta_2 \left(T_i \times \mathbb{1}(I_i > 0)\right) + \gamma X_i + \epsilon_i \tag{1}$$

Equation 1 is similar to the "individual indices" specification of Equation 1 in the main paper, but includes a multiplicative interaction term between each index T and confidence in information. The indicator function, $\mathbb{1}(I_i > 0)$, denotes whether a respondent i had an index value for infor-

	12 months (Ind. Indices)	Prepare (Ind. Indices)	12 months (All Indices)	Prepare (All Indices)
Safety (Syr.)	0.012**	0.073*	0.009	0.038
	(0.006)	(0.039)	(0.006)	(0.037)
Economic well-being (Syr.)	0.021***	0.075*	0.015**	0.021
	(0.006)	(0.037)	(0.006)	(0.024)
Services (Syr.)	0.018***	0.087***	0.011**	0.053*
	(0.006)	(0.030)	(0.005)	(0.027)
Networks (Syr.)	0.016**	0.094**	0.012**	0.084**
	(0.006)	(0.039)	(0.005)	(0.040)
Economic well-being (Leb.)	0.008	0.186**	0.002	0.204**
	(0.008)	(0.074)	(0.008)	(0.082)
Services (Leb.)	−0.010*	0.026	−0.012**	−0.020
	(0.005)	(0.020)	(0.006)	(0.018)
Networks (Leb.)	0.008	0.020	0.003	−0.014
	(0.005)	(0.019)	(0.006)	(0.027)
Social well-being (Leb.)	0.001	0.026	0.008	−0.004
	(0.005)	(0.022)	(0.006)	(0.028)
Legal conditions (Leb.)	−0.003	0.004	−0.003	−0.013
	(0.006)	(0.023)	(0.006)	(0.025)
Log travel distance	0.002	−0.079*	0.007	−0.022
	(0.008)	(0.043)	(0.008)	(0.037)
Log household size	0.006	−0.066*	0.010	−0.098**
	(0.009)	(0.033)	(0.012)	(0.045)
Confidence in information	0.022**	0.055*	0.018**	0.016
	(0.008)	(0.027)	(0.008)	(0.025)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table A25: Additive indices. While the results in the paper use the first principal component of independent variables, here we construct the independent variables using additive indices. First, we standardize each variable to have zero mean and unit standard deviation. Then we summed these variables and standardized the sum again using the same method. The first two models present the results using one index in each regression for the return in 12 months outcome (model 1) and the preparation to return outcome (model 2). The last two models present the results using all the additive indices in one regression for the return in 12 months outcome (model 3) and the preparation to return outcome (model 4).

mation confidence above the mean. X denotes the same vector of covariates as the main paper's Equation 1. Figure A13 presents regression results, displaying the estimated marginal effect of a one standard deviation change in each index for people with low (below-average) confidence in information compared to high (above-average) confidence in information. To obtain standard errors, we used bootstrapping. The confidence intervals here represent the 97.5th (95th) and the 2.5th (5th) percentiles of coefficients across all bootstraps and imputations.

The results in Figure A13 suggest that the relationship between conditions in Syria and return intentions and preparations is shaped by respondents' confidence in their information sources for some key factors. Specifically, we find evidence that information is a significant moderator for the role of regime control and economic prospects in shaping return intentions. Next, we see a differential relationship between the availability of services in people's hometowns and both return intentions and return preparations. Last, we see a differential relationship between networks in Syria and return preparations depending on information confidence.

An alternative way of obtaining confidence intervals is by finding the empirical variance of

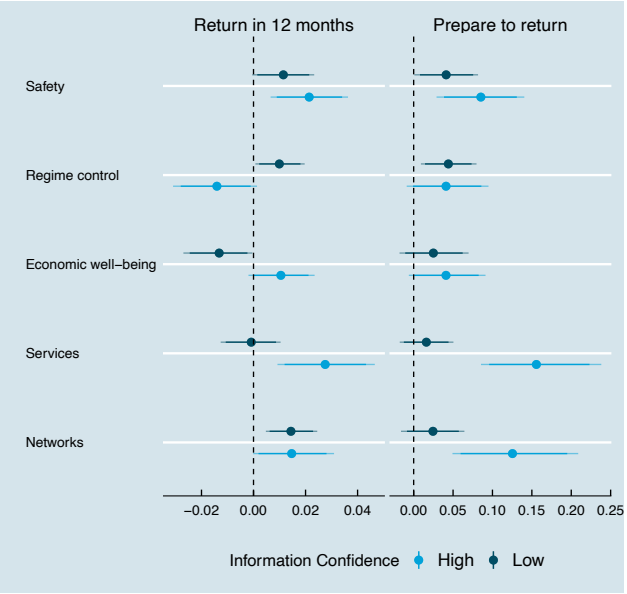

Figure A13: Interactive Effects of Information Confidence on Intentions and Preparations. Estimated marginal effects are presented with the corresponding 95% confidence intervals, with standard errors clustered by locality in Lebanon.

coefficients across bootstraps within each imputation then finding the pooled variance across all imputations using the rules of Rubin (1987). Figure A14 presents results from this approach.

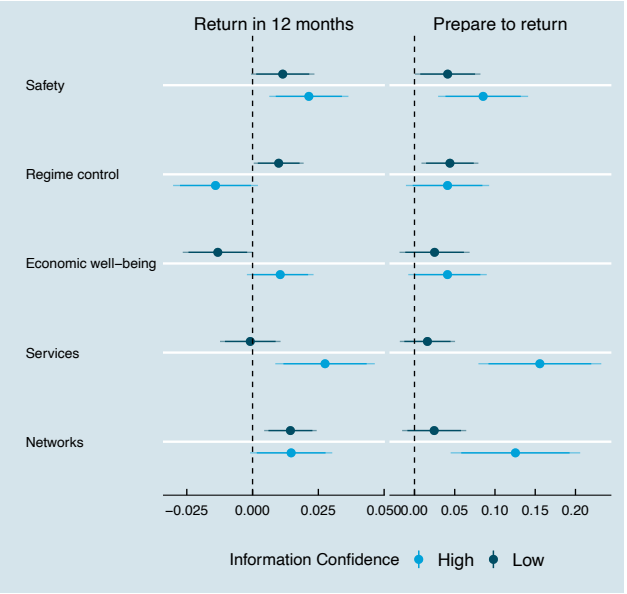

Figure A14: Interactive effects of information on intentions and preparations to return

6.7 Testing the models' predictive power

Below we present prediction plots, including OLS (same models as in the PAP) and lasso, with AUC results for ROC and PR. PR is often as a better performance metric than ROC for predicting rare outcomes. This is due to the fact that the ROC allows for relatively "good" performance by predicting all zeros, which is not the case with PR plots.

Looking at the results in Figures A15–A16, we see that the trends are consistent with our main findings but the differences across models are small. Looking at the ROC plot, we witness a ∼2.5 percentage point increase (∼5%) comparing the push model to the pull model, and a ∼6 percentage point increase (∼9%) moving from push to full (i.e., push + pull).[2] Looking at the PR curves, we see a ∼1.5 percentage point increase (∼17%) comparing the push model to the pull model, and a ∼2.5 percentage point increase (∼29%) moving from the push model to the full model (i.e., push + pull). The gains in terms of PR AUC are large in percentage terms, although not absolute terms.

The results suggest a few key takeaways about the predictive power of the models presented in the paper. First, the Syria model is a better predictor than the Lebanon model, aligning with our main results. Second, the Syria+Lebanon model is the best predictor, suggesting that push factors are relevant, just less than pull factors. Third, the gains across models are large in percentage terms, but not in absolute terms. Lastly, overall predictive power is low and prediction is hard in our case, possibly due to studying a rare outcome, but also highlighting that understanding the aggregate drivers of return will not necessarily allow policymakers to make reliable predictions about whether an individual household will return or not.

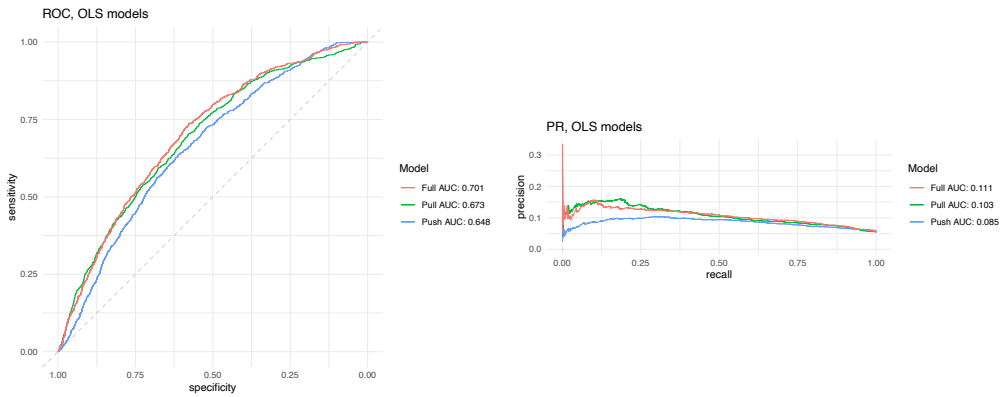

Figure A15: ROC—OLS models with pre-specified inputs of indices as predictors (left) and Principal-response curve—OLS models with pre-specified inputs of indices as predictors (right)

[2] We indicate that these differences are approximate since their precise magnitudes will vary across different simulations.

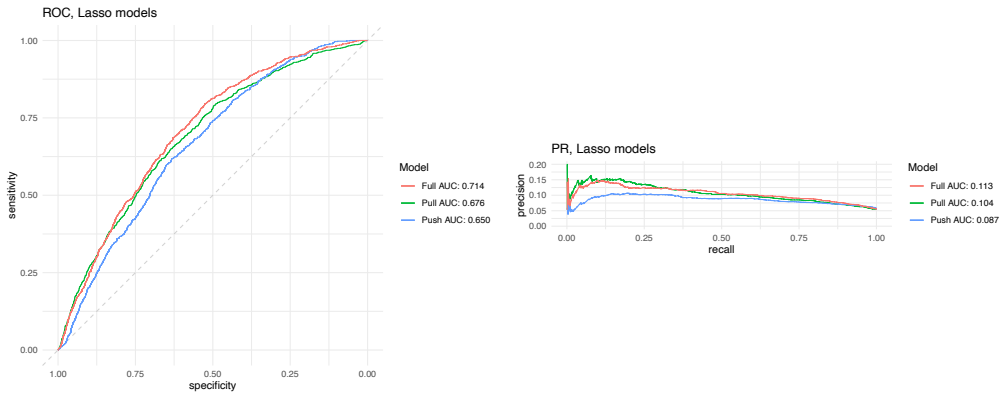

Figure A16: ROC—Lasso models with pre-specified inputs of indices as predictors (left) and Principal-response curve—Lasso models with pre-specified inputs of indices as predictors (right)

6.8 Map of conditions in Lebanon

This section further demonstrates the variation in push factors within Lebanon. Using all variables use to measure push factors in Lebanon, we construct a single index for conditions in Lebanon (extracting the first principal component from PCA). Figure A17 shows the weighted average of this index by district. It should be noted that this index was constructed to have zero mean and unit standard deviations, so the variation can be measured in terms of standard deviations. The map in Figure A17 shows that push factors in Lebanon varied substantially across districts. The difference between the district with the worst conditions for respondents (El Minieh-Dennie) and the district with the best conditions for respondents (El-Nabatieh) is about 2.4 standard deviations. This map suggests that variation across districts was associated with a significant change in push factors in our survey. It should be noted that disaggregating results from a nationally representative survey does not produce estimates that are representative of the subnational units (here, districts). Nonetheless, Figure A17 clearly demonstrates that respondents in our sample exhibit significant variation in push factors across Lebanon's different districts.

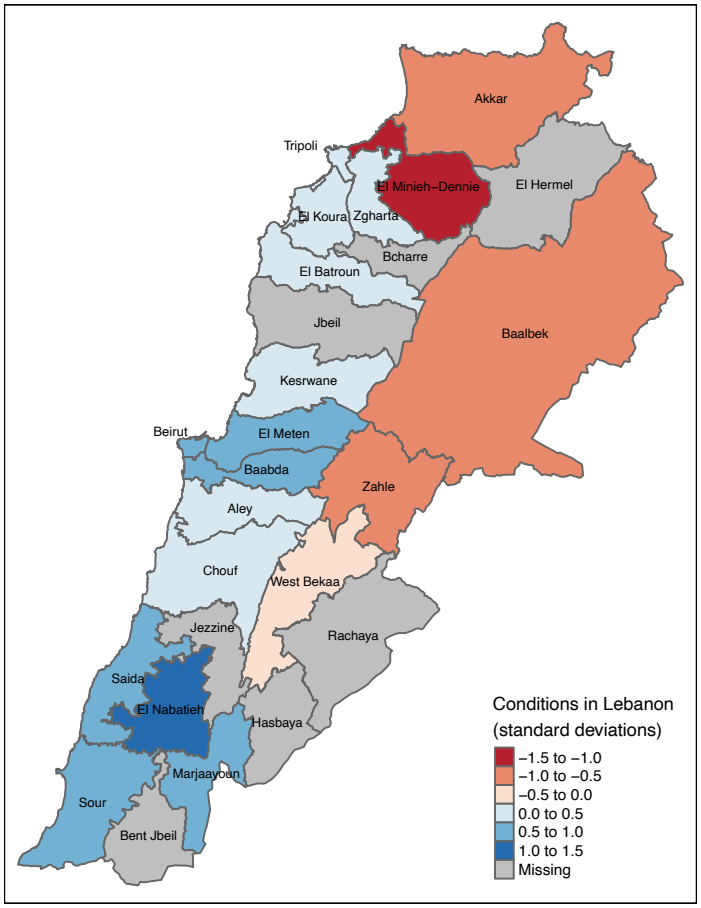

Figure A17: Map of respondents' conditions in Lebanon.

7 Jordan survey

In this section, we present the set of questions we used to construct each index from the Jordan survey data. As with the Lebanon indices, we constructed these indices by extracting the first components from PCA of the input variables. Some of the questions differ from the wording used in Lebanon in order to fit the Jordan context. Due to space constraints in the Jordan survey, the survey did not contain the full set of questions used in Lebanon.

8 Ethical Considerations

The ethical imperative to do no harm is especially pressing in research with refugees, given their extreme vulnerability (Masterson and Mourad, 2019). We designed this project to reduce potential harm, maximize policy relevance, and increase opportunities for direct benefits to research partici-

Safety in Syria	**Economic well-being in Syria**	**Services in Syria**	Networks in Syria
How would you describe the risk to civilians physical safety (such as fighting, kidnapping, IEDs, crimes) in your place of origin?	How would you describe the availability of jobs at present in your place of origin?	As far as you know, how many hours per day is there electricity in your place of origin?	Approximately how many of your relatives or friends in Jordan have gone back to Syria?
What were the main reasons for you to leave your home country Syria? (Violence/bombardment selected)		As far as you know, how many hours per day is there running water in your place of origin?	
		As far as you know, are schools operating in your place of origin?	
		As far as you know, are health centers operating in your place of origin?	

Table A26: Jordan questionnaire: Pull factors in Syria

pants. The achieve the first two goals, the authors drew on exploratory fieldwork, interviews with international and local humanitarian actors, and a research planning workshop with the humanitarian community in Beirut during which we discussed our research plan and questionnaire to minimize potential harm and ensure that the design can provide the humanitarian community with required evidence to fulfill their needs. To address the first and the third goals, the research team partnered with NGOs to provide protection training to enumerators and established a referral mechanism through which research participants in need of humanitarian services were connected to available resources. Below we discuss some of the measures we have taken.

Compensation

Survey interviews took about 30-40 minutes to complete and we provided survey respondents with $10 cash compensation for their time. We decided to compensate respondents following extensive interviews with humanitarian actors, as many humanitarian actors suggested that it is only fair to compensate participants for their time following a somewhat long survey.

Engagement with humanitarian actors

To strengthen the research contribution, we closely consulted with humanitarian actors at all stages of the project. At the outset, we worked with humanitarian actors to conduct focus groups and meetings with refugees in Lebanon and Jordan. We consulted with humanitarian actors closely on the development of the questionnaire. Prior to data collection, we conducted a workshop with humanitarian actors in Beirut, Lebanon, in March 2019 to present the whole project and get feedback, including on the questionnaire. We then revised the questionnaire based on comments (for instance, we removed questions that directly measured political attitudes, which they suggested were too sen-

Economic well-being in Jordan	Services in Jordan	Networks in Jordan	Social well-being in Jordan	Legal situation in Jordan
If you think back about the situation one year ago in terms of access to assistance, has the situation improved, stayed the same, or deteriorated?	On 24 January [2019] the Prime Minister's office signed a decree informing that Syrian refugees in MOH hospitals and health centers will be requested to pay directly 80% of the applicable 'foreigners rate,' whereas from November 2014, they were treated like Jordanians who did not have health insurance and were able to access health services at subsidized rates. Has it impacted your ability to access health services?	In the last 12 months, how often did you share a meal with Jordanians who are not part of your family?	Does anyone in your family face verbal or physical harassment, meaning verbal or other actions meant to annoy, threaten, intimidate, or make someone feel scared for their safety, in the area around your house?	Do all your family members have a valid registration with UNHCR?
Have you received assistance from [government organizations, local organizations, NGOs, UN in the last two months]?	Number of school aged children out of school	Please think about the Jordanians in your phone contacts. With how many of them did you have a conversation–either by phone, messenger chat, face-to-face, or text exchange–in the last week?	Would you describe the relations with your neighbors as mostly positive/neither positive nor negative/mostly negative?	Do you have a government service card, currently called "MOI Card"?
Are your currently working?		Please think about the Syrians in Jordan in your phone contacts. With how many of them did you have a conversation–either by phone, messenger chat, face-to-face, or text exchange–in the last week?	What is your level of education?	
Is any member of your household currently working?			When did the first member of your family arrive to Jordan?	
Do you or any of your household have a work permit?			How connected do you feel with Jordanian society?	
Have you received food vouchers during the last month?			How often do you feel like an outsider in Jordan?	
			Have you received help from your neighbors?	

Table A27: Jordan questionnaire: Push factors from Jordan

Outcome	Confidence in information
Do you think it will ever be possible to return to your place of origin in Syria?	What do you have information about [in Syria]: Respondent selects all that applies. 1 if selected: (a) Safety/security news (b) Status of infrastructure in a particular location (c) News about friends/family in Syria
	How confident are you in your knowledge about conditions in your place of origin?

Table A28: Jordan questionnaire: Outcome and confidence in information

sitive for the context). In June 2020, we conducted several online workshops with humanitarian actors in Lebanon to share results and answer questions from humanitarian actors.

Referrals

When we presented the research design and questionnaire to humanitarian organizations in March 2019, they suggested developing a referral strategy for refugees who require or ask for help. The goal of referrals is to help facilitate refugees' access to services by either (1) putting individuals in need of services directly in contact with the service providers or (2) enabling people to seek assistance and support them in receiving assistance. Enumerators and team leaders who participated in data collection received special training from a humanitarian organization in Lebanon on the goals of referrals, when to refer respondents, and how to refer respondents. In addition to the training, enumerators received guidelines that were built using material that is used by humanitarian organizations. Humanitarian organizations in Lebanon collaborate on a centralized referral platform

called Referral Information Management System (RIMS), in addition to traditional referral methods (by reaching out directly to the responsible organization). As RIMS was not available for use by researchers (only humanitarian actors could use it), the research team collaborated with a local humanitarian organization to conduct the referrals and compensated them for the time they spent carrying out needed referrals for this project's research participants.

Figure A18 summarizes the referral process and was included in the enumerator guide:

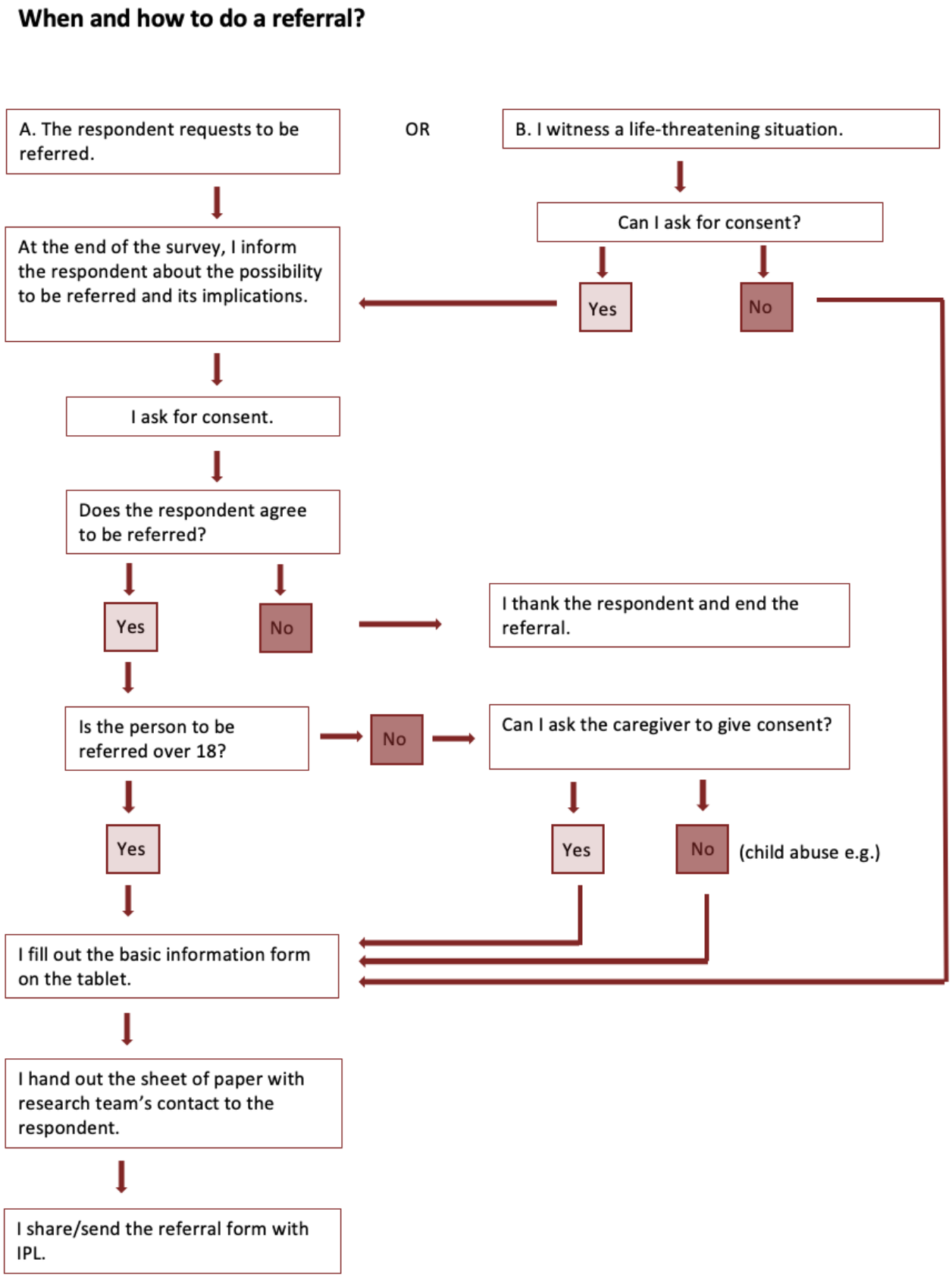

Figure A18: Referral guide summary

References

Abadie, Alberto, Susan Athey, Guido W Imbens and Jeffrey Wooldridge. 2017. "When should you adjust standard errors for clustering?".

Hainmueller, Jens, Dominik Hangartner and Giuseppe Pietrantuono. 2017. "Catalyst or crown: Does naturalization promote the long-term social integration of immigrants?" *American Political Science Review* 111(2):256–276.

Kling, Jeffrey, Jeffrey Liebman and Lawrence Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75(1):83–119.

Masterson, Daniel and Lama Mourad. 2019. "The Ethical Challenges of Field Research in the Syrian Refugee Crisis." *APSA MENA Politics* 2(1).

Rubin, Donald B. 1987. *Multiple imputation for nonresponse in surveys.* John Wiley & Sons.