The Refugee's Dilemma: Evidence from Jewish Migration out of Nazi Germany

Johannes Buggle, Thierry Mayer, Seyhun Orcan Sakalli, Mathias Thoenig

The Refugee's Dilemma: Evidence from Jewish Migration out of Nazi Germany *

Johannes Buggle† Thierry Mayer‡ Seyhun Orcan Sakalli§ Mathias Thoenig¶

April 14, 2022

Abstract

We estimate the push and pull factors involved in the outmigration of Jews facing persecution in Nazi Germany from 1933 to 1941. Our empirical investigation makes use of a unique individual-level dataset that records the migration history of the Jewish community in Germany over the period. Our analysis highlights new channels, specific to violent contexts, through which social networks affect the decision to flee. We first estimate a structural model of migration where individuals base their own migration decision on the observation of persecution and migration among their peers. Identification rests on exogenous variations in local push and pull factors across peers who live in different cities of residence. Then we perform various experiments of counterfactual history to quantify how migration restrictions in destination countries affected the fate of Jews. For example, removing work restrictions for refugees in the recipient countries after the Nuremberg Laws (of 1935) would have led to an increase in Jewish migration out of Germany in the range of 12 to 20%, and a reduction in mortality due to prevented deportations in the range of 6 to 10%.

Keywords: Refugees, Migration Policy, Counterfactual History, Nazi Germany

JEL Codes: F22, N40, F50, D74

*Helpful comments from Alberto Bisin, Mirna Safi, Shanker Sattyanath, Philipp Schmidt-Dengler, Joachim Voth, Katia Zhuravskaya, Yanos Zylberberg, as well as from participants of seminars and conferences at ASREC 24-hour online conference, Bonn, Bristol, Cambridge, CEU, CfP HiCN Workshop in Göttingen, 14th Conference on Migration and Development in Luxembourg, EIEF, EHA conference in Atlanta, ENS Lyon, EPCS conference in Jerusalem, Fribourg, Geneva, Global Challenges Seminar, Graduate Institute, Hebrew University of Jerusalem, IDC Herzliya, Lausanne, Nova Lisbon, NYU, OECD/CEPII conference on "Immigration in OECD Countries", Paris Dauphine, PSE, Queen Mary, SIOE conference in Stockholm, ALUM-CEPR conference in Siracusa, SNEE conference 2021, Tinbergen, Political Economics Workshop at Zurich University, Zurich ETH, VfS conference in Freiburg, University of Vienna, and WU Vienna are gratefully acknowledged. We thank Elio Bolliger, Yannick Joller, Nathanael Moser, Finn Wendland, and Athanasia Zarkou for excellent research assistance. Johannes Buggle, Seyhun Orcan Sakalli, and Mathias Thoenig acknowledge financial support from the SNF-Grant "The Refugee's Dilemma: Uncertain Threat at Home or Costly Asylum Abroad? Evidence from Jewish Emigration in Nazi Germany" (182242). We kindly thank Dr. Tanja von Fransecky of the German Federal Archives and the members of the association "Tracing the Past e.v." for sharing their data with us. This paper is accompanied by an Online Appendix.

†Department of Economics, University of Vienna.

‡Department of Economics, Sciences Po, Paris; CEPII; and CEPR.

§King's Business School, King's College London.

¶Department of Economics, School of Business and Economics, University of Lausanne; and CEPR.

1 Introduction

Violence is a major driver of migration: In 2020, 82.4 million people were forcibly displaced because of wars, civil conflicts, mass-killings, and other forms of persecution (UNHCR, 2021). Yet, we do not have a precise understanding of the process behind the decision to emigrate at the individual level in a violent context. How do individuals factor in the threat to personal security that a conflict poses, and at which point are they ready to leave their home behind? How do migration policies in potential destination countries affect their survival prospects in their origin countries? Beyond case studies and anecdotes, these questions remain overlooked from both a causal and quantitative perspective. This lack of systematic evidence is worrying given the high policy relevance of the link between conflicts and migration as exemplified by the recent waves of refugees fleeing Syria and Ukraine.

This paper studies the push and pull factors that were involved in migration decisions of Jews facing persecution in Nazi Germany from 1933 to 1941 (the last year before migration was banned). By the end of 1938, more than two-thirds of the Jewish community was still in Germany despite years of persecution. This puzzling fact has attracted a lot of attention from historians who have contrasted two main explanations, namely (i) migration frictions and (ii) the underestimation of the actual threat by the Jewish community.[1] Our analysis aims at assessing quantitatively their relative contributions. Our natural premise, backed by many historical records, is that network effects and social interactions within the community played a pivotal role in the decision to flee by impacting both migration prospects and perceptions of the threat. Our empirical investigation makes use of a unique individual-level dataset that records the migration history of a sample intended to cover the universe of the Jewish population living in Germany over the period. We estimate a structural model of migration where individuals base their migration decision on the observation of persecution and migration among their peers. Identification rests on exogenous variations in *local* push and pull factors across peers who live in different cities of residence than the decision maker. Equipped with structural parameter estimates, we perform various experiments of counterfactual history to quantify how migration restrictions in destination countries affected the fate of Jews living in Germany.

Our modeling of social interactions features elements that, we believe, are inherent to contexts where violence is pervasive. First, we emphasize how social networks aggregate information on the extent of persecution and consequently shape outmigration incentives. We call this channel the *threat* effect. Second, we investigate how past migration of peers affect current outmigration incentives. Here we consider two different migration spillovers. The *exodus* effect acts as a network-driven push factor that operates, in particular when population displacement becomes massive. A shrinking social network in the origin country lowers the prospects of staying. The reasons are numerous and pertain to less frequent social interactions between group members, a fall in real wages (e.g., in-group business network), fewer in-group amenities (food, culture, etc.), the statistical targeting of the remaining group members, and the migration of peers signaling how seriously they factor in the threat. The *diaspora* effect is a network-driven pull factor that has been

[1] See for example Strauss (1980); Kaplan (1999); Nicosia and Scrase (2013).

extensively documented in the migration literature (McKenzie and Rapoport (2010) and Beine *et al.* (2011) for early contributions). An expanding social network in a destination country facilitates future migration to this destination by lowering frictions (job market search, housing, etc.).

Our empirical investigation makes use of rich information about the Jewish residents of Germany during 1933 to 1945. The dataset, known as the *Resident List*, was compiled by the German Federal Archives, on behalf of the Federal Government that, in 2004, commissioned a scientifically sound and complete list of all Jews that lived in pre-war Germany (see Zimmermann, 2013).[2] This dataset records biographic information, as well as a detailed migration or deportation history, including the timing of migration movements, the destination countries, and/or the deportation date and place. To the best of our knowledge, we are the first to use this dataset for a scientific quantitative study. Therefore, a first key objective of our study is to establish the Resident List as a reliable historical source, and to describe the characteristics of Jewish emigration based on this data. We then exploit the available information on individuals' city of birth to reconstruct (part of) their social network. Our assumption is that individuals of comparable age (± 5 years) and born in the same city are likely to know each other—a reasonable view given that the *"Gemeinden"* (communities) were the focal point of the Jewish social life at the local level (Maurer, 2005).[3] These communities were spread all across the German territory, were relatively small and spatially sorted even in big cities. We restrict our measure of the social network to the subset of *distant peers* (DP) only, namely individuals from the same age group, from the same city of birth but living in a different city. As a result, decision makers and their peers are exposed to different local push and pull factors of migration.

The identification of a causal impact of peer effects on violence-induced outmigration is challenging for several reasons. A first issue is that violence usually prevents the collection of exhaustive data. Particularly important, it is rare in those episodes to have data covering extensively both the migrants and the "stayers". Perhaps even more important, there are challenges related to the identification of peer effects. Measurement of social networks is notoriously hard and requires fine-grained information. In addition, peers often live in the same place and therefore experience the same unobserved conditions (e.g., localized violence and economic deprivation), leading to correlated effects. Our data and context provide a unique setting to tackle these issues. Indeed, the disaggregated nature of our dataset allows us to control for a large battery of fixed effects that absorb many unobserved correlated effects. Moreover, cross-city mobility of the German population was high after the collapse of the German Empire in 1918. Hence, when Hitler came into power, many peers, friends, and relatives were living in different cities of residence and were consequently exposed to different migration incentives at the local level. Our measure of social network exploits this fact; our focus on distant peers allows us to further filter out potential correlated effects.

[2] The Archives drew on more than 1,000 different sources (including emigration lists, membership lists of Jewish parishes, all German municipal archives, foreign archives, deportation lists and registers of concentration camps) over a period longer than a decade to trace emigration and deportation at the individual level for about 300 thousand Jews living in Germany in the 1930s.

[3] Jewish communities, known as "Jüdische Kultusgemeinde" or simply "Gemeinde" hereafter, were public corporations, collected taxes and organized local Jewish life by financing religious and secular institutions, such as synagogues and more than 5,000 Jewish associations (Gruner and Pearce, 2019).

Our main dataset offers rich information on the location of the inter-war Jewish community within Germany and their migration decisions. However, it lacks information on individuals' education, income, and wealth, which potentially affect migration prospects.[4] This prevents us from assessing the effect of financial constraints on migration decisions. Note that our causal analysis of networks effects exploits identifying variations that are likely to be uncorrelated with individual characteristics: Push and pull factors affecting peers living in different cities of residence. Furthermore, our results are barely affected by the inclusion of individual fixed effects, which capture time-invariant omitted individual characteristics, such as those related to economic status, political orientation, or religiosity. Another data limitation is the absence of information on the intensity of social ties. Therefore, we cannot investigate how strong and weak social ties differentially affect migration prospects.

We begin our empirical analysis with a reduced-form estimation of the determinants of outmigration decision. Our estimates show that both past detainment and migration of distant peers impact positively the individual-level likelihood of migrating. We interpret this finding as preliminary evidence of the threat effect and the (joint) influence of the migration spillovers (exodus and diaspora). In order to disentangle diaspora and exodus effects, we then proceed with a structural approach. We exploit the available information on choices of destination to discriminate between the two effects. Migration of peers to a given country increases future migration *only* to that country according to the diaspora effect, while it raises the odds of migrating to *all* destinations according to the exodus effect. To allow for an integrated framework of the outmigration and location decisions, we build a random utility model of migration with network spillovers and specify it as a nested logit (a standard setup for considering multi-stage discrete choices). The model allows for the estimation of migration decisions in two steps: A lower model explains destination choices and yields a gravity equation of city-to-country migration; and an upper model explains the decision to migrate out of Germany at the individual level. The structural estimation of destination choices reveals the underlying parameters driving the response to migration frictions, in addition to estimates of the "core attractiveness" of each destination country for every year between 1933 and 1941. Those estimates are then used to construct a theory-consistent measure of expected utility for each individual in the outmigration decision. Results from estimating the outmigration model show that peers' past migration impacts positively the likelihood of emigration. A one-standard-deviation increase in the past migration of network members, i.e., the exodus effect, increases the annual emigration probability by 1.1 percentage points (20% of the sample mean). Increasing the expected utility by one standard deviation, which encompasses diaspora effect together with destinations' attractiveness and migration frictions, increases the annual probability of emigration by 1.3 percentage points (24% of the sample mean).

Building on recent advances in the analysis of policy scenarios in trade (often referred to as

[4] It is likely that economic means affected the capacity and motivations of Jewish outmigration decisions in Nazi Germany. It is however also likely that the influence of income and wealth for instance were ambiguous. On the one hand, richer individuals could more easily cover the costs associated with travel and relocation. On the other hand, well-off households might have been more sensitive to the expropriative taxes imposed upon departure. A high level of education (associated with income and wealth) did not either necessarily translate into better migration prospects, since some specialized professional skills could less portable abroad (e.g., lawyers) than others (e.g., scientists, entrepreneurs).

Exact Hat Algebra), we use the model to conduct a number of counterfactual policy scenarios. The policies we simulate are motivated by actual historical circumstances (events and proposals discussed in the period). The simulation results show substantial effects of policies reducing migration frictions, especially when magnified by the social spillovers. For example, removing work restrictions for refugees in the destination countries after the Nuremberg Laws (of 1935) would have led to an increase in Jewish migration out of Germany in the range of 12 to 20%, and a reduction in mortality due to prevented deportations in the range of 6 to 10%. Moreover, we document that the diaspora and exodus effects are both at work with quantitatively close magnitudes. Finally, our quantifications indicate that migration frictions in the destination countries contributed more to the low rates of migration out of Nazi Germany than the underestimation of the actual threat by the Jewish community did.

The paper is structured as follows. Section 2 discusses the related literature. Section 3 presents a brief historical background and then describes and validates the data. In Section 4, we discuss the role and measurement of social networks. Section 5 provides a preliminary "reduced-form" analysis of the data and documents the threat effect. In Section 6, we build and estimate a structural model of outmigration. Section 7 displays the counterfactual exercises. Section 8 concludes.

2 Related Literature

Our paper contributes to the economic analysis of the determinants of migration, in particular the push factors of migration. When it comes to refugees and asylum seekers, the literature on forced migration under threat has shown that violent conflict and natural disasters are first-order push factors: Chin and Cortes (2015), and Becker and Ferrara (2019) provide excellent reviews. This literature mostly consists of aggregate studies and only few analyses have examined the connection between violence and migration at the individual level. The most prominent are Engel and Ibáñez (2007) and Ibáñez and Vélez (2008) who study the determinants of population displacement during conflict in Colombia using a household survey: They find that the threat of violence and the presence of paramilitary and guerilla groups were strongly associated with outmigration. In the same vein, Bohra-Mishra and Massey (2011) examine the effects of exposure to violence on individual decision to migrate during the Nepalese civil conflict. In all these papers, and contrary to ours, the authors observe only a small fraction of individuals affected by the conflict. Moreover, their main focus is on internal displacement in the context of developing countries, while we investigate international migration under threat in the 1930s Germany.

A small set of papers focuses on the persecution and migration of Jews. Spitzer (2015) investigates the effects of anti-Jewish mob violence in the Russian Empire and Jewish migration networks established in the U.S. on the migration of Jews to the U.S. between 1881 and 1914. In comparison, our study focuses on individual migration decisions of Jews during a period of persecution as opposed to post-persecution migration decisions aggregated at the district level. Blum and Rei (2018) look at Jewish emigration during the Holocaust for a sample of migrants that travelled from Lisbon to New York between 1940 and 1942. Becker *et al.* (2021) highlight a specific aspect of emigration from Nazi Germany: They focus on the role of professional networks in the emigration of Jewish academics. We instead look at migration decisions of the *entire* Jewish community and simulate counterfactual history experiments to assess how migration would have reacted to less restrictive refugee policies.

Our paper also relates to a literature on the economic determinants and impacts of persecution of Jews both in historical times and during the Holocaust.[5] Anderson *et al.* (2017) and Grosfeld *et al.* (2020) emphasize the importance of negative economic shocks in explaining the timing of violence against Jews in the Middle Ages and in the twentieth century, while Becker and Pascali (2019), Jedwab *et al.* (2019), and Grosfeld *et al.* (2020) identify occupational complementarities and competition between Jews and Gentiles as important determinants of historical violence against Jews in Europe. Voigtländer and Voth (2012) document a persistence in antisemitic attitudes between the Middle ages and Nazi Germany, whereas Adena *et al.* (2015) show that radio propaganda of Nazis increased antisemitism in Germany, more so in places with a greater level of historical antisemitic sentiment. Acemoglu *et al.* (2011), Grosfeld *et al.* (2013), Akbulut-Yuksel and Yuksel (2015) find a negative effect of the persecution of Jews during the Holocaust on economic development. Huber *et al.* (2021) document that the removal of Jewish managers had negative effects for the stock prices of firms, while Waldinger (2011, 2010) study how the expulsion of Jewish scientists affected university researchers and doctoral students in Germany. Moser *et al.* (2014) document a positive effect of Jewish immigrants from Nazi Germany on scientific output in the US.

We also contribute to the large literature that studies how social networks influence migration. The existing literature has focused on diaspora networks and chain migration as pull factors of migration (Munshi, 2003; McKenzie and Rapoport, 2010; Beine *et al.*, 2011; Comola and Mendola, 2015; Giulietti *et al.*, 2018). We document the important role played by another social network spillover, the exodus effect, which acts as a push factor that reduces the prospects of staying when relatives and peers migrate out.

Regarding methods, our work is most closely related to the recent stream of papers building quantitative spatial economics models as surveyed by Redding and Rossi-Hansberg (2017) and Redding (2020). In particular, we use the same micro-foundations for estimating migration gravity equations that explain the integrated (nested) decisions to (i) migrate out of Germany, and (ii) which country to settle in, conditional on emigration. The estimation procedure enables to retrieve two key structural parameters driving the response of migrants to migration costs at both levels of the decision tree. Equipped with these critical elasticities, and in line with the approaches used in this branch of the literature (Monte *et al.*, 2018; Fajgelbaum *et al.*, 2018; Bryan and Morten, 2019;

[5]Besides, this study speaks to a large historical literature that has studied Jewish persecution and emigration in Nazi Germany, and the reactions and policy responses in receiving countries. On the description of Jewish life in pre-war Germany see for example Kaplan (1999); Maurer (2005); Matthäus and Roseman (2010); Nicosia and Scrase (2013). On the response of foreign countries, notably the U.S., see for example Strauss (1980, 1981); Friedman (2017). Existing historical research has unravelled important patterns of Jewish emigration. These earlier studies provided estimates of the number of Jewish emigrants per year and by geographical regions as well as the demographic composition of Jewish emigrants, notably Strauss (1980, 1981) and Rosenstock (1956). These studies describe and discuss a spatial pattern of migration over time that is similar to what our data reveals, such as the initial rush to Western European countries, and the later shift to further away destinations, such as the U.S. and Shanghai (see Section 3). Note however, that the numbers provided in this earlier research are limited in scope and precision, as they are often based on rough estimates by statistical authorities, and not based on comparable micro-data used in this paper.

Tombe and Zhu, 2019; Caliendo *et al.*, 2020, for instance), we then conduct counterfactual policy experiments. A contribution to that literature is that those experiments are among the first to be implemented on historical data. The inclusion of network effects in both levels of the migration decision is also an innovation.

Finally, we contribute to the empirical literature on migration by providing a cleaned and cross-validated version of the Resident List to the communities of economists, historians, and social scientists. We believe that the coverage and quality of this dataset makes it valuable not only for tackling the specific case of Jewish migration in Nazi Germany but also for studying refugees' migration in general. The reasons why the dataset is so unique—even compared to many modern datasets on refugees—are numerous: It covers most of the population at risk. It provides information not only on the migrants, but also on the stayers. It includes multiple destinations allowing us to study the choice to outmigrate *and* the choice of destination (which feedbacks on the incentives to outmigrate). It contains fine-grained information on individuals and localities. To the best of our knowledge, none of the data sources covering modern or historic contexts cover all these features.

3 Historical Background and Data

3.1 Historical Background

Jewish life in Germany. When Hitler took power in 1933, about 503 thousand Jews were living in Germany according to the Census of 1933, of which about four-fifth were of German origin, and about one-fifth were of foreign nationality (Statistisches Reichsamt, 1936).[6,7] Compared to 65 million people living in the German Reich, the Jewish community was small and made up less than 1% of the total population. Its members were spread over the entire area of the German Reich inhabiting more than 5,000 different towns, however, a large fraction lived in urban centers.[8]

Social relationships within the Jewish community (*"Gemeinde"*) were the cornerstone of Jewish life. Jews interacted intensively with their extended family, Jewish friends, and community members. The *"Gemeinden"* consolidated all Jewish activity in an area (Gruner, 2011). They collected taxes to finance religious and communal institutions, such as synagogues, schools, newspapers, and charities (Gruner and Pearce, 2019). Germany's Jews participated in more than 5,000 different Jewish local and national associations. While religiosity declined in the early twentieth century, synagogues remained a crucial center of social interactions. Loyalty towards the Jewish community was high.[9] Even Jews who were not practicing their faith participated in community meetings

[6] The Census of 1933 recorded Jewishness based on religion. Zimmermann (2013) estimates that about 600 thousand individuals of Jewish ethnicity were living in Germany in 1933.

[7] See also Online Appendix F for additional details of the historical background.

[8] Maurer (2005, page 273), referring to resources other than the ones we use, reports that 55% of the entire Jewish population lived in the top 10 cities in 1933, a number consistent with the share we obtain with our data.

[9] According to Kaplan (2005), only about 23,000 Jews converted in the German Empire from Judaism to Christianity (between 1871 and 1919), which given a population of Jewish people estimated around 500 thousand would be about 4%. Data from the pre-war period from 1933 to 1939 for a selected set of cities from Mendelsohn (1999) suggest small conversion rates in a range between 0.7% to 4.5%, with the outlier being Hamburg at 10,5%. The motives for conversion were often non-religious, as most converts tried to escape discrimination. Even Jews that converted to Christianism did not give up on their Jewish relationships, and the social relationships of converts tion with their Jewish friends and

and Jewish organizations and donated to Jewish charity. Besides the community, socialization took place within the extended family, and matchmaking created extensive family ties. Family and friends provided networks of support and stood in regular contact, either by phone, letter, or through personal visits on weekends and religious holidays, even when living apart. Despite increasingly socializing outside of the Jewish community in professional and German communities, the circle of close friends remained Jewish: *"Generally, friendships among Jews were more frequent and closer than friendships with non-Jews. For some families their entire circle of friends was exclusively Jewish. [...] neither acculturation nor the abandonment of religious observance brought about close relations with non-Jews."* (Maurer, 2005, p. 335). To sum, socialization within the community and the extended family created strong social ties among Germany's Jews at the eve of Hitler's rise to power. For Jewish citizens of Germany, Jewish networks gained further in importance with the start of Nazi persecution in 1933 (Maurer, 2005, p. 273).

Persecution. Immediately after the Nazis rose to power in January 1933, anti-Jewish legislation, state-led antisemitic actions, and violence began.[10] The Nazi government sought to push Jews out of the country by taking away legal rights and economic opportunities and by excluding them from social life. Just weeks after the new government was elected, a nation-wide boycott of Jewish business took place, which marked the first planned act of Jewish persecution. During the boycott, numerous shops were attacked and destroyed, and their owners were taken into "protective custody" (*"Schutzhaft"*). "Protective custody", the Nazi euphemism for arbitrary and indeterminate detainment, was one of the most powerful instruments of the Nazis to persecute individuals that they deemed unwanted. It was officially framed as being necessary to protect the detained Jews from the "righteous" wrath of the German population. Individuals could be taken into custody by members of the Nazi Party's organizations (SA and SS) and the Secret State Police (Gestapo) without judicial warrant or justified reason. The public spurred detainment, for example by reporting cases of "race-defilements", or by denouncing business competitors of alleged crimes (Wünschmann, 2010). After days or weeks of detainment, the detainees returned to their family often severely beaten and emaciated. In many cases (particularly after November 1938) prisoners were required to sell their belongings and to emigrate within the next few months as a condition to be released. Especially in the years prior to November 1938, detainment was not organized centrally. The historical evidence suggests that who and how many people were detained was largely idiosyncratic at the local level. It depended on local antisemitic sentiments and arbitrary decisions of local party members, which created an environment of fear (Bartrop and Dickerman, 2017). We document a strong positive correlation between detainment and local antisemitism using our dataset in Online Appendix Section G.2.

In the first years of the Nazi rule, incidences of violence intensified and culminated in the infamous 1938 November pogroms, known as "*Kristallnacht*" or the "Night of Broken Glass". During *Kristallnacht*, hundreds of synagogues were attacked, several thousand businesses were destroyed, and thousands of men were taken into custody. With the start of WWII, the Nazi policy changed

family remained intact (Kaplan, 2005).

[10] See Online Appendix Section F.5 for a description of the main antisemitic events.

from encouraging Jewish outmigration to the extermination of all Jews. At the end of 1941, emigration from Nazi Germany was officially forbidden, and the mass deportations to the concentration camps began shortly afterward.

When the Nazis took power, most of the Jewish population believed that the Hitler regime will only be short-lived. *"Hitler used the Jews as propaganda, now you'll hear nothing more about the Jews,"* or *"Such an insane dictatorship cannot last long,"* illustrate popular sentiments (Nicosia and Scrase, 2013). However, imprisonments and other antisemitic actions created fear in the Jewish community and made the danger more apparent. Anecdotal evidence, presented in Online Appendix Section F.2, documents the exchange of information about incidences of persecution within the Jewish community. These personal and indirect experiences with antisemitic events provided important information about the extent of persecution.

Deciding to stay or leave. Jews faced many dilemmas and uncertainties while deciding to stay or migrate, and where to go. While emigration was voluntary, it involved high costs. Migrating meant to leave behind traditional lives, to split up from family and friends, and to suffer a loss of economic status and wealth. The Nazi government sought to benefit from Jewish emigration by levying several taxes to expropriate migrants.[11] Migrating also involved bypassing bureaucratic hurdles, such as filing applications and paying for visas. Obtaining a visa was often extremely difficult and came with additional requirements, such as a personal contact in the destination country. Settling in a new country was burdensome, as migrants frequently did not master the local language, and as their skills and academic qualifications acquired in Germany were often useless in the destination countries.[12] Thus, migrating implied a future in which living conditions, status, and social relationships were highly uncertain. Anecdotal evidence, presented in Online Appendix Section F.3, documents that the emigration question, i.e., migration plans and prospects, were heavily discussed in personal visits and letters within the Jewish community.

Immigration policies. As the world economy recovered slowly from the Great Depression (1929), many countries imposed restrictions to immigration. Policies aimed at curbing immigration included quotas and visa restrictions. The allocation of entry visas depended on qualifications, financial means, resident relatives and friends, age, or state of health. Already in 1924, the U.S. had fixed a quota that restricted entry to 27,370 migrants per year from Germany and Austria, independent

[11]Since 1931, migrants from Germany had to pay a flight tax ("*Reichsfluchtsteuer*"). In 1934, the Nazi government lowered the levels of income / wealth above which individuals migrating were required to pay this tax. From May 1934 onwards, migrants had to pay a flight tax of 25 percent if their yearly income exceeded 20,000 Reichsmark (equivalent of 5,000 USD in 1934), or if they possessed assets worth 50,000 Reichsmark (equivalent of 12,500 USD in 1934). On top of that, after the pogroms of November 1938, Jewish emigrants were also required to pay an emigration levy of between one to ten percent ("*Auswandererabgabe*") for assets above 1,000 Reichsmark that they wished to transfer. In addition, the Nazi government heavily restricted the transport of wealth and private belongings outside of Germany. Financial assets had to be moved to a domestic account from which only small fractions could be transferred abroad. Ritschl (2019) estimates that over the period from 1933 to 1937, these policies resulted in an effective tax rate of 77% on migrants.

[12]Germany's Jews often lacked transferable skills that could meet foreign labour demands in sometimes less developed emigration destinations. For example, to prepare for emigration to Palestine, young Jews got trained in agriculture and crafts. Some of the academic qualifications were less useful abroad, especially if they had a distinctive German component, as it was the case for lawyers trained in German civil law (e.g., Heusler and Sinn, 2015).

of religion (Stiftung Jüdisches Museum Berlin, 2006). As the situation for Jews within Germany aggravated during the 1930s, more and more countries closed their doors. Only a limited number of destinations, such as the "Shanghai International Settlement" (Shanghai, hereafter), had little to no restrictions and remained open for Jewish refugees until 1941 (Stiftung Jüdisches Museum Berlin, 2006). In July 1938, 32 countries met at a conference in Evian, France, to discuss solutions to the Jewish refugee crisis. The conference, however, did not result in an agreement regarding how to allocate the flow of Jewish migrants, as none of the participants, except for Dominican Republic, wanted to commit to accept additional refugees. The reluctance of destination countries to accept more Jewish refugees was often backed by sentiments of antisemitism and hostility towards migrants shared by politicians and the wider public.[13] As the world's doors gradually closed and Nazi terror intensified, it became increasingly difficult for Jews to leave Germany.

3.2 Data on Migration and Deportation

We base our empirical investigation on individual-level information on migration, detainment, and deportation of Jews living in Germany in the 1930s. Our main source of information is "the List of Jewish Residents in the German Reich 1933–1945" (hereafter, the "Resident List") that was compiled by the German Federal Archives (Zimmermann, 2013). We supplement it with the 1939 Census to recover further information on the non-migrants.

The Resident List provides information on individuals' first and last name, birth date, birthplace, gender, place of residence. Particularly valuable for the purpose of our study, it records (when relevant) the migration and/or deportation history: timing of migration movements, the first and second destination countries, and/or the deportation date and place. Moreover, it conveys information on whether and when a person was detained (subject to "*Schutzhaft*").

To the best of our knowledge, these sources had never been used for quantitative research. An important part of our work is hence devoted to transforming the raw information into estimation datasets and validating its suitability for econometric analysis. In this section, we provide a summary of our data construction procedure, with more details in Online Appendix A. In order to estimate our model of outmigration decision, we need to define and collect information on the overall Jewish "population at risk" (of persecution) over the period. Then, we explain a number of sample cuts dictated by data availability and our procedure for encoding migration trajectories. Finally, as for any new dataset meant to be available for the research community, we conduct a number of validation exercises.

3.2.1 Defining and Measuring the Population at Risk

In this subsection, we explain the definition of population at risk of persecution in Nazi Germany and how the Resident List relates to it.

[13] Opinion polls conducted in 1938 and 1939 in the US, shown in Online Appendix Figure F1, illustrate that a majority of Americans was against hosting Jews persecuted in Germany.

Decision makers. Our period of study starts in 1933 and ends in 1941—the last year before outmigration was banned. We aim to cover the population likely to be persecuted because of being Jewish according to the Nazis. We therefore need definitions of (i) Jewishness, and (ii) who is a decision maker in the migration choice. Regarding the first point, we chose to work with the Nazis' definition of "ethnic Jews" according to the Nuremberg Laws of 1935. The second point involves focusing on adult (i.e., excluding children) residents of Nazi Germany that were not forcibly displaced. The individuals who migrated and those who stayed are considered to be decision makers. Similarly, individuals who stayed in Germany and passed away before 1941 are considered to be decision makers until their death.

Nuremberg Laws definition. According to the Laws of September 15, 1935, individuals with three or four Jewish grandparents were considered ethnic Jews. Those with one or two Jewish grandparents were considered mixed race of second and first degree, respectively. First-degree mixed-race individuals (with two Jewish grandparents) were considered ethnically Jewish under some circumstances, in particular if they practiced the Jewish religion or if they were married to a Jew.[14]

Estimation of the number of Ethnic Jews residing in Germany in 1933. About 503 thousand people were registered to be Jewish (based on their confession) in the Census of June 16, 1933 (Statistisches Reichsamt, 1936).[15] The Federal Archives, however, estimate that a total of 600 thousand people of ethnic Jewish origin lived in Germany between 1933 and 1941 (Zimmermann, 2013). While being less precisely measured, this figure is more relevant for our purpose of working with an ethnic definition of Jewishness. Starting from the 503 thousand persons recorded as Jews according to their religious affiliation in the Census of 1933, the Archives added (i) an estimate of (non-religious) ethnic Jews according to the definition of the Nuremberg Laws, (ii) newborns in the period between 1933 and 1941, and (iii) so-called "half-breeds", i.e., individuals with one Jewish grandparent or non-religious individuals with two Jewish grandparents. To obtain an estimate of the number of ethnic Jews that were not in the 1933 Census (where Jewishness is based on religious confession), the Archives used the 1939 Census, which contains information on *both* ethnic (according to the Nuremberg Laws) and the religious definitions of Jewishness. As explained in Zimmermann (2013), the Archives first compute the ratio of ethnic over confessional Jews surveyed in the 1939 Census. This reveals that 8.44% of ethnic Jews did not belong to the Jewish religious community. The Archives then apply this proportion to the population of 503 thousand confessional Jews covered in the 1933 Census, adding approximately 42 thousand Jews. Second, 13 thousand ethnic Jews individuals who were born after 1933, according to the 1939 Census, were added. Third, the

[14] The official wording of the laws is that Jewishness is based on race. However, practically, it seems to be have been based at least in part on participation to the Jewish community or on religion practice. For instance, Jews who recently converted to Christianity were still considered Jewish, and Christians who converted to Judaism were also considered Jewish. (See for example the full text of the Reichsbürgergesetz (1935), as well as an interpretation of the law from March 1938 in the wake of the occupation of Austria (DOEW, 1988)).

[15] The two censuses taken in 1933 (January and June) report the number of Jews as adherents to Judaism. 499,682 Jewish people were recorded in the Census of June 1933, this number increases to approximately 503,000 taking into account Jews living in Saarland (see Statistisches Reichsamt, 1936).

Archives add an estimate of the number of so-called half-breeds in 1933 to arrive at 600 thousand individuals, a number that Zimmermann (2013) regards as consensual among scientists.

The Resident List. Our source of information for the migration and deportation status of Jews, as well as for their detainment status and cities of birth and residence, is the Resident List. In 2004, the Federal Government of Germany commissioned the Federal Archives to construct a scientifically sound and complete list of all Jewish residents who had lived in Germany between 1933 and 1945. What is unique with the Resident List is its coverage of the resident population of ethnic Jews at the individual level.[16] The aim of the project was "to establish a list as completely and precisely as possible of the approximately 600,000 Jewish residents from a variety of sources, who [at any time] between 1933 and 1945 had resided in Germany [...]." (Zimmermann, 2013). However, the total number of individuals in the raw data of the Resident List (provided to us by the Federal Archives on March 25, 2019) is larger, at 812,566. The difference comes from the fact that the Federal Archives chose a criterion of inclusiveness in the Resident List that was *broader* than what is needed for our purpose. The first difference is that they included all residents they could find with at least one Jewish grandparent.[17] The second difference, which emerges when we investigate the characteristics of the set of individuals that lack information on their city of residence, seems to be related to the inclusion by the Resident List of a large number of deported Jews who might have transited through Germany after 1941 (see Online Appendix B).

3.2.2 Treatment of the Resident List

In this subsection, we briefly describe our treatment of the Resident List from the raw data to the estimation samples.

Cleaning the Resident List. We clean the raw data from the Resident List to construct social networks, to build city-level measures, and to merge it with other datasets. In particular, we harmonized first names and city names within the Resident List, a task made necessary by large variations in which names appear (e.g., German vs Polish, Czech, Russian, Ukrainian).

Table 1 summarizes the different steps of how we narrow down the sample from the raw data source to our estimation samples. The number of individuals in the raw data is reported in row 0 of the table. Our empirical framework requires that the individuals included in the data were

[16] It is important to note that the individual-level data of the 1933 Census have been lost (according to our personal correspondence with the Federal Archives). The only official (un-digitized) data on Jewish population in 1933 we managed to find from the 1933 Census are aggregated at the district level. The (digitized) 1939 Census does report individual-level information on Jews but at a date where many of them already left the country. While we make use of the 1939 Census to recover information on a subsample of the Jewish population, we cannot use it to re-construct the whole population at risk in 1933.

[17] Upon our inquiry, in a written reply dated March 20, 2021, the Archives confirmed that all Jews listed in the Resident List have at least one Jewish grandparent. Quoting from the reply, translated by the German native-speaker among coauthors: "The Resident List documents all Jews who, between 1933 and 1945, voluntarily took up residence in the German Reich within the borders of 31 December 1937, regardless of their nationality. In the Resident List we work with a broad definition of "Jewishness", i.e., all persons are documented who had between one and four Jewish grandparents, i.e., all so-called "full Jews," "half Jews," and "quarter Jews." This leads to a correspondingly high number of persons."

indeed at risk, in particular being residents of Nazi Germany between 1933 and 1941. The Resident List lacks information on the city of residence for a large number of individuals (see Section 3.2.3 and Online Appendix B for a thorough analysis of this issue).[18] Dropping those individuals (along with 7,473 duplicate entries), the sample gets to 550,908 individuals, corresponding to row 1. This figure is much closer to the estimate of the number of Ethnic Jews living in Nazi Germany given by Zimmermann (2013). We call this sample the "Core Resident List" (CRL).

Table 1: Process leading from the raw Resident List to our sample

	Individuals present in the data		
Steps	Resident List (+ data available)	1939 Census	# of ind.
(0)	Yes		812,566
(1)	+ City of Residence (CoR)		550,908
(2)	+ Date/City of Birth (CoB)		476,727
(3)	+ geo-coded CoR/CoB		430,691
(4)	+ adults only (16+)		343,560
(5)	+ migration /deportation date		173,816 #, ‡
(6)	+ in Germany *pre* 1939,	Yes	83,959 ‡
(7)	+ death in Germany/expelled to Poland	No	16,624 ‡
(8)=(4)-(5)-(6)-(7)	Without further information after 1933	No	69,161

NOTE: The symbols # and ‡ signal individuals who are part of the F sample (full spells) and the FP sample (full and partial spells), respectively. (See Online Appendix Section A.3.1 for the construction of these estimation samples.)

Our measurement of social networks is based on the date and city of birth of individuals, along with their city of residence. Dropping observations without such information brings us to a sample of 476,727, corresponding to row 2. Geocoding cities is crucial for computing spatial distance between individuals to identify the distant peers in the social networks. We collected latitude and longitude for 6,006 cities. For a set of cities, it was impossible to recover coordinates. These cities are for the most part very small settlements. There were misspellings in city names with no obvious way to correct, or no further information to unambiguously identify places with identical names. In addition, the geocoding procedure revealed that a small number of residence cities (1,307 individuals in 35 cities) were not within the borders of Germany on December 31, 1937 (the date chosen by the Federal Archives for establishing the Resident List). Dropping the latter two sets of cities, the sample size falls to 430,691, corresponding to row 3. Figure 1 illustrates the spatial distribution of this population. Panels (a) and (b) display the size of the Jewish population across residence cities and birth cities, respectively.

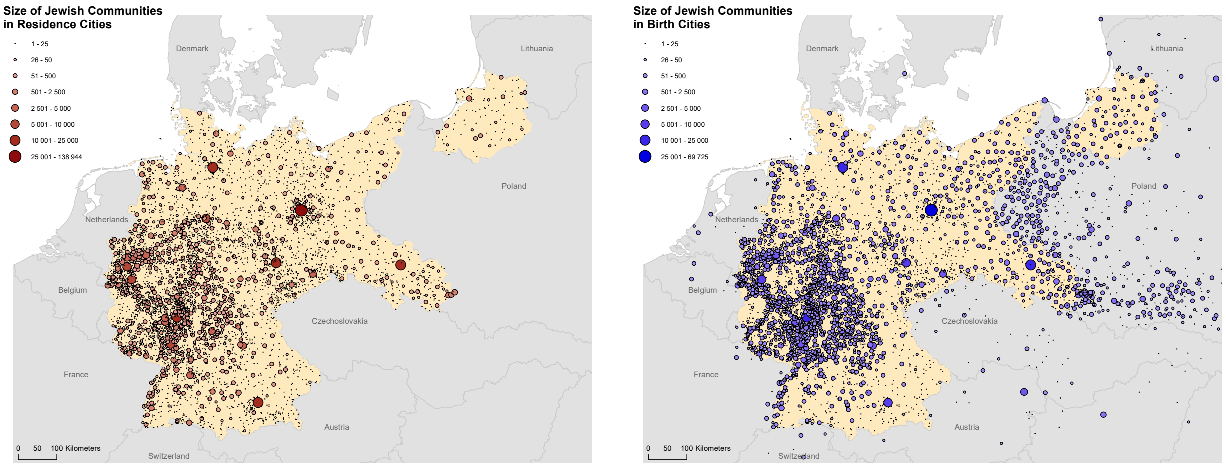
Keeping only "adults". We reduce the sample to the population most likely to be true autonomous decision makers. The legal majority status was set at 21 in Germany over that period. However, 16

[18] There are up to three variables that list cities of residence for each individual. The first one is providing the last known residence, while the two others refer to older residence cities (but with no date attached). In our benchmark analysis, we work with the last known city of residence, but investigate the robustness of our results to using previous cities of residence.

Figure 1: Spatial Distribution of Jewish Population

(a) Cities of residence

(b) Cities of birth



NOTE: Boundaries of Germany as of February 28, 1938. The sample comprises 430,691 individuals with known city-of-birth and city-of-residence coordinates, corresponding to row 3 of Table 1.

is thought by historians (see Online Appendix Section G.1) to be the earliest age at which someone could be economically independent at this period: The end of compulsory schooling was at age 14, to which 2 years apprenticeship were typically added. Empirically, we observe a clear jump in the probability of migrating at age 16, a pattern that is documented in Online Appendix Figure G1. In our baseline analysis, we therefore restrict the sample of choosers to individuals aged 16 or above in 1933. (We also adopt the legal definition of adulthood by restricting the sample to individuals above 21 in Online Appendix Section G.13). This step involves cutting the sample to 343,560, corresponding to row 4 of Table 1.

Estimation samples. In this step, we transform individual-level migration information into a single "spell" of zeros and ones to estimate a discrete-time duration model of outmigration over 1933-1941. A spell consists of zeros between 1933 and the last year before exiting the sample; the year of migration being coded as one. Spells of individuals who did not migrate are made of zeros only. Treatment of death and right-censoring are described in details in Online Appendix Section A.3.

Our baseline econometric analysis considers the sample of individuals with full spells: those with a recorded deportation or migration date in the Resident List, corresponding to row 5 of Table 1. We call it the Full-spell sample, in short *"F sample"*, made of 173,816 adults. In the robustness analysis, we extend the estimation to adults with censored spells as well, adding the 83,959 individuals who can be matched with the 1939 Census but have no migration year or deportation information in the Resident List, corresponding to row 6. We also add 16,624 individuals who died before 1939 and those who were expelled to Poland in 1938/39, corresponding to row 7. Adding such partial spells, we get to our Full-and-Partial-spell sample, in short *"FP sample"*: 274,399 out of the 343,560 adults. Because many individuals miss information after 1939, the FP sample can be used for estimation purpose only over the 1933–1938 subperiod, while the F sample can be used over the entire 1933–1941 period.

Jewish Ethnicity in estimation samples. We investigate in Online Appendix Section C.1 whether the estimation samples mainly consist of people at risk of being persecuted, i.e., identified as ethnic Jews by the Nazi authorities. We find that 97.5% of individuals in the F sample whom we can match with the Census of 1939 (migrants and deportees still present in Germany at that date) unambiguously meet the Nuremberg Laws criteria for Jewishness (three or four Jewish grandparents). As such, they were clearly in danger of Nazi persecution. In the FP sample, 77.5% of individuals were ethnic Jews. The remainder had two or less Jewish grandparents. Jews with a single Jewish grandparent were citizens of the Reich, while those with two Jewish grandparents were considered as citizens if they were "culturally assimilated." Depending on their classification, these individuals could be subject to various antisemitic measures, including economic restrictions, forced labor, and deportation. Whether these individuals should be included in the population at risk in our econometric analysis is an open question that ideally should be addressed on a case-by-case basis. In the absence of a definitive answer to this difficult question, it reinforces the case for using the FP sample for robustness analysis (F sample being the baseline).

3.2.3 Validating the Quality of the Resident List and Estimation Samples

Working with historical datasets comes with a necessity to deal with data imperfections. A particular feature is probably that a large part of the information that would ideally be needed has been lost along the years (during the war of course, but also later). This "missingness" problem is of concern if it implies a selection bias, where the individuals with non-missing data would exhibit specific characteristics and behaviors. Since our goal is to provide estimates and counterfactuals for the *entire* population at risk, we now turn to describe our investigations regarding the existence of patterns in data missingness. This section briefly summarizes the main findings; the details are relegated to Online Appendices B and C. We present the robustness checks to test the extent of potential biases due to data selection in Section 6.

From the raw data to the Core Resident List. The Resident List lacks information on the city of residence for a large number of individuals (254,185). These individuals, together with 7,473 duplicate entries, account for the difference between row 0 and row 1 of Table 1, that is between the raw data and the Core Resident List. The evidence we present in Online Appendix Section B.1 suggests that many of them do not belong to the population at risk that we are interested in (i.e., Jewish residents in 1933) as they were likely transitory residents (people detained in an occupied territory, who transited through Germany at some point before or during deportation). As such, the fact that they are not included in the Core Resident List is not too concerning for our estimations, since a large part of them are actually not decision makers.

Validating the Core Resident List. For our current study, but also more generally, we want to assess whether the Core Resident List (which is new to the literature) can be considered as a reliable historical source. For this purpose, Online Appendix Section B.2 compares the Core Resident List to two alternative sources that also contain information on Jewish communities at the city level: (i)

"The Encyclopedia of Jewish Communities in the German Language" by Alicke (2014); and (ii) the Census of 1925. The latter is official data, but for a date 8 years before our sample period. Also its definition of Jewish population is based on religion and not ethnicity. Alicke's Encyclopedia is not an official statistical source but is available for years closer to 1933. The take-away of this validation exercise is that the Core Resident List, when aggregated at the city level, offers a measurement of the population at risk that is validated by both external sources.

Missing information on city of birth or geo-coordinates. In Online Appendix Section B.3, we look at the characteristics of the 120,217 individuals in the Core Resident List for which we lack information either on their place and/or date of birth, or on the geo-coordinates of their cities of birth and/or residence. In our data construction procedure, this set of individuals corresponds to the sample cut between row 1 and row 3 in Table 1. We use information on their city of residence to compare them with the 430,691 individuals in the Core Resident List for whom information on place and date of birth and city geo-coordinates are available. Balancing tests show that both groups share a large array of similar city characteristics, which is reassuring regarding the (lack of) bias implied by the data cut based on availability of birth information and geo-coded city of birth and residence.

Representativeness of estimation samples. We investigate the representativeness of our estimation samples, the F and FP samples, in Online Appendix Section C.2. We analyze whether the 69,161 adults from the population at risk that we are unable to use in our empirical investigation because they (partially or fully) lack information on migration and deportation (row 8 of Table 1) are different from those individuals that make up our estimation samples.

For 24,868 of those out-of-sample individuals, we know that they were migrants but we do not observe their migration date. We can still compare them to in-sample migrants. We find that both city-or-residence and individual characteristics are balanced between the two categories of migrants. Hence, missing the exit date for a subset of migrants is unlikely to introduce a systematic selection bias in our estimated coefficients.

However, most of out-of-sample individuals have *unknown* trajectories: We do not know whether they migrated or stayed. When compared to in-sample individuals, their city-of-residence characteristics are balanced but we detect some mean differences along individual characteristics. Therefore, we report in Section 6 an important robustness check to assess how missing information on migration/deportation trajectories affects estimates. Our approach builds upon the entropy balancing method introduced by Hainmueller (2012), and used in the applied economic history literature to achieve balancedness across covariates between treatment and control groups (Satyanath *et al.*, 2017; Voigtländer and Voth, 2020; Angelucci *et al.*, 2021; Caprettini and Voth, 2021, among others). In a nutshell, we reweight observations of the F sample, using entropy weights, so that it matches the observable characteristics of individuals with unknown trajectories. The spirit of the approach is to evaluate the robustness of our estimates in the case where the individuals in the estimation sample would "look like" out-of-sample individuals. We find that the point estimates of our variable of interests remain stable across several reweighting schemes. This suggests that dropping individuals with unknown trajectories is unlikely to introduce a systematic selection bias in our estimated coefficients.

Overall, the results presented in Online Appendices B and C suggest that the Core Resident List provides an accurate coverage of the population of ethnic Jews living in Germany in the 1930s. The findings make us confident that our baseline estimation sample, i.e., the F sample, covers well the members of the Jewish community at risk. Finally, the robustness checks motivated by the validation exercises all point towards a remarkable stability of coefficients.

3.3 Descriptives of Migrants and Deportees.

We next describe several regularities observed in our data that are related to the characteristics of migrants and the aggregate migration flows. These descriptive statistics give a first idea of the broad patterns of Jewish migration. They can be regarded as a further validity check of our data in light of some of the well-known facts that the historical literature has documented. (i) Compared to non-migrants, migrants were on average younger (by almost 10 years) and more likely to be male (17 percentage points, see Online Appendix Table G2). (ii) While the average age of migrants was about 42, we observe a large jump in the probability to outmigrate starting at age 16, see Online Appendix Figure G1. (iii) Few individuals migrated in the early years of Nazi ruling, while after 1941 it was almost impossible to emigrate (Online Appendix Figures G2 and G3). (iv) Jews initially migrated to neighboring countries, such as France and Netherlands, in a false sense of security. About 11% of all emigrants were later deported (Online Appendix Table G2). The risk of getting deported was significant for those who migrated to neighboring countries. For example, about 50% of the Jewish migrants in the Netherlands were eventually deported (Online Appendix Table G10). As a result, after the war broke out in 1939, as the danger of German occupation in proximate countries increased, Jews fled to far away destinations, such as the USA, Shanghai, or Argentina (see the top 10 destinations for each year in Online Appendix Table G14).[19] (v) Before the November pogroms of 1938, incidents of persecution as measured by detainment ("*Schutzhaft*") were relatively rare events. However, persecution increased dramatically during and after the Night of the Broken Glass (see Online Appendix Figures G5 and G6). (vi) The largest number of people were deported in the years 1941, 1942, and 1943 (see Online Appendix Figure G7). Deportation was a death sentence: The median year of death for deported Jews is 1942. As shown in Online Appendix Table G3, based on incomplete information on the fate of individuals, we estimate that about 11% of deported individuals survived. Overall, the stylized facts based on our individual-level dataset depict a similar picture as earlier historical studies relying on other, more aggregate sources of information (in particular Rosenstock, 1956).

[19] See Online Appendix Figure G4 for the total migration by destination country over the entire period from 1933 to 1945.

4 Social Networks: Role and Measurement

Based on our reading of the historical literature on migration and communication about persecution within the German Jewish community, we hypothesize that, in a situation of political violence, social networks can affect outmigration decisions through at least two channels.

1. Threat Effect: Social networks in the origin country (here: Germany) aggregate available information on the extent of persecution. These information spillovers impact the incentives to outmigrate. The direct exposure of an individual to violence and persecution might be limited (e.g., until 1938 being taken into "protective custody" was rare—less than 1% of the Jewish population, see Online Appendix Figure G6), and public information in the radio or newspapers might be unavailable, or at least partly unreliable because of propaganda. Thus, individuals can extract information about the actual threat to their lives by observing the extent of persecution of their peers. We therefore expect an individual to be more likely to outmigrate if her peers were persecuted. We label this the *threat effect*.

2. Migration Spillovers: A larger number of peers who emigrate increases an individual's incentives to migrate out. On the one hand, this is the result of an expanding network in each destination country, which facilitates future emigration to this specific country. For example, peers abroad can provide private information about returns to skill in the destination country. They can also help with legal or logistic procedures in the visa process, facilitate finding a job and housing, and lower assimilation costs. We label this component of the migration spillover the *diaspora effect*. Empirically, it has been well established in many different contexts that larger diaspora networks in destination countries pull migrants to that country (e.g., Munshi, 2003; Beine *et al.*, 2011; Spitzer, 2015). Importantly, besides the diaspora effect, we identify an additional and novel channel through which peers' migration increase outmigration incentives in a situation of violence. As peers migrate out, the network in the origin country shrinks. We label this the *exodus effect* and it has two components. First, the fall in the size of the community leads to fewer business opportunities with group members and a depreciation of wages. It also comes with less in-group amenities and social connections. Moreover, it might result mechanically in more persecution (i.e., statistical targeting) of the remaining members of the group. Second, the observation of peers' migration can give a signal on the threat, as individuals can filter out the economic motives behind their peers' migration decision. Thus, the migration of family members or friends provides credible information about the danger of staying in Germany. We view this exodus effect as operating mostly when violence is pervasive and population displacement becomes substantial.

To illustrate, think of a person called Jakob fleeing Germany to the United States: The migration of Jakob not only increases the future likelihood that his friend Elisabeth also moves to the U.S. (diaspora effect), but also increases the likelihood that she moves to a country other than the U.S., for example Palestine (exodus effect). This is how information on destination choices included in the Resident List can be used to discriminate between the diaspora and exodus effects: Migration of peers to a given country increases future migration *only* to that country according to the diaspora effect, but to *all* destinations according to the exodus effect.

Measurement of the social networks and identification. Our key econometric challenge is to identify whether past detainment and migration *of peers* causally influence the outmigration decision of an individual. We might observe a (spurious) correlation between these variables even in absence of peer effects. This arises especially when individuals face a common environment and are exposed to the same time-varying shocks that affect simultaneously the migration incentives of an individual and her peers.

To address these so-called correlated effects, we restrict the construction of social networks to *Distant Peers*. We code as distant peers of an individual i all the individuals: (a) covered in the Core Resident List (so they are Jewish themselves); (b) born in the *same city of birth* as i; (c) *within the same age bracket* as i (± 5 years); and (d) living in a *different city of residence* than i and at least *5 km away from the city of birth*.[20] This focus on distant peers excludes family members, such as spouses and children, from the list of peers, who are likely to decide jointly and (at least try to) comove. The two maps featured in Figure 2 illustrate two examples of distant-peer networks observed in our data. Panel (a) displays the network of an individual born in Stuttgart in 1912. Her network is composed of all individuals born in Stuttgart in a 5-year window around 1912, and who have moved out of Stuttgart at some point: 53 distant peers that live in 33 cities of residence. Panel (b) displays the network of another individual who was born in Bonn in 1893 and who has moved to Stuttgart at some point where she is living during the Nazi period, according to the Resident List. Her network consists of all individuals also born in Bonn in a 5-year window around the year 1893, and who have moved out of Bonn to a city different than Stuttgart: 52 distant peers that live in 27 cities of residence.

Our approach regarding the influence of social networks therefore rests on two assumptions:

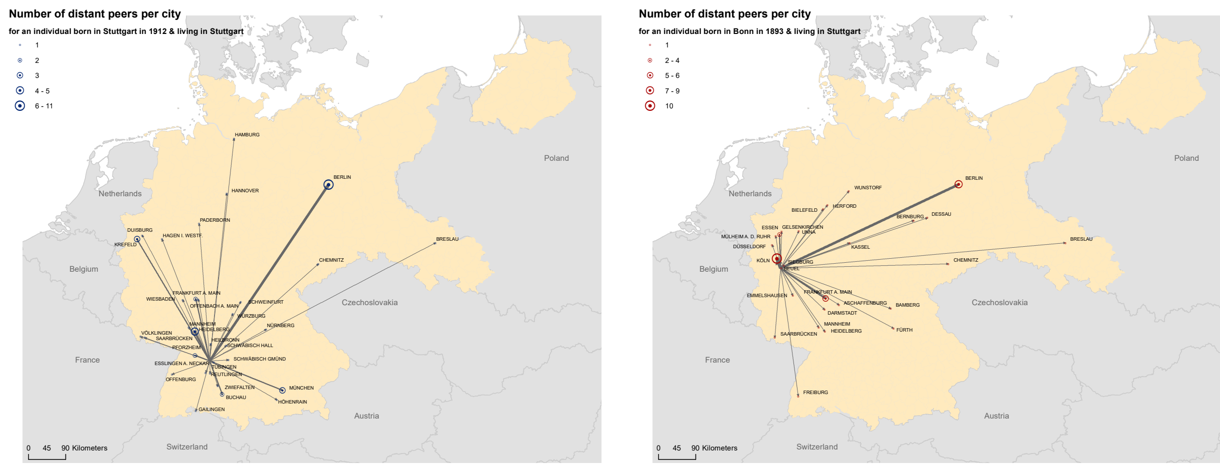
A.1 Social ties are created between Jewish individuals who were born in the same city and are of comparable age (criteria b and c above). We view this as a reasonable assumption given the historical context, as the social relationships of Germany's Jews were concentrated within *Gemeinden*. These Jewish communities were cohesive, relatively small, and spatially concentrated in most cities (the median community is comprised of 365 adults, see row 1 of Table 2). Even in large cities, Jews were living in similar neighborhoods, and they socialized within the community, for example in the synagogue, in Jewish associations, and in schools and shops. Later in the analysis, we run a robustness test where the largest and smallest cities of birth are sequentially dropped from the sample (Online Appendix Section G.13).

A.2 At least parts of the social ties created during childhood persist in the long-run and are robust

[20] To fix ideas, let us take the illustrative example of four individuals, A, B, C, and D. They are all present in the Resident List, therefore satisfying criterion (a). They are all born in the same city, Frankfurt, satisfying criterion (b). They are all born within 5 years from one another, satisfying criterion (c). At some point, B moved to Berlin, C moved to Hamburg, and D moved to München, while A stayed in Frankfurt. Based on criterion (d), distant peer network of individual A will be composed of all peers who moved to another city that is more than 5 kilometers away than the city of birth: individuals B, C, and D. For individual B, it will be composed of individuals C and D; for individual C, it will be composed of individuals B and D; for individual D, it will be composed of individuals B and C. Excluding individual A from the peer networks of B, C and D when building network variables is intended to abstract from the pull and push shocks experienced in Frankfurt (and therefore by individual A), which can simultaneously cause the migration of individuals B, C, and D.

Figure 2: Illustration of Distant-peer Networks of two Stuttgart Residents

(a) Born in Stuttgart in 1912 | (b) Born in Bonn in 1893



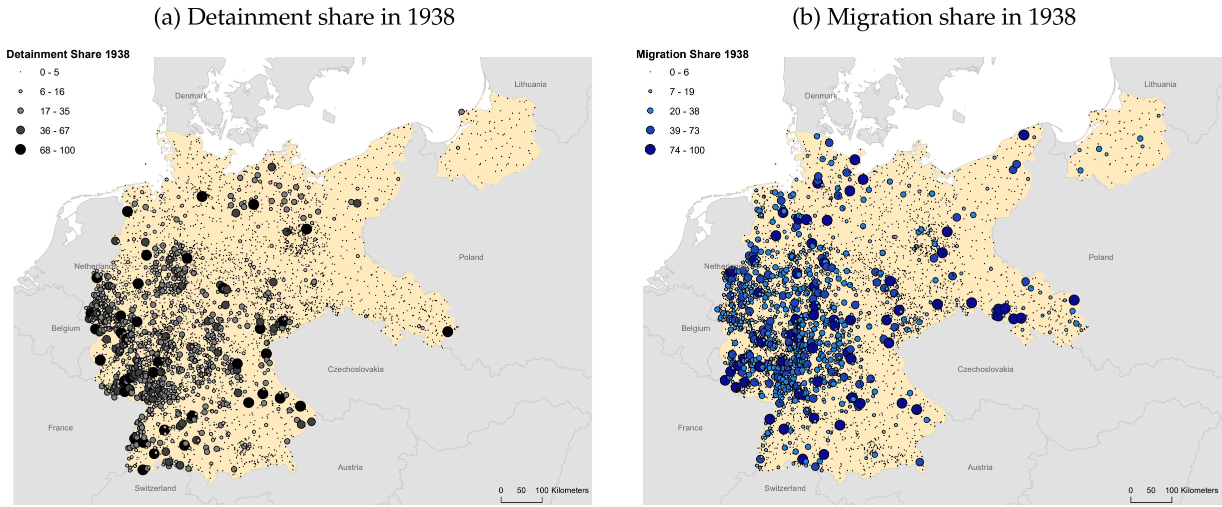
NOTE: This figure shows the distant-peer networks of two individuals both residing in Stuttgart in the period of our investigation. Figure in panel (a) displays the network of an individual that was born in Stuttgart in 1912. Figure in panel (b) displays the network of an individual that was born in Bonn in 1893, and who later moved to Stuttgart.

to spatial mobility (criterion (d) above). Evidence from social network studies suggests that a substantial part of links survives even after relocating. The strength of ties matters naturally: Family ties survive greater distances than friendships, which survive more often than networks with work colleagues, etc. (Fischer, 1982; Larsen *et al.*, 2006). Moreover, survival of long-term relationships is favored in particular when individuals do not invest sufficient time in extending their local networks—as is the case for individuals that relocate often (Viry, 2012). A limitation in our construction of the social networks is that we cannot distinguish between kin/non-kin or strong/weak ties; nor can we observe which ties survive when an individual moves to another city.

Our identification strategy exploits variations in the spatial distribution of distant peers across individuals. We compare outmigration choice between decision makers who live in the same city of residence but have different distant peers because they originate from different cities of birth and/or they were born in different years. Crucially, those distant peers themselves are exposed to push and pull factors in their own city of residence that are exogenous to the decision makers: (i) pull factors, such as connections to foreign countries, as well as (ii) push factors, notably related to the degree of persecution.[21] Together, local push and pull factors create spatial variation in outmigration of distant peers across cities, which we later use for identification in our estimation framework. Figure 3 visualizes the spatial variations in local persecution (i.e., detainment) and outmigration across cities in 1938.

[21] As detailed above, detainment was not centrally planned, and the historical literature suggests that it depended largely on the antisemitism of the local population and SA members. Indeed, in Online Appendix Table G4, we document a strong positive association between city-level detainment based on our individual-level data and measures of antisemitism taken from Voigtländer and Voth (2012). The idiosyncratic nature of detainment leads to time-varying variations in persecution across cities of residence.

Figure 3: Detainment and Migration across Cities of Residence in 1938



NOTE: The figures displays the variation in the share of detained individuals in 1938 in panel (a) and the share of emigrants in 1938 across cities of residence in panel (b), respectively. The sample comprises 430,691 individuals for which geo-coordinates are available, corresponding to row 4 of Table 1. Detainment share is computed as the number of detained individuals in 1938 divided by population in 1933. Migration share in 1938 is computed as the number of migrants in 1938 divided by population in 1933. The map represents Germany as of December 31, 1937.

Table 2 summarizes the characteristics of distant-peer network for the sample of adult Jews living in Germany in 1933 (row 4 of Table 1). We compute our network measures, i.e., migration and detainment of peers, only for decision makers whose distant-peer network comprises at least 5 people. Panel A presents the characteristics of all adults, while Panel B focuses on the F sample. The average size of Jewish communities in birth cities is about 7,750 adults in the all-adults sample, while the median size is considerably smaller with only 365 adults (Berlin explaining most of the right-skewness of the size distribution). Roughly four-fifth of the all-adults sample has at least 5 distant peers. For adults with at least 5 distant peers, the mean size of distant-peer network comprises 138 individuals (median = 49). The average size of Jewish communities is slightly lower in the F sample with mean value of 6,965 and median value of 285. Again, roughly four-fifth of the adults in this sample have at least 5 distant peers. For adults with at least 5 distant peers in the F sample, the mean distant-peer network comprises 127 individuals (median = 39), and peers live on average about 268 kms apart from each other.

For the set of individuals in the F sample, we compute cumulative migration and persecution rates among distant peers. Cumulative network detainment rates takes into account persecution from 1933 up to year t while cumulative network migration excludes observed network outmigration in year t (including migration from 1933 up to year $t-1$) to abstract from potential joint decisions taken by decision makers and their peers in year t. Empirically, we measure persecution by incidences of detainment ("*Schutzhaft*"). As detainment was a rare event, the average cumulative detainment rates were low and from about 0.1% in 1933 to reach 5.4% in 1941. In the beginning of the period, cumulative network migration was similarly low, with on average only 4.9% of distant peers having migrated by the end of 1933. The share of distant peers that left Germany increased

Table 2: Network Descriptives

	Mean	Median	SD	Min	Max	Obs.
Panel A: All adults						
Size of Jewish community in birth cities	7,749.6	365.0	15,936.9	1.0	47,172.0	343,560
Share of individuals with 5+ distant peers %	79.1	100.0	40.7	0.0	100.0	343,560
Size of distant-peer network	137.8	49.0	171.1	5.0	608.0	262,610
Panel B: Adults in the Full-spell Sample						
Size of Jewish community in birth cities	6,965.4	285.0	15,293.6	1.0	47,172.0	173,816
Share of indidivuals with 5+ distant peers %	78.4	100.0	41.1	0.0	100.0	173,816
Size of distant-peer network	126.7	39.0	168.2	5.0	608.0	132,421
Distance between individuals and distant peers	268.1	279.8	94.5	7.0	1,026.3	132,421
Network detainment rate by 1933 %	0.1	0.0	0.8	0.0	25.0	132,421
Network detainment rate by 1941 %	5.4	4.5	6.0	0.0	83.3	132,421
Network detainment rate between 1933–1941 %	2.4	1.9	2.7	0.0	35.6	132,421
Network migration rate by 1933 %	4.9	3.5	5.8	0.0	80.0	132,421
Network migration rate by 1941 %	41.8	40.3	21.0	0.0	100.0	132,421
Network migration rate between 1933–1941 %	21.7	20.0	13.2	0.0	93.3	132,421

NOTE: This table displays network summary statistics. Panel A provides numbers for individuals aged 16 and above in 1933 (row 4 of Table 1). Panel B concentrates on the F sample (row 5 of Table 1).

to 41.8% until the end of 1941. The average network migration rate from 1933 to 1941 is at 21.7%.

Characteristics of distant peers. Distant peers are movers, i.e., individuals who have moved out of their city of birth to a city within Germany at some point in their lives. In Online Appendix Section C.3 we compare characteristics of movers with those of individuals in the estimation samples (decision makers). Our tests show that city-of-residence characteristics are well balanced between the two groups. This provide evidence against strong selection bias regarding our network variables of interest which exploit variations across cities of residence of distant peers. However, movers are on average born in much smaller cities. This feature suggests one robustness check: excluding from the estimation sample all individuals living in cities with large Jewish communities (e.g., Berlin) or, reversely, excluding small communities in birth cities. In terms of individual characteristics, our balancing tests show that movers differ from the rest of the population. To address this concern, we run a robustness check where we restrict our analysis to movers as decision makers on the left-hand side of the migration regression (Online Appendix Section G.13).

5 Outmigration and Social Networks: Preliminary Analysis

In this section, we conduct a first "reduced-form" pass on the data. We put relatively more emphasis on the threat effect in migration decisions, namely how persecution of peers affects migration incentives. The in-depth analysis of the two migration spillovers (exodus and diaspora) relies on the construction and estimation of a structural model. Thus, it is relegated to Section 6, together with our instrumentation strategy and robustness analysis.

An empirical model of out-migration. The structure of our data naturally calls for a discrete time duration model of the migration decision. Section 3.2.2 explains how we transform individual-level migration information into spells, where each individual can outmigrate once. That is, our outcome of interest, Migrate_{it}, is a series of zeros until it takes on the value one in the year t in which individual i migrates (should he/she do so). In the year $t+1$ after the migration occurs, the individual exits the sample. Individuals also exit the sample when they are no longer in capacity to act as decision makers, i.e., after they were deported or after their known date of death. This also means that the econometric model excludes the alternative of re-entering in Germany in $t+s$ to an individual who has migrated out of Germany in t.[22] We study migration decisions from 1933 until 1941, since after October 1941 emigration was officially forbidden. This leads to the following specification, where the unit of observation is an individual i living in city $r(i)$ with network $n(i)$, at time t:

$$\text{Prob}[\text{Migrate}_{it} = 1] = \Phi\left[\mu \times \texttt{Mignet}_{n(i)t} + \gamma \times \texttt{Detainment}_{n(i)t} + \mathbf{X}'_{it}\delta + \text{FE}_{r(i)t}\right], \qquad (1)$$

where Φ is a functional form that depends on the estimation procedure. Equation (1) can be estimated with a Logit model, a Complementary Log-Log Model (Cloglog), which is particularly well adapted for dealing with discrete time duration models (Cameron and Trivedi, 2005), or a Linear Probability Model (LPM). While non-linear estimators (Logit and Cloglog) are preferable given our data structure and structural model (detailed in the next section), LPM offers several advantages: ease of dealing with high-dimensional fixed effects, allowing for a rich clustering structure, and transparent interpretation of the coefficients of interaction terms. We only report LPM estimation results in Table 3. The non-linear estimations, which yield comparable results when expressed in terms of marginal effects, can be found in Online Appendix Table G7.

Our first main variable of interest is $\texttt{Mignet}_{n(i)t}$ that measures the post-1933 share of distant peers who had already left Germany (strictly) before time t. Its coefficient μ captures the *joint* influence of the two migration spillovers, namely the exodus and diaspora effects, whose respective contributions will be disentangled when estimating the structural model. Our second variable of interest is detainment among networks members, $\texttt{Detainment}_{n(i)t}$, that we define as the share of distant peers that were detained until year t (included). The coefficient γ captures the threat effect that we expect to be positive. We control for a vector of individual characteristics $\mathbf{X}'_{it}$, comprising gender, age and its square, an indicator for whether the person was born outside of Germany, and an indicator for whether the person was ever detained herself, in the past or in year t.[23]

Importantly, our specifications take into account city-of-residence fixed effects that capture local geographical and cultural features that were important in the period, such as the clustering of

[22] We believe that it is a reasonable assumption in our context as the migration inflows of Jews in Germany in the 1930s is negligible. According to Niederland (1993), between 1933 and 1935 only about 10% of migrants returned to Germany; in 1935, German authorities started to place returning migrants in training/concentration camps.

[23] Those and additional variables used in the empirical analysis are described in Online Appendix E.

Jewish populations in some cities/areas of Germany (in particular in the West of Germany, see Figure 1), differences between rural versus urban places, larger share of secular Jews in urban localities (Kaplan, 2005), distance to borders and emigration facilitators (like visa offices, counseling centers of the *"Hilfsvereins der Deutschen Juden"*, etc.). In most specifications, the vector $\texttt{FE}_{r(i)t}$ corresponds to city-of-residence $\times$ year fixed-effects that absorb all time-varying local push and pull factors affecting co-residents of a city, such as local outbreaks of persecution and violence or adverse economic shocks. With more than 1,700 cities of residence, and over 15,000 city of residence $\times$ year combinations, the estimation of fixed effects is extremely demanding from the data.

The model is estimated on the F sample.[24] Since the construction of the network of distant peers is based on the city of birth, standard errors are clustered at the level of the birth city $\times$ year. Equation (1) is a particular case of the canonical empirical network setup (Bramoullé *et al.*, 2009) where the outcome variable is impacted by peer mean outcome (i.e., endogenous network effects), here $\texttt{Mignet}_{n(i)t}$. The estimation challenges have been extensively discussed in the literature and, in our case, pertain to correlated effects, namely shocks that affect the migration incentives of both the individuals and their peers. We believe that most of the concern is alleviated thanks to restricting the construction of social networks to distant peers and controlling for a fine-grained structure of fixed effects. However, we go one step further in Section 6 by using exogenous shifters of peers' past migration in instrumented regressions.

The first two columns in Table 3 report the LPM estimation results of equation (1) for two different sets of fixed effects. In column 1, we only include year fixed effects with the idea of validating our main findings with a coarse structure of fixed effects. Column 2 considers the full battery of city-of-residence $\times$ year fixed effects. Results are comparable across the two specifications. The coefficients of the control variables reassuringly have the expected signs. Individuals who are young and male are more likely to migrate out (a pattern observed in many outmigration contexts). Individuals who were ever detained themselves also display a significantly higher probability to outmigrate compared to individuals who were not taken into custody. In column 1, we explicitly control for a first order local push factor (absorbed in the fixed effects in column 2), namely city-of-residence-level persecution, measured by the cumulative share of co-residents who were detained in city $r(i)$ until year t. As expected, we find a positive and statistically significant impact of city-level persecution on outmigration. This finding is particularly relevant for our instrumental variable framework in Section 6 that exploits the link between detainment and migration in peers' city of residence.

Second, and more importantly for our purpose, network migration and network detainment both display a positive effect on outmigration probability that is significant at the 1% level. We interpret this finding as supportive evidence for the migration spillovers and the threat effect. In column 2, the magnitude of the migration spillovers is as follows (a breakdown along the exodus/diaspora categories is provided in Section 6.3): A one-standard-deviation increase in migration of network members (0.16) translates into a 1.2 percentage point increase in the annual migration probability (22% of the sample mean). Regarding the threat effect, a one-standard-deviation

[24] See Online Appendix Table A4 for a breakdown of the sample cuts from the F sample to the observations used in the regressions.

increase in peers' past persecution (0.043) increases the annual migration probability by 0.18 percentage points (3.3% of the sample mean). As a way of comparison, personal detainment has a quantitatively larger effect: Having been detained in the past increases the migration probability by 2.1 percentage points (38% of the sample mean). Note that these numbers should be interpreted in a conservative manner. Indeed, given the reduced-form nature of the regressions, our quantification exercise can only reflect the "static", and partial, impact on migration probability of the explanatory variables under consideration. Their full impact, which relies on dynamic externalities, will be assessed thanks to the counterfactual simulations of the structural model in Section 7.

Interpreting the threat effect: Learning. Which individuals react more to new information on persecution? We expect individuals to update more strongly their beliefs about the level of threat and danger of staying in Germany, if information about the victimization of their peers comes as a surprise. For example, detainment of peers elsewhere provides new information for people who live in localities where antisemitic sentiments are low or for secular individuals.

To study learning, we estimate heterogeneous effects of peers' persecution by interacting detainment of network members with characteristics at the individual level (identity and personal experience of persecution), characteristics at the level of the city of residence (historic antisemitism, and size of the local Jewish community), and characteristics of peers. Our empirical hypothesis is that a larger observed behavioral response, measured in terms of outmigration propensity, reveals a stronger belief updating.

We begin by interacting (demeaned) network detainment with individual characteristics. In column 3 of Table 3, we study the interaction effects with the Jewish Name Index, a proxy for how integrated an individual's family was to the German society based on the first name of an individual. The Jewish Name Index takes on higher values if the first name of the individual is more distinctively Jewish, such as Abraham and Rachel, and lower values if the first name is more distinctively German, such as Otto and Hildegard.[25] We interact network detainment with an indicator for whether the Jewish Name Index is above the median of the distribution. The coefficient of this interaction term, displayed in column 3, is negative and statistically significant. It shows that individuals with a more German-sounding first name tend to respond *more* to their peers' victimization. This is in line with anecdotal evidence that secular Jews who identified strongly with the German society were more likely to underestimate the level of actual threat (e.g., Nicosia and Scrase, 2013; Heusler and Sinn, 2015). Under this interpretation, those individuals were more likely to update substantially their beliefs when observing detainment in their network, and therefore reacted more strongly to this information. The main effect (i.e. linear coefficient) of a more Jewish sounding name is also negative—an explanation could be that individuals with a higher JNI

[25] *"Namings connected families to previous generations and to religious traditions [..] and confronted Jews directly with the vexed issue of tradition versus acculturation as families passed on secular first names—or "Germanized" old Jewish names."* (Kaplan, 2005, p. 239). See the Online Appendix D for details on the construction of the the Jewish Name Index. In Online Appendix Table G6 we validate that the JNI is highly positively correlated with Jewish ancestry based on a sample of individuals observed in the 1939 Census, which recorded the number of Jewish ancestors each individual had. Online Appendix Table G5 tabulates the most popular 30 names in the Resident List by gender, along with the number of times they appear and their computed Jewish Name Index (JNI).

Table 3: Outmigration Decision: 1933–1941

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Migration Decision						
						≤ 1938	> 1938
Migration of network members	0.085***	0.074***	0.074***	0.074***	0.074***	0.104***	0.023***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.007)
Detainment in city of residence	0.050**						
	(0.020)						
Detainment of network members	0.051***	0.043***	0.079***	0.047***	0.084***	0.080***	0.015
	(0.012)	(0.010)	(0.014)	(0.011)	(0.014)	(0.017)	(0.013)
$\times$ Jewish Name Index ($>$ *Median*)			-0.068***		-0.068***		
			(0.015)		(0.015)		
$\times$ Ever detained				-0.062*	-0.067*		
				(0.036)	(0.036)		
Jewish Name Index ($>$ *Median*)			-0.002***		-0.002***		
			(0.000)		(0.000)		
Ever-detained	0.028***	0.021***	0.021***	0.024***	0.024***	-0.039***	0.051***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Age	-0.042***	-0.039***	-0.039***	-0.039***	-0.039***	-0.040***	-0.051***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.010)
Age squared	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.028***	0.027***	0.027***	0.027***	0.027***	0.026***	0.031***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)
Born outside Germany	0.000	0.001	0.001	0.001	0.001	0.003**	-0.005*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Year FE	✓						
CoR $\times$ Year FE		✓	✓	✓	✓	✓	✓
Observations	974,308	972,603	948,038	972,603	948,038	715,807	256,796
R-squared	0.06	0.10	0.10	0.10	0.10	0.07	0.13
Mean of dependent var.	0.055	0.055	0.055	0.055	0.055	0.042	0.092
SD of dependent var.	0.227	0.227	0.227	0.227	0.227	0.199	0.288

NOTE: LPM estimation with the dependent variable being an indicator for migration in year t. The unit of observation is an individual in year t. The sample consist of individuals who were at least 16 years old in 1933. All continuous variables in interactions are demeaned. Detainment of network members measures the cumulative share of network members that have been detained until year t. Standard errors clustered at the city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

were less educated (see Online Appendix Table G6), and therefore likely to have had fewer means to outmigrate. Our interpretation that behavioral responses are attenuated by prior awareness of the danger is further strengthened by column 4, in which we document a negative interaction of peers' detainment and personal detainment: Individuals who were themselves detained in the past respond less to their peers' persecution. The interaction effects of network detainment with the JNI, and with personal detainment, are very similar if we estimate them in the same regression model (column 5). Finally, in columns 6 and 7 we split the sample into a period prior to and after

Kristallnacht (pre/post 1938). We find that the effect of network detainment is significant only in the pre-1938 period. After the November pogroms, when Jews across Germany realized that staying posed a significant danger to their lives, additional information on persecution coming from the network ceases to be important.

In Online Appendix Table G8, we consider characteristics of residence cities interacted with peers' persecution. The table documents that detainment among network members has a smaller effect for individuals living in cities with a larger local Jewish community.This negative interaction is suggestive of a competition taking place between the local and distant (peers) informational sources: In localities with larger local Jewish community, individuals could draw on more sources of information. With a larger sample of observations, they are therefore likely to extract a more accurate signal about the threat. It is also likely that in those large cities, individuals were more exposed to extreme (negative) realizations over the distribution of persecution events. Both mechanisms would contribute, in large cities, to reduce the importance of private information obtained from distant peers. Moreover, the table shows that detainment among network members interacts negatively with city-level antisemitism, measured by the occurrence of a pogrom in Medieval times, or the vote share for the NSDAP in the 1928 election. The negative interaction between detainment and antisemitism is again interpreted as a surprise effect: In cities where historical and recent antisemitism are ingrained in the collective memory of the local community, information about persecution coming from distant peers comes less as a surprise.

In addition, in Online Appendix Table G9, we investigate how peer characteristics affect the learning process. The table documents that individuals react more to information from peers living in greater distance, in line with the idea that private information from farther away places provide new information about the level of threat in addition to what the individual can observe locally. We also explore the importance of homophily in behavioral responses to victimization of peers. The findings show a stronger reaction of individuals to detainment among their peers who are culturally similar to them, i.e., have first names close to theirs in terms of JNI, which indicates a strong homophily in behavioral response to persecution of peers.

6 A Structural Model of Outmigration under Violence

In this section, we investigate how past migration of peers affects current migration incentives. We estimate the two network effects that drive migration spillovers, namely the diaspora and exodus effects. As explained above, using information on the destination chosen by migrants is key for discriminating between those. To this end, we build a random utility model of migration with network spillovers and estimate it in a nested logit setup. The structural model also serves the purpose of simulating counterfactual policies in destination countries (which we do in Section 7).

The nested logit model is a natural framework for the data at hand: We observe migration decisions as a binary outmigration choice repeated every year between 1933 and the year of exit for migrants and 1941 for the stayers. Conditional on deciding to migrate, we observe the discrete choice of a destination country. The overall decision process therefore comes in with a hierarchical structure of discrete decisions that the nested logit model was designed to fit. Outmigration is considered as a definitive exit—an assumption supported by anecdotal evidence (see footnote 22). Our analysis considers a repeated *static* choice, in the sense that individual migration decisions do not factor in the expected future realizations of relevant migration determinants.[26] Finally, our setup does not model the general equilibrium feedback effect of migration on economic activity, neither in origin country (Germany) nor in destination countries. Besides gaining in tractability, this modeling choice can be justified at the light of the historical context of the 1930s: The Jewish community in Nazi Germany represented less than 1% of the overall population. Similarly, in our data, inflows of German migrants were small in comparison with the destination countries' populations.

6.1 The Nested Random Utility Model of Migration

The random utility model applied to migration decisions starts with a specification of utility, U_{idt}, enjoyed by an individual i when located in country $d \in D$ in year t. In our specific case, the choice set D includes Germany as well as all potential destination countries during that period. Each individual selects the destination d^* that maximizes her utility: $d^* = \arg\max_{d \in D} U_{idt}$. The optimal choice can lead to staying in Germany or fleeing to another country. There are observable and unobservable determinants of that utility. We assume that the unobservable component of utility, ϵ_{idt}, follows a Generalized Extreme Value (GEV) distribution such that the decision-making process will result in a nested logit structure. More specifically, the distributional assumptions regarding the error terms are such that it is possible to decompose the choice into two "sequential" nested decisions: whether to *stay* in Germany, nest $B_s = \{\text{DEU}\}$, or *migrate* to a country belonging to nest $B_m = \{\text{USA}, \text{GBR}, \text{FRA}, ...\}$. In a second step, one must decide in which destination $d \in B_m$ to settle in. The decision to stay in Germany or migrate is referred to as the upper-level model while the destination choice is called the lower model.

There are two types of observable components in the utility function. The upper part component, W_{ikt}, varies across nests with $k \in \{s, m\}$. The second component, V_{idt}, varies across alternatives d within a nest:

$$U_{idt} = W_{ikt} + V_{idt} + \epsilon_{idt}, \quad \text{for} \quad d \in B_k. \tag{2}$$

We interpret ϵ_{idt} as the unobserved costs and benefits of emigration to destination country d for individual i in a specific year. For a given individual, part of that random component of utility is destination-specific (e.g., whether speaking the language of d). Part of the error term is likely to be *correlated* across all foreign destinations $d \in B_m$—for instance how portable are the skills of i but also how financially constrained she might be.

[26] We abstract from sophisticated forward-looking strategies where individuals, in spite of their high willingness to emigrate, would postpone their movement in order to free-ride on the migration effort of their peers (e.g., let them migrate first to a destination country d and then settle afterwards in the same destination in order to benefit from their experience, support, and help to lower migration frictions). Such a beachhead effect is conceptually appealing but comes at the cost of bringing additional elements of complexity without a clear gain in terms of empirical relevance. While it is possible that these strategic elements have played a role in the long-run dynamics of migration in other less-violent contexts, historical records do not emphasize that it played a first-order role in the post-1933 Jewish migration where time constraint was binding and persecution risk was high.

Regarding the observable components of utility at each level of the decision tree, we retain intuitive and flexible functional forms that are standard in the literature on the impact of migration networks. For foreign destinations, $d \in B_m$, the lower-level utility is assumed to take the following form:

$$V_{idt} = A_{dt} - \ln \tau_{idt} + \alpha \times \texttt{Diasp}_{n(i)dt}, \tag{3}$$

where A_{dt} represents the overall (log) attractiveness of destination country d (e.g., economic prospects) — a component that is estimated in our empirical analysis with destination $\times$ year fixed effects.[27] The term $\ln \tau_{idt}$ corresponds to migration "frictions" that we can either observe or capture with fixed effects (e.g., distance to border, availability of sea and ground transports, administrative efficiency, etc). Note that those frictions are modeled ad-valorem, i.e., they shift down the utility from traditional determinants in a proportional manner ($\tau > 1$). Finally, the term $\texttt{Diasp}_{n(i)dt}$ corresponds to our measure of peers living in destination d at time t (the exact functional form used for this variable is detailed in Section 6.2). Peers of i constitute the set of individuals within her social network, denoted with $n(i)$. We expect $\alpha > 0$, since the empirical migration literature has extensively documented positive effects of the existing set of migrants on later migration (which inspired a very large set of papers using Bartik instruments to predict migration shocks).

Turning to upper-level utility, we start by noting that since the outmigration decision is a simple binary choice between Germany and the rest of the world, only relative levels of covariates matter and we need to specify them only for one of the alternatives. The utility of individual i who decides to stay in Germany in time t is therefore affected by the component $W_{i\texttt{DEU}t}$, specified as:

$$W_{i\texttt{DEU}t} = -\beta \times \texttt{Mignet}_{n(i)t} - \gamma \times \texttt{Detainment}_{n(i)t} + \mathbf{X}'_{it}\delta + \mathbf{FE} \tag{4}$$

The variable $\texttt{Mignet}_{n(i)t}$ measures the post-1933 share of peers who had already left Germany (strictly) before time t. It differs from the diaspora variable in the lower-level decision, equation (3), in one key aspect: It encompasses all possible destinations. We expect $\beta > 0$, since the total accumulation of departures in an individual's network should affect negatively her utility to stay in Germany. This feedback effect is our empirical measure of the *exodus* effect. It is important to note here that although past migration flows of peers appear both in the lower and upper levels of the decision, both diaspora and exodus effects are clearly identified. Intuitively, in the lower level, it is the share of individuals having chosen each destination d relative to other countries that creates the diaspora effect. Identification stems from comparisons across destinations within migrant peers. At the upper level, what matters for identifying the exodus effect is the number of peers left relative to the initial stock of Jews who were susceptible to emigrate from Germany (which does not feature in the lower level estimation).

As defined in the previous section, the second variable of interest in equation (4), $\texttt{Detainment}_{n(i)t}$,

[27] It is typical in models featuring endogenous migration to model $A_{dt} = \ln\left(\frac{M_{dt} w_{dt}}{P_{dt}^{\theta}}\right)$, where M measure amenities of d and w_{dt}/P_{dt}^{θ} is the real wage, with θ being the share of non-tradables in the consumption basket of individual i. Since we will not model the general equilibrium effects of Jewish migration in destination countries, we can let A_{dt} capture all relevant determinants, seen as exogenous from the point of view of prospective migrants.

corresponds to the share of peers who were detained until year t. We expect its coefficient γ to be positive (threat effect). The vector $\mathbf{X}'_{it}$ represents a set of observable individual-level characteristics that influence the utility to stay in Germany. Finally, **FE** corresponds to a battery of fixed effects that varies across specifications. Particularly, we can include city-of-residence $\times$ year fixed effects that crucially capture all the local push and pull factors that are common across individuals living in a given city. The richness of our individual-level data therefore allows to control for a very broad spectrum of local differences that pushed Jews to emigrate.

As in Anderson *et al.* (1992), Train (2003), and Cameron and Trivedi (2005), we characterize the nested choice as two logit equations. The probability of choosing a foreign destination $d \in B_m$ is decomposable into the product of conditional and marginal probabilities: $\text{Prob}_{idt} = \text{Prob}\left(d \mid \text{mig}_{it} = 1\right) \times \text{Prob}\left(\text{mig}_{it} = 1\right)$. We follow Anderson *et al.* (1992) regarding the specific form taken by the GEV distribution assumed for the unobserved preference term :

$$F(x_1 \cdots x_n) = \exp\left[-H(e^{-x_1} \cdots e^{-x_n})\right] \quad \text{with} \quad H(x_1 \cdots x_n) = \sum_k \left[\sum_{d \in B_k} x_d^{1/\lambda_2}\right]^{\lambda_2/\lambda_1}. \tag{5}$$

Imposing this distribution on ϵ_{idt}, the conditional probability of choosing a given destination (lower-level model) can be written as:

$$\text{Prob}\left(d \mid \text{mig}_{it} = 1\right) = \exp\left(V_{idt}/\lambda_2 - I_{it}\right), \quad \text{with} \quad I_{it} \equiv \ln \sum_{d \neq \text{DEU}} \exp(V_{id't}/\lambda_2). \tag{6}$$

The log-sum term I_{it} is also called the inclusive utility, since it has been shown that $\lambda_2 I_{it}$ is the expected utility of being able to choose among all options at this level of the choice, destination countries in our case (Small and Rosen, 1981; Anderson *et al.*, 1992; Train, 2003).

Again, given that B_s is a singleton (all potential migrants initially live in Germany), the upper-level decision is a binary choice, and only relative levels of determinants to migrate out of Germany matter. This means that we can normalize without loss of generality $V_{i\text{DEU}t} = 0$. The marginal probability of choosing nest B_m and outmigrating (upper-level model) takes the following logit form:

$$\text{Prob}\left(\text{mig}_{it} = 1\right) = \frac{1}{1 + \exp\left(\frac{W_{i\text{DEU}t} - \lambda_2 I_{it}}{\lambda_1}\right)}, \tag{7}$$

where it is immediate that the set of utility determinants making Germany more attractive to i ($W_{i\text{DEU}t}$) reduces the migration probability, while the inclusive utility term ($\lambda_2 I_{it}$), which summarizes all the relevant information coming from the possibility of choosing one of the destination countries, increases it.

The parameters λ_2 and λ_1 play several critical roles in our model. From the GEV assumption, λ_1 and λ_2 capture respectively the between-nest and within-nest heterogeneity of the error term; their ratio λ_2/λ_1 is an inverse measure of correlation of the error term within the lower nest, which means the destination country choice here.[28] An important theoretical requirement is $\lambda_2 \leq \lambda_1$

[28] The nest s is a singleton in our case, therefore no within-nest correlation structure is attached to it.

for the model to be consistent with utility maximization for all possible values of the explanatory variables (Anderson *et al.*, 1992). The lower λ_2 is, the more correlated are the shocks to individual utility brought by alternative destination countries. With $\lambda_2 = \lambda_1$, the shocks are totally uncorrelated within a nest, and the model collapses to the standard multinomial logit where all choices are at the same "level". Those structural parameters also have a second role and interpretation in the model. From equation (6), it is clear that all coefficients relevant in the choice of destination country contained in V_{idt} will be scaled by $1/\lambda_2$. In equation (3), the determinants of the lower-level utility feature bilateral migration costs with unitary coefficient. Therefore $1/\lambda_2$ is also the elasticity of attractiveness of a destination country to both its expected real wage and migration impediments. This elasticity features as the migration cost elasticity in recent papers from the quantitative spatial literature.[29] Since λ_2 is reflecting the degree of heterogeneity in the idiosyncratic tastes of individuals with respect to locations, the intuition is straightforward: When individuals have very similar, correlated tastes (λ_2 approaching 0), they all flow to the country with highest expected income (after discounting for migration costs). Any change in expected real income of migration costs generates infinite flows of migrants. Following the same logic, $1/\lambda_1$ is the elasticity driving the response to all determinants of outmigration in the upper-level choice: Both upper-level variables $W_{i\text{DEU}t}$ and the expected maximum utility of migration $\lambda_2 I_{it}$.[30]

Estimation equations. We now discuss how we turn the structural equations (6) and (7) into an econometric estimation. The estimation proceeds in two steps. We start with the migration destination choice estimation (lower level), which we use to construct the inclusive utility. We then include it in the binary migration decision (upper level). Regarding the first step, because our data lack predictors of destination choices that would vary at the individual level, we work with an aggregated version of the model. The main characteristics we have on individuals are the place of birth and the last known place of residence. We therefore aggregate migration decisions at the city-of-residence $\times$ city-of-birth level. This yields a (triadic) gravity equation which still allows to retrieve the structural inclusive utility term. For the upper-level model, we can perform estimation at the individual level. The different aggregation levels involved in different steps of estimation make it natural to proceed with sequential regressions. A well-known disadvantage of such procedure applied to nested logit is that, although it retains consistency, it is not as efficient as a joint estimation (Train, 2003).

We start with the lower-level model. Aggregating the decision-making process means that migration frictions and diaspora can be decomposed into city-of-residence and city-of-birth compo-

[29] Fajgelbaum *et al.* (2018), Bryan and Morten (2019), Tombe and Zhu (2019) and Caliendo *et al.* (2020) are recent examples of papers providing estimates for this elasticity .

[30] An alternative presentation of the nested logit model, featured in Train (2003) and Cameron and Trivedi (2005) for instance, imposes $\lambda_1 = 1$. In most cases, this normalization is natural since λ_1 and λ_2 are impossible to identify separately. In our case however, the lower-level equation has a natural variable entering with unitary elasticity in the indirect utility: income per capita. This allows separate identification of both parameters, which is important for the counterfactual analysis and the study of complex "substitution patterns". For example, considering whether a change of attractiveness in one country d diverts migrants mostly from alternatives d' or from Germany will be driven by the values taken by those two parameters.

nents: $\ln \tau_{idt} \equiv \ln \tau_{rdt} + \ln \tau_{bdt}$ and $\texttt{Diasp}_{n(i)dt} \equiv \texttt{Diasp}_{rdt} + \texttt{Diasp}_{bdt}$. The first component is related to the easiness to move from city of residence to country d. The second component allows for individuals born in different towns to have access to varying levels of information about country d and exhibit different levels of "proximity" with it. We can re-express the lower-level observed part of utility for an individual born in b and living in r as:

$$V_{rbdt} = A_{dt} - \ln \tau_{rdt} - \ln \tau_{bdt} + \alpha_1 \times \texttt{Diasp}_{rdt} + \alpha_2 \times \texttt{Diasp}_{bdt}. \tag{8}$$

The probability of selecting $d \neq$ DEU conditional on migrating becomes:

$$\text{Prob}\left(d \mid \text{mig}_{it} = 1\right) = \exp\left(V_{rbdt}/\lambda_2 - I_{rbt}\right), \quad \text{with} \quad I_{rbt} \equiv \ln \sum_{d' \neq \text{DEU}} \exp(V_{rbd't}/\lambda_2). \tag{9}$$

We obtain the empirical counterpart of this probability by measuring at the rbt-cell level the share of migrants who fled to d rather than to another country outside Germany.[31] Following this logic and combining equations (8) and (9), we obtain a triadic gravity regression for the expected share of rb migrants going to a specific country as:

$$\mathbb{E}\left[\frac{\text{mig}_{rbdt}}{\text{mig}_{rbt}}\right] = \exp\left(\frac{A_{dt} - \ln \tau_{rdt} - \ln \tau_{bdt} + \alpha_1 \texttt{Diasp}_{rdt} + \alpha_2 \texttt{Diasp}_{bdt}}{\lambda_2} - I_{rbt}\right), \tag{10}$$

where mig_{rbdt} is the yearly flow of migrants from cell rb to country d; mig_{rbt} is the yearly total outflow from rb. The variable $\texttt{Diasp}_{rdt}$ captures the cumulative flows of individuals from r who have migrated to d until year $t-1$; the one-year lag is aimed at mitigating simultaneity bias. The measurement of $\texttt{Diasp}_{bdt}$ follows the same logic. Empirically, we will capture A_{dt}/λ_2 and I_{rbt} with two sets of fixed effects,[32] and the migration frictions with geodesic distances, such that $\ln \tau = \rho \ln \texttt{dist}$.[33] The resulting estimating equation is:

$$\mathbb{E}\left[\frac{\text{mig}_{rbdt}}{\text{mig}_{rbt}}\right] = \exp\left(\text{FE}_{dt} - \tilde{\rho}_1 \ln \texttt{dist}_{rd} - \tilde{\rho}_2 \ln \texttt{dist}_{bd} + \tilde{\alpha}_1 \texttt{Diasp}_{rdt} + \tilde{\alpha}_2 \texttt{Diasp}_{bdt} + \text{FE}_{rbt}\right). \tag{11}$$

The tilde on coefficients α and ρ denotes that the structural parameters driving frictions and diaspora effects are divided by λ_2 when considering the impact on migration flows, i.e., $\tilde{\alpha}_1 \equiv \alpha_1/\lambda_2$. There are two sets of fixed effects. The first set, FE_{dt}, is defined at the destination $\times$ year level;

[31] It is important to note that we define the destination as the first emigration destination after leaving Germany. For a small fraction of individuals (8%), the data reports in addition a second destination. Online Appendix Section G.8 explores the timing of first and secondary migration movements, and the countries concerned. An important finding is that we only observe few people moving twice the same year, suggesting that the first destination in our data represents not just a transitory country, but the outcome of a real choice.

[32] We restrict the sample to the 35 destinations that belong (at least once in the sample period) to the set of countries that make up 95% of total migration. We use the estimated A_{dt}/λ_2 fixed effects to estimate the migration cost elasticity in an auxiliary regression. For those fixed effects to be comparable, destinations need to belong to the "largest connected set" (connections occurring because rb cells do send migrants to several countries every year). This restriction reduces the number of destinations further to 29.

[33] In our baseline analysis, τ_{rdt} (τ_{bdt}) is measured with the distance from the city of residence/birth to the closest point along the border of destination d, defined as of February 28th, 1938. We also allow the distances to have different effects in each year, by interacting (log) distance with the time dummies.

it captures the attractiveness of each destination country ($\text{FE}_{dt} = A_{dt}/\lambda_2$).[34] The second (high-dimensional) set, FE_{rbt}, is crucial for alleviating a source of estimation bias coming from what the gravity literature refers to as multilateral resistance (see Head and Mayer (2014) for a survey). Comparing equations (9) and (11) reveals that the latter fixed effects have a structural interpretation as the inclusive utility times -1, i.e., $\text{FE}_{rbt} = -I_{rbt}$. They therefore capture the expected utility from the lower-level decision, by accounting for the fact that once the migration decision is made, individuals from rb at that time t will choose the best destination available outside of Germany. It accounts in particular for the spatial distribution over destinations of people from the community that have emigrated since then.

The triadic gravity regression corresponds to our econometric implementation of the lower-level decision model. Estimating equation (11) raises several empirical issues. One that has attracted a lot of attention relates to whether the researcher should simply take logs of equation (11) and run OLS, or whether to estimate it in natural form using the Poisson Pseudo-Maximum Likelihood estimator recommended by Santos Silva and Tenreyro (2006). The latter method is more robust to potential heteroskedasticity in the error term, which was the main point of Santos Silva and Tenreyro (2006). It is also the consistent estimation procedure of equation (11). Indeed, when a multinomial discrete choice model such as equation (6) contains no covariate with a chooser (i) $\times$ choice (d) dimension, Guimaraes *et al.* (2003) established that it can be estimated with either conditional logit or PPML with identical results. Finally, PPML is also a natural estimator when so many observations have an observed value of zero (in our sample, about 98% of the potential combinations of rb and d have zero migration flows).[35] Accounting for the zeros is particularly important in the two steps approach we follow here. With no outmigration from a cell rbt, we cannot estimate FE_{rbt} and therefore cannot either recover directly the associated inclusive utility. However, we can still use estimates from the gravity regression and combine them with the observables to reconstruct the inclusive utility according to its definition in equation (13).[36]

A potential issue associated with the existence of zero flows is selection bias. Indeed, endogenous selection into positive flows might generate a correlation between the observable determinants and the error term in the regression. To take one example, since networks of previous migrants lower migration frictions, destinations with low diaspora levels should be associated with a positive shock on the unobservable taste for the migrants-destination pair, explaining that we observe this flow. This will tend to bias the diaspora effects downwards. One solution proposed in the literature (Mulligan and Rubinstein, 2008; Paravisini *et al.*, 2015) is to focus on the part of the sample that is "far from selection", i.e., for which the impact of the idiosyncratic term should be smaller. In our case, we implement this approach by verifying that our main results hold when

[34]Throughout the paper, for the sake of notational clarity, we do not use a specific notation for distinguishing between a theoretical parameter and its point estimate. The reason is that we reserve the use of hat-notation, $\hat{}$, for denoting a different type of variable, namely counterfactual changes in Online Appendix Section H.1.

[35]The combination of cities of residence and cities of birth with the 29 destination countries result in a total of 854,101 $rbdt$ cells, for which there has been at least one migrant from rb in year t to any destination. If nobody migrates from rb in t, the left-hand side variable, the share of migrants to d is undefined. Out of those, 39,899 (or 4,6%) of cells experienced a positive migration flow to country d in year t, for the remaining we impute a zero-migration flow.

[36]Note that the properties of PPML ensure that the two ways to recover the inclusive utility yield identical results when the flow is not zero, as noted by Fally (2015).

restricting the sample to the cities that send a large enough number of migrants every year.

We now turn to the estimation of the upper-level model that we can estimate as a binary logit of the migration decision at the individual level. To this purpose, we simply start from the logit form in equation (7) that we combine with the observed utility from equation (4):

$$\text{Prob}\left(\text{mig}_{it}=1\right)=\Lambda\left(\frac{\lambda_2}{\lambda_1}I_{rbt}+\check{\beta}\texttt{Mignet}_{n(i)t}+\check{\gamma}\texttt{Detainment}_{n(i)t}+\mathbf{X}'_{it}\check{\delta}+\mathbf{FE}\right), \tag{12}$$

where $\Lambda(x)=1/(1+\exp(-x))$, and structural parameters are now scaled by λ_1, i.e., $\check{\beta}\equiv\beta/\lambda_1$. The inclusive utility I_{rbt} is generated using the right hand side of equation (9):

$$I_{rbt}=\ln\sum_{d\neq\text{DEU}}\exp\left(\text{FE}_{dt}-\tilde{\rho}_1\ln\texttt{dist}_{rd}-\tilde{\rho}_2\ln\texttt{dist}_{bd}+\tilde{\alpha}_1\texttt{Diasp}_{rdt}+\tilde{\alpha}_2\texttt{Diasp}_{bdt}\right), \tag{13}$$

and accounts for all determinants in the expected utility gains of migrating that come from the choices of destination country. For instance, upward or downward swings in the business cycle of France compared to the Netherlands will be captured in I_{rbt}, since it uses the dt fixed effects which capture all potential attractiveness factors common across migrants when choosing country d in year t. The coefficient on the inclusive utility is λ_2/λ_1. It informs us both about the validity of the decision-tree structure assumed, which is not rejected as long as the ratio is smaller than one, and on the upper-level outmigration elasticity λ_1 (since the lower level gives us an estimate of λ_2).

Our econometric equation (12) uses the structure of the model to distinguish and quantify the three channels of social interactions that we emphasized in the introduction. The threat effect is captured by the effect of $\texttt{Detainment}_{n(i)t}$. The exodus effect is measured by the coefficient $\check{\beta}$ on $\texttt{Mignet}_{n(i)t}$, while the diaspora effect is channeled through the impact of the lower-level inclusive utility, I_{rbt}. In the counterfactual analysis, the latter two migration spillovers are likely to dynamically amplify the initial impact of a change in the immigration policy of a destination country.

6.2 Lower Model Estimation

We now turn to estimate the triadic gravity model shown in equation (11). The most important variables for our purpose are the ones capturing diaspora effects ($\texttt{Diasp}_{rdt}$ and $\texttt{Diasp}_{bdt}$). The literature has followed several routes to measure the effects of previous migrations on current flows. The simplest approach is to consider $\text{stock}_{d,t-1}$, the cumulative stock of peers that chose d until year $t-1$, taking logs of the (non-zero) stock to account for the multiplicative nature of the gravity regression. A problem with this approach is that those counts are often zero until a certain date.[37]

[37] This problem of zeros in the diaspora variable is particularly severe in our case, since we work with a high degree of spatial detail (rather than national flows); annual migration flows (rather than commonly used 10-year windows); and a relatively small initial population at risk of migrating. We only include cities-of-residence r and cities-of-birth b if the number of Jews in 1933 living in r, or originating from b, is positive, and if we have information on the migration date of at least one adult individual from this rb cell for the overall period. The average size of the Jewish community in 1933 in the city of residence, pop_{r,t_0}, is 10,361 individuals, and the mean of Jewish communities in 1933 in cities of birth, pop_{b,t_0}, is 507. In the end, 45% of the $rbdt$ cells used in the lower-level estimation display a non-zero diaspora from the city of residence r in year t. The figure is 20% for non-zero diaspora cells from the city of birth b in year t.

Hence, many papers have chosen to measure $\texttt{Diasp}_{rdt}$ as $\ln(1 + \text{stock}_{d,t-1})$. This $+1$ can be rationalized by the fact that the potential migrant is considering its own addition to the observed stock of migrants. However, this functional form is distorting the distribution of the variable, in particular when stocks are low. An alternative is to consider the stock of migrants in levels rather than in logs, or its relative level, i.e., the cumulative share of migrants from r that chose d until year t (not included). Our measures are therefore written as:

$$\texttt{Diasp}_{rdt} \equiv \frac{\text{stock}_{rd,t-1}}{\text{pop}_{r,t_0}} \qquad \text{and} \qquad \texttt{Diasp}_{bdt} \equiv \frac{\text{stock}_{bd,t-1}}{\text{pop}_{b,t_0}}, \tag{14}$$

where pop_{r,t_0} is the observed population of Jews in city of residence r, measured at the start of the sample (1933) and $\text{stock}_{rd,t-1}$ is the cumulative stock of Jewish residents that chose d until year $t-1$. These variables go from 0 (before any peer from same city of residence or birth has moved to d) to (almost) 1 if all peers have already moved to d. However, this functional form does not impose a constant elasticity to the impact of peers' previous moves (because it takes levels rather than logs), which makes coefficients harder to interpret. Since these variables are critical in our estimations and counterfactual exercises, we organize our results mostly around the different approaches to the measurement of the diaspora effect. In order to ease comparison, the regression tables will systematically report the average elasticity of the probability to choose d with respect to the two diaspora variables.

The triadic gravity model is estimated over 1933-1941 on an aggregated version of the F sample (see Online Appendix Section A.3.3). Counts of migrants are summed at the residence-birth-destination-year level to construct the share of migrants by destination, $\text{mig}_{rbdt}/\text{mig}_{rbt}$, which is the dependent variable in equation (11). Standard errors are clustered at the city-of-residence $\times$ city-of-birth $\times$ year level. As our model recommends, all regressions control for destination $\times$ year, and city-of-residence $\times$ city-of-birth $\times$ year fixed effects, which requires the use of high-dimensional panel data estimation techniques.[38] The migration gravity literature frequently uses shift-share instruments to address concerns about endogeneity of diasporas. In our setup, this would require some measure of the pre-1933 bilateral jewish migration stock between each German city and each destination country, which is not available. Given that our analysis uses historical data at a very granular level, there is no easy alternative. However, we see three reasons why, in our case, the concern should be limited. First, the rich structure of fixed effects that we allow for should do a good part of the job of filtering out unobserved heterogeneity. In particular, we control for destination $\times$ year fixed effects: The destination fixed effects for 1933 capture the (unobserved) initial stock of diasporas in each destination and more generally the pre-sample attractiveness of a country for the German Jewish community. Hence, the remaining problematic component pertains to city-to-country attractiveness factors. There, the overall connectivity of a city to the rest of the world is already captured by city-of-residence $\times$ city-of-birth $\times$ year fixed effects. Moreover, although we have no data on the quality of bilateral transport infrastructure at the city $\times$ destination level, we

[38] The use of linear multi-way fixed effects packages such as `reghdfe` is now standard. The econometric procedure we use for the high-dimensional fixed effect PPML estimation is `ppmlhdfe`, recently developed by Sergio Correia, Paolo Guimaraes and Thomas Zylkin for Stata.

control for distance from cities to each specific destination and this captures the likelihood of having a good bilateral transport connection between German cities and a specific destination country. Second, existing papers looking at the impact of diasporas on destination choice in migration do not find a major difference between instrumented and non-instrumented results (see Beine *et al.* (2011) for instance). Third, we find below that our estimates of the diaspora coefficients are quite close to the ones found in the literature. This fact reassuringly suggests that our estimation is not contaminated by pervasive endogeneity biases.

Table 4 displays the results. Column 1 follows a classical setup in empirical gravity equations, taking logs of equation (11), and running OLS on the sample of *rbdt* cells with non-zero migration flows. In this column, the diaspora variable takes the often-used functional form of log of 1 + cumulative counts of migrants. Column 2 uses our preferred measure of diasporas with same estimation method as column 1. Columns 3 and 4 turn to PPML regressions, first on positive flows, and then on the entire sample including *rbdt* cells with zero migration flows. Columns 5 replicates the regression of column 4, but adds interaction terms between distances and year dummies to evaluate the evolution of the effect of distance over time.[39] Since column 5 allows for more flexibility in the impact of migration frictions, we take its set of coefficients and parameters as our baseline for computation of the inclusive utility. Finally, columns 6 and 7 replicate the regressions of columns 3 and 4, restricting the sample to the top 100 cities in terms of the population at risk (both in terms of birth and residence cities) in order to limit selection bias concerns (see Section 6.1 for a more detailed discussion).

Migration costs. Regarding the influence of migration costs, it is reassuring that the coefficients on the distances from the city of residence/birth to the different destination countries are negative and significant in almost all columns, since it confirms the large literature that has estimated migration gravity regressions on modern times samples (Beine *et al.*, 2016, being a good survey of that literature). As expected, the location of the city of residence is more important than the location of the birth city: The coefficient in column 2 implies that a 1 percent increase in the distance from the city of residence (birth) is associated with a 0.096 (0.011; not significant) percent decrease in the share of migrants. Turning to the PPML estimation technique on column 3, while keeping the same sample of strictly positive flows, does not change massively the impact of distance. The distance estimates for city of residence to destination in columns 1 and 2 are smaller (in absolute value) than the elasticities obtained on more recent samples (Beine *et al.*, 2011; Ortega and Peri, 2013; Bertoli and Moraga, 2013, 2015). This might be because we are here mostly identifying out of internal distances from different parts of Germany to contiguous countries (since distance to the USA is roughly constant across German cities), combined with a different time period. Note that the impact of distance becomes very much in line with findings in the literature when accounting for zeros in columns 4, 5 and 7.

[39]Overall, results are not affected very much by this different treatment of distance. In Online Appendix Figure G11 we report the coefficients of these interaction terms: Distance effects vary across years but their overall time-series profile is rather flat. Online Appendix Table G13 replicates all columns Table 4 with distance-year interactions.

Table 4: Triadic Gravity Estimation (Lower Stage)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variables:	Share Migrants (log)	Share Migrants (log)	Share Migrants > 0	Share Migrants	Share Migrants	Share Migrants > 0	Share Migrants
	All Cities					Top 100 Cities	
	OLS	OLS	Poisson	Poisson	Poisson	Poisson	Poisson
log (Cumulative # Migrants_{rdt}+1)	0.085***						
	(0.009)						
log (Cumulative # Migrants_{bdt}+1)	0.168***						
	(0.012)						
Diaspora_{rdt} = Cumulative Share of Migrants_{rdt}		0.435*	0.487**	5.250***	5.435***	1.713***	6.268***
		(0.233)	(0.192)	(0.205)	(0.209)	(0.557)	(0.716)
Diaspora_{bdt} = Cumulative Share of Migrants_{bdt}		2.083***	1.315***	5.184***	5.294***	3.614***	9.399***
		(0.262)	(0.209)	(0.207)	(0.208)	(0.609)	(0.940)
(log) Distance_{rd}	-0.073***	-0.096***	-0.076***	-0.460***		-0.063***	-0.376***
	(0.013)	(0.013)	(0.011)	(0.013)		(0.019)	(0.026)
(log) Distance_{bd}	0.001	-0.011	-0.014*	-0.123***		-0.023	-0.083***
	(0.009)	(0.009)	(0.008)	(0.011)		(0.014)	(0.025)
(log) Distance_{rd} $\times$ Year					✓		
(log) Distance_{bd} $\times$ Year					✓		
Observations	15,010	15,010	15,010	839,402	839,402	7,664	144,920
Observations incl. singletons	39,899	39,899	39,899	854,101	854,101	11,086	150,819
R-squared	0.82	0.81					
Pseudo R-squared			0.75	0.18	0.19	0.74	0.18
Avg. Elasticity Diaspora_{rdt}	0.085	0.010	0.011	0.070	0.073	0.028	0.064
Avg. Elasticity Diaspora_{bdt}	0.168	0.057	0.036	0.130	0.132	0.067	0.112
Mean of dependent var.	-1.462	-1.462	0.317	0.035	0.035	0.240	0.036
SD of dependent var.	0.973	0.973	0.190	0.176	0.176	0.186	0.164
No CoR	599	599	599	1,956	1,956	75	98
No CoB	1,303	1,303	1,303	3,510	3,510	100	100

NOTE: The unit of observation is a city-of-residence x city-of-birth x country in year t. The sample consists of individuals who were at least 16 years old in 1933, includes 29 destination countries, and the estimation period is 1933-1941. Diaspora_{rdt} (Diaspora_{bdt}) is defined as the cumulative share of migrants from a city of residence (city of birth) to the destination country. All regressions control for country $\times$ year fixed effects, and city-of-residence $\times$ city-of-birth $\times$ year fixed effects. Standard errors clustered at the city-of-residence $\times$ city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Turning to our main variables of interest, we find that, in all specifications, the diaspora networks from the city of residence and city of birth have a positive effect on the choice of destination: Although the effects are less precisely estimated in columns 2 and 3, both diaspora networks have a large and statistically significant impact in our preferred specifications of columns 4 and 5 (PPML with zero flows). Interpreting those variables is more involved than for distance. The simplest case is column 1. Since we log the diaspora variables, the coefficients are elasticities with respect to that the stock of migrants *after adding one*. Those elasticities are again smaller than in the recent literature (Beine *et al.*, 2016, refer to a 0.4 elasticity as being consensual). Measurement issues in our sample (remember that we can only start the stock in 1933), combined with a very different context, could explain the discrepancy. An additional difference is that we consider annual migration flows, whereas the literature cited in Beine *et al.* (2016) mostly uses *decadal* migration flows. As frequent in the literature, the diaspora variable in column 1 adds 1 to the stock to ease the interpretation of the coefficients. Our preferred specification relies on shares (equation 14). The associated average elasticities, reported in the bottom of the table and calculated by multiplying the estimates (semi-elasticities) by the mean value of the relevant variables (when positive), are lower, but still significantly positive. The results of both columns 4 and 5 imply that a 10-percent increase in the cumulative stock of migrants from the city of residence having moved to country *d* increases the proportion of migrants further choosing country *d* by 0.7%. The effect of networks from city of birth is larger with a elasticity close of 1.3%. As for the impact of distance, the two diaspora coefficients are sensitive to the inclusion of zero migration flows in the regression, as shown by comparing columns 3 and 4. This calls for a detailed investigation of the issue.

Zero migration flows, limited mobility bias, and selection. In our context, the inclusion of cells with zero migration flows is particularly important. Since we work with a discrete choice framework, all coefficients are identified out of variation in the characteristics of choices available to the chooser. The chooser here is a *rbt* combination. The fact that no migrant went from a *rbt* cell to a given country is informative about the underlying attractiveness of this country. Another way to put it is that all choosers face the same choice set.

Furthermore, including the cells with zeros helps with the proper estimation of the destination-time fixed effects, FE_{dt}. As emphasized in the employer-employee literature, estimation of multi-way fixed effects relies critically on "connectivity" (Abowd *et al.*, 1999; Andrews *et al.*, 2008; Card *et al.*, 2013). In our context, if we consider only positive flows, a country needs to be chosen by several *rb* that themselves chose several destinations in that year. If there are very few of such cells, the estimate of FE_{dt} will be noisy, which is an incarnation of the limited mobility bias emphasized by labor economists. Including the zeros in the choice set increases connectivity, since the absence of migrants in a particular destination provides a valid comparison point to positive flows.[40] Restating our migration problem as a bipartite network of cities (choosers) and countries (choices), the connectivity measure proposed by Jochmans and Weidner (2019) reaches its maximum value when including the observed zero migration flows to the matrix.

Online Appendix Table G15 illustrates the point. For every year and destination in our sample, we compute the number of "connecting" *rb*, that is the chooser cells that actually sent migrants to at least two destinations this given year. For many countries and years, such cells are rare, even in the years when migration is stronger, raising concern for limited mobility bias. With zeros, this table would constitute of columns uniformly filled with the total number of *rb* that year. The consequence of including the zeros can also be seen in Online Appendix Figure G12 where we plot the rank of estimated FE_{dt} against the rank of a country in the total share of Jewish migrants in a given year (constructed such that a higher value means more attractiveness). Panel (a) shows the scatter plot corresponding to column 3 of Table 4, while panel (b) presents results from column 4. The version including the zeros corresponds to a better fit with less variance as expected from the insights of the labor literature. Including the zeros in columns 4 and 5 are therefore our preferred specifications.

Columns 6 and 7 of Table 4 reduce the sample to the set of birth and residence cities that are

[40]Note that this is due to the fact that a zero in our setup is a "true" null flow. This is a notable difference with the employer-employee case where we do not actually know what the wage of an individual would be in a firm with which the individual did not actually match.

large enough that the selection into positive flows should be a small concern. As expected from our discussion in Section 6.1, this sample restriction alleviates the (downward) selection bias and therefore leads to an increase in the absolute value of coefficients. However, the increase is much less pronounced in the case where the regression includes the zeros (column 7 compared to column 4). Furthermore, the elasticities reported in the bottom of the table are very close between these columns 7 and 4. This stability reinforces our decision to consider the specifications of a PPML regression with zeros included as our preferred ones.

Revealing the attractiveness of countries. From the gravity estimation of column 5 in Table 4, we recover the estimated fixed effects FE_{dt}, which measure the attractiveness of destination d in year t. Panel B of Online Appendix Table G14 ranks the top 10 destinations in each year, as revealed by these fixed effects. The ranking of those country $\times$ year fixed effects follows fairly closely the list of top destinations in terms of observed migration flows displayed in panel A of the same table.[41]

Table 5: Determinants of Country Attractiveness

	(1)	(2)	(3)	(4)
Dependent variable:	Country Attractiveness (FE_{dt})			
(log) GDP per capita	3.270*	3.499**	2.951*	2.600*
	(1.747)	(1.610)	(1.480)	(1.331)
(log) GDP per capita $\times$ Post 1938		-1.231**	-1.124**	-1.148**
		(0.518)	(0.522)	(0.522)
(log) Distance $\times$ Post 1938			0.355*	0.171
			(0.190)	(0.192)
German Occupation				-1.508***
				(0.462)
Observations	230	230	223	221
R-squared	0.61	0.63	0.65	0.66
Mean of dependent var.	-2.70	-2.70	-2.65	-2.66
SD of dependent var.	2.16	2.16	2.17	2.18

NOTE: The unit of observation is a destination country-year combination. The dependent variable are the estimated fixed effects FE_{dt} recovered from the estimates in column 5 of Table 4. All regressions control for country and year fixed effects. Standard errors clustered at the country-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Our next step is to assess how revealed attractiveness of a country-year correlates with observables that theory predicts to affect indirect utility. In most micro-foundations of the migration decision, A_{dt} relates to real income per capita of the destination-year combination with a unitary coefficient (Monte *et al.*, 2018; Tombe and Zhu, 2019; Caliendo *et al.*, 2020, are three recent examples). In Table 5, we therefore show results of an auxiliary regression of the estimated fixed effects

[41] The last row of Online Appendix Table G14 shows the pairwise correlation coefficients of the ranks of countries based on their observed migration shares and estimated attractiveness. Those correlations vary over the years but are consistently high, ranging from .66 to .88.

FE_{dt} on the log income per capita of the countries.[42] From equation (10), the structural interpretation of the coefficient is $1/\lambda_2$, which we estimate to be 3.27 in column 1.[43] Note that from the same equation (10), we see that $1/\lambda_2$ is also the migration cost elasticity. This elasticity has been recently estimated by several papers using very similar theoretical motivation: Monte *et al.* (2018) find 3.3, quite close to our value. Tombe and Zhu (2019) preferred estimate for China is 1.5, while Bryan and Morten (2019) report 3.2 for Indonesia, and 2.7 for the USA. Fajgelbaum *et al.* (2018) also use U.S. data to estimate a migration cost elasticity at 1.73. At the other end of the spectrum in that literature, Caliendo *et al.* (2020) find 0.5 on a pan-EU sample. In column 2, we find that the relation between FE_{dt} and income is weaker after 1938 (reduced by about one-third), reflecting the well-documented fact that after the "*Kristallnacht*" economic considerations became much less important for the decision whether and where to go. Consistent with aggregate patterns of migration, we see in column 3 that distance from Germany increased attractiveness after 1938. This is to be expected since, by that time, countries nearby Germany were either already occupied or threatened to be. Column 4 validates this interpretation by adding a dummy that turns on when a country gets occupied by Nazi Germany. This makes the distance effect disappear. Overall, the consistency between the estimated and observed destination attractiveness, and the expected behavior of coefficients with respect to historical facts, suggest that our estimation framework is relevant for the location choices of Jewish emigrants in the inter-war period.

6.3 Outmigration Model

We now turn to the estimation of the upper-level model equation (12). This is the structural version of equation (1), i.e., taking into account the lower-level destination choice through inclusive utility. We construct I_{rbt}, the inclusive utility for individuals living in r and born in b, using estimates from column 5 of Table 4 with formula given in equation (13).

Panel A of Table 6 displays the non-instrumented results; only the variables of major interest are reported. Our structural model calls for the use of a non-linear estimator. We start with the traditional binomial logit (reporting coefficients in column 1 and marginal effects in column 2). Column 3 considers the Cloglog—a duration model—as an alternative. As with the reduced form estimation of Section 5, in columns 4 and 5 we also provide estimates of LPM, which allows for standard treatment of high-dimensional fixed effects and two stage least squares (in panel B, see below). Across all specifications we find that the two migration spillover variables (network migration and inclusive utility) have a statistically significant positive sign. This is a first indication that both the exodus and diaspora effects are at work in the data. As for the threat effect, we see that despite the inclusion of the two migration spillover variables, the impact of past detainment in the network is extremely close to its corresponding reduced-form estimate (comparing column 2 of Table 3 and column 5, panel A, of Table 6, which use identical estimation methods).

Marginal effects are reported in columns 2 and 3, in order to make estimates comparable to

[42]This method is similar to one of the ways Eaton and Kortum (2002) use to recover the trade elasticity in a gravity setup for trade flows.

[43]As explained above (head of Section 6), the concern for endogeneity of the real income per capita of destination countries for Jewish migrants is much less severe than in those papers.

LPM results in columns 4 and 5. Comparing across the last four columns, we see that the method of estimation does not make a large difference regarding the marginal effect of inclusive utility. The structural interpretation of this variable, λ_2/λ_1 from equation (12), is confined to column 1. As discussed above, a theoretical requirement is that $0 \leq \lambda_2/\lambda_1 \leq 1$, ensuring that the assumed tree structure of the location choice is consistent with utility maximization. The theory-consistent estimator of column 1 finds that $\lambda_2/\lambda_1 \simeq 0.75$, confirming that our nested logit structure is compatible with revealed preferences. Finally, combined with our lower-level estimate of $\lambda_2 = 1/3.27 = 0.31$, we can reveal $\lambda_1 = 0.31/0.75 = 0.41$. The magnitude of the estimated coefficients implied by column 5 are large: A one-standard-deviation increase in the migration of network members (0.16), i.e., the exodus effect, increases the annual migration probability by 1.1 percentage points, or 20% of the sample mean. A one-standard-deviation increase in the inclusive utility (0.89), which encompasses not only the diaspora effect but also destinations' attractiveness and migration frictions, increases the annual probability of migration by 1.3 percentage points, or 24% of the sample mean, according to column 5.

Instrumentation. By construction, individuals and their distant peers originate from the same city of birth but do not live in the same city of residence. This network construction ensures that our empirical design is immune to migration shocks that are common across individuals living in the same city. We believe this construction deals with the first-order exogeneity concern in our regressions. However, because decision makers and their peers are born in the same city, there could be some unobserved shocks driving simultaneously their outmigration. For instance, having been exposed to similar secular/religious education could affect the overall propensity to migrate later in life as well as the destination choice. This threatens the exogeneity of the network migration and the inclusive utility variable. We tackle this issue, called *homophily* in the network literature, by building two exogenous shifters of distant peers' migration decisions, which should not be related to the city of birth. The idea is to exploit the push and pull factors that are specific to the *city of residence of distant peers*, which are orthogonal to the direct determinants of migration choice of the decision maker.

The first shifter captures push factors related to persecution, building on the observation that detainment in the city of residence positively impacts outmigration (see our discussion of the estimates in column 1 of Table 3). For an individual i, it is defined as the average past detainment share in the residence cities $r(j)$ of her distant peers j up to year $t-1$:

$$\texttt{Push}_{i,t} \equiv \sum_{1933 \leq s < t-1} \frac{1}{N_{i,s}} \times \left[\sum_{j \in n(i,s)} \texttt{Detainment}_{r(j)s} \right], \tag{15}$$

where $n(i,s)$ is the network of $N_{i,s}$ distant peers still living in Germany in year s.

The second shifter relates to the pull factors affecting distant peers' migration as captured by their *partial* inclusive utility, namely the components of peers' I_{rbt} that neither relate to their city-of-birth nor to diasporas in equation (13). For each distant peer $j \in n(i,s)$, we retrieve from the gravity estimates her partial inclusive utility and we average it across distant peers who still live

Table 6: Outmigration Decision with Inclusive Utility: 1933–1941

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Migration Decision				
Panel A: Non-Instrumented	Logit coef.	Logit dy/dx	Cloglog dy/dx	LPM coef.	LPM coef.
Detainment of network members	0.480***	0.023***	0.023***	0.042***	0.042***
	(0.143)	(0.007)	(0.007)	(0.012)	(0.010)
Migration of network members	0.378***	0.018***	0.014***	0.074***	0.071***
	(0.089)	(0.004)	(0.004)	(0.009)	(0.009)
Inclusive Utility	0.747***	0.036***	0.034***	0.040***	0.015***
	(0.043)	(0.002)	(0.002)	(0.003)	(0.004)
Pseudo R-squared	0.13	0.13			
R-squared				0.07	0.10
Panel B: Instrumented	Logit coef.	Logit dy/dx	Cloglog dy/dx	LPM coef.	LPM coef.
Detainment of network members	0.489***	0.023***	0.023***	0.044***	0.044***
	(0.152)	(0.007)	(0.007)	(0.013)	(0.011)
Migration of network members	1.864***	0.089***	0.074**	0.140**	0.184***
	(0.710)	(0.034)	(0.033)	(0.061)	(0.049)
Inclusive Utility	0.811***	0.039***	0.038***	0.035***	0.016
	(0.161)	(0.008)	(0.007)	(0.012)	(0.015)
R-squared				0.02	0.02
Pseudo R-squared	0.13	0.13	.		
Year FE	✓	✓	✓	✓	
CoR × Year FE					✓
Observations	974,308	974,308	974,308	974,308	972,603
Mean of dependent var.	0.055	0.055	0.055	0.055	0.055
SD of dependent var.	0.227	0.227	0.227	0.227	0.227
F-Stat				22.17	55.69

NOTE: The dependent variable is an indicator for migration in year t. The unit of observation is an individual in year t. The sample consists of individuals who were at least 16 years old in 1933. Migration of network members measures the cumulative share of network members that emigrated until year $t-1$. Detainment of network members measures the cumulative share of network members that have been detained until year t. All regressions control for age, age squared, gender, personal detainment, and a dummy for whether the individual was born outside Germany, as in Online Appendix Table G7. Marginal effects reported in columns 2 and 3. Standard errors clustered at the city-of-birth × year in parentheses. Panel B: Columns 1, 2 and 3 use control function approach, while columns 4 and 5 estimate two-stage least squares. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in Germany up to year $t-1$ to generate the second shifter:

$$\texttt{Pull}_{i,t} \equiv \sum_{1933 \leq s < t-1} \frac{1}{N_{i,s}} \times \left[\sum_{j \in n(i,s)} I^{\text{p}}_{r(j)s} \right] \quad \text{where,} \quad I^{\text{p}}_{r(j)s} \equiv \ln \sum_{d' \neq \texttt{GER}} \exp\left(FE_{ds} - \tilde{\rho}_{1s} \ln \texttt{dist}_{r(j)d} \right). \tag{16}$$

We use $\texttt{Push}_{i,t}$ and $\texttt{Pull}_{i,t}$ as exogenous shifters of $\texttt{Mignet}_{n(i)t}$ and I_{rbt} in our structural equation (12). These two variables capture exogenous changes in the relative attractiveness of Germany compared to the rest of the world for the distant peers of decision makers. Both variables induce peers' migration which impacts directly $\texttt{Mignet}_{n(i)t}$ and indirectly I_{rbt} (via $\texttt{Diasp}_{bdt}$ in equation 13).[44]

Panel B of Table 6 presents the results from instrumented specifications. For non-linear estimators (Logit and Cloglog), we use a control function approach (Cameron and Trivedi, 2005). In the last two columns, instrumented LPM is estimated with standard two-stage least squares that allow for testing for weak instruments (Kleibergen-Paap F-statistics). Moreover, 2SLS provide appropriately corrected standard errors. Note that the statistical level of significance for the three variables of interest obtained with 2SLS reassuringly stays in line with the ones from non-linear estimators. First-stage estimation results are displayed in Online Appendix Table G17 and confirm the statistical power of both instruments.[45]

In all specifications we estimate coefficients of network migration and inclusive utility that are positive and statistically significant. The first-stage F-Statistics reported at the bottom of the table in column 4 (22.17) and column 5 (55.69) underline the relevance of the instrumental variables. Compared to their non-instrumented counterparts (panel A), the instrumented point estimates of inclusive utility (panel B) are quite stable. In terms of the theory, the ratio λ_2/λ_1 is recovered from column 1. Combined with the value of λ_2 this reveals that $\lambda_1 \simeq 0.31/0.81 = 0.38$. The magnitude of the effect of network migration is more sensitive to instrumentation. This will matter for the structural interpretation of those coefficients. We will therefore run the counterfactual with the two sets of parameters (instrumented/non-instrumented).

Robustness. We now investigate the sensitivity of the estimated migration spillovers, i.e., migration of network members and inclusive utility, to a battery of robustness checks. We benchmark on column 1 of Table 6, panel A. We briefly summarize the checks we perform here and discuss them in detail in Online Appendix Section G.13. In Table G18, we investigate how sensitive our estimates are to the age of decision makers. In Table G19, we relax the 5-year age bracket between decision makers and their distant peers when building social networks. In Table G20, we change the minimal number of distant peers required for the computation of network measures. In Table G21, we explore the sensitivity of the estimates to using the final rather than the first migration destination, and to considering a previous city of residence instead of the last known city of residence. In Table G22, we test the robustness of our findings to excluding decision makers who migrated in early years of the Nazi rule. In Table G23, we explore how missing information on migration and

[44]The validity of the instruments relies on the assumptions that push and pull affect an individual's migration decision only through the *actual* migration of their peers and inclusive utility (which captures the diaspora effect). In other terms, we assume that individuals (i) base their own migration decision on "hard facts" about the migration/detainment of their (first-degree) distant peers; (ii) but do not react to "soft information" about the persecution and migration of second-degree peers, i.e., co-residents of their distant peers who to them are strangers.

[45]We present the intention-to-treat (or reduced-form) results where the two endogenous network effects are replaced by their shifters in Online Appendix Table G16. The coefficients of the shifters capture the network externalities driven by push/pull factors. They both load positively confirming that fiercer persecution and/or better migration prospects in the cities of residence of distant peers increase the propensity to migrate of individuals. While the intention-to-treat approach is immune to potential violations of the exclusion restriction, it does not disentangle the migration spillovers.

deportation status for some distant peers affects the estimated effects of migration spillovers. In Tables G24 and G25, we check the robustness of our estimates to restricting our sample to a subset of residence cities whose city-level population figures match better external sources (Alicke (2014) and the Census of 1925). In Tables G26 and G27, we drop cities of residence and birth from the sample if they are in the lower and upper tails of the population distribution. In Table G28, we only consider movers as decision makers, and impose a minimum distance between the cities of residence of decision makers and their distant peers.

Figure 4 summarizes the point estimates from a total of 45 robustness checks for the two network migration variables. Each light gray diamond represents the point estimate along with the 90% confidence intervals obtained from a robustness specification; the dark blue circles represent our baseline point estimates that we benchmark on. Overall, the estimated effects are very stable in magnitude and significance across specifications. Our preferred specification from column 1 of Table 6 yields point estimates that are located around the median of the distribution of estimates.

Moreover, we assess how missing information on migration and deportation trajectories affects estimates in Online Appendix Section G.14. We weight the F sample such that it matches the observable characteristics of out-of-sample individuals, i.e., individuals for which the dependent variable of the model cannot be constructed due to missing information. Our weighting approach builds upon the entropy balancing method introduced by Hainmueller (2012). Results are displayed in Table G31 (logit) and Table G32 (LPM) for several weighting schemes. All the estimated coefficients of the weighted regressions are close to their baseline counterparts. Hence, restoring balancedness between in-sample and out-of-sample observations does not change meaningfully the point estimates.

Last, we perform a placebo test, in which we reshuffle observed networks across decision makers. To each decision maker, we randomly assign all three network variables which are observed for another individual, and re-estimate the specification reported in column 1 of Table 6, panel A. Online Appendix Figure G13 presents the coefficients of the network migration variables obtained from 1,000 repetitions. This figure shows that the likelihood of obtaining the benchmark estimates by chance is less than one in a thousand.

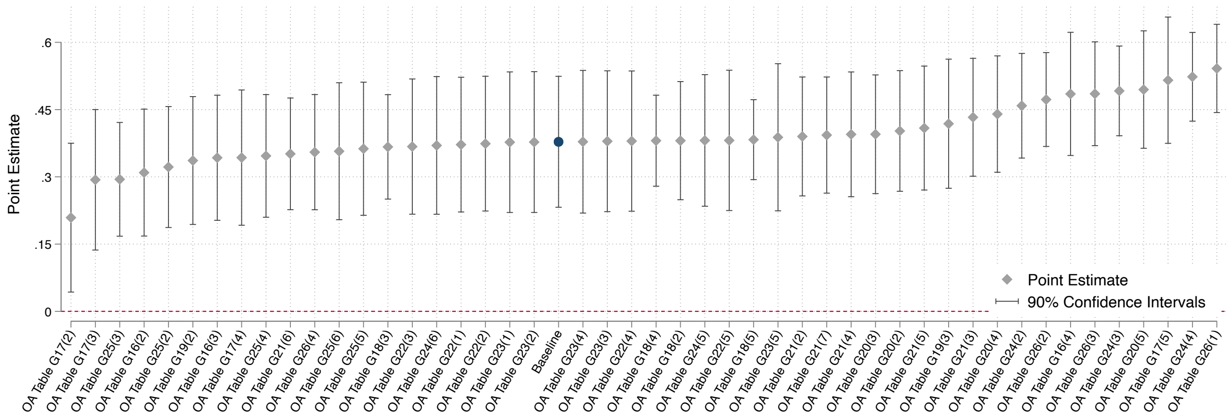
Unobserved individual characteristics. In our data, we have sparse information on individuals' characteristics. In particular, we do not observe wealth, income, and education. Economic means were a likely, *but ambiguous*, determinant of migration decisions in Nazi Germany (see footnotes 4 and 11). In addition, we do not observe other important drivers of migration at the individual level such as political ideology, e.g., Zionist or Socialist, and religiosity.[46]

What are the implications of this data limitation for our empirical design? First, it could have some consequences for the interpretation of our main coefficients of interest. Indeed, our estimates of the migration spillovers are to be interpreted as an *average* marginal effect across individuals.

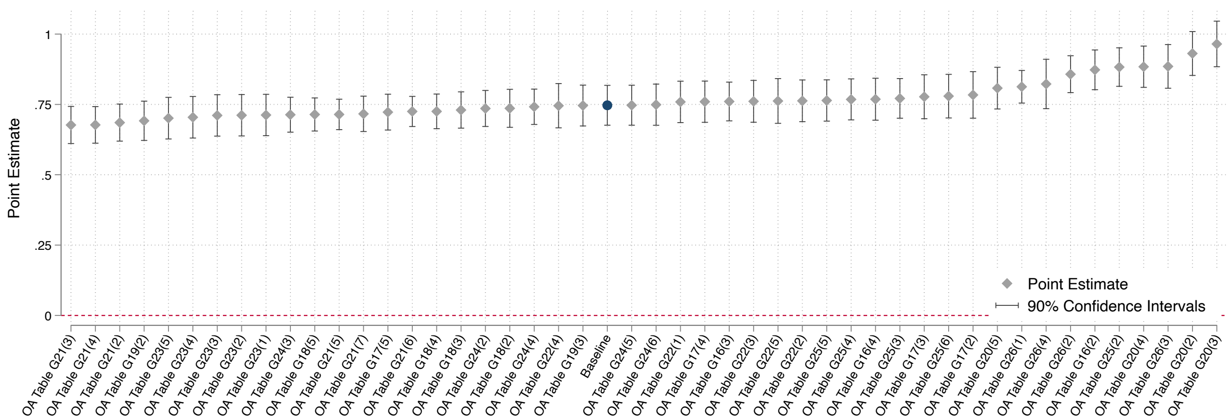
[46]It is likely that Zionist and left-wing individuals migrated earlier (Rosenstock, 1956). In Online Appendix Table G22, we exclude decision makers who migrated in early years of the Nazi Rule. Moreover, using the Jewishness of first name as a proxy for religiosity, we explore in Table 3 the differential response of individuals with varying levels of religious identity towards the threat of persecution.

Figure 4: Summary of Point Estimates on Migration Spillovers from Robustness Checks

(a) Migration of network members



(b) Inclusive utility



NOTE: This figure visualizes the point estimates along with the 90% confidence intervals obtained from the robustness checks presented in Online Appendix Section G.13. Each diamond and circle sign represent point estimates obtained from a different regression. On the x-axis we indicate the Online Appendix Table (G17–G26) and column (1–7) that contains the corresponding point estimate.

Marginal effects might vary across individuals as their decision to migrate can be more or less elastic to their peers' migration. How heterogeneous is this elasticity across individuals? Is the average marginal effect mostly driven by wealthy, educated people? Our data do not enable us to shed a definitive light on these questions, and we can only note that the answers are *a priori* unclear, given the ambiguous relationship between economic means and migration incentives.

Second, if wealth and income are correlated between decision makers and distant peers, nationwide shocks and policies could affect their migration incentives simultaneously. As explained above, our instrumented specifications are well suited for addressing this type of concerns, thanks to the construction of migration shifters that are exogenous to the characteristics of decision makers. For the sake of completeness, we now consider an alternative way of dealing with homophily. To directly account for unobserved heterogeneity across individuals, we re-estimate the outmigration model with individual fixed effects. These fixed effects capture not only heterogeneity in economic means (wealth, income), but all other unobserved heterogeneity at the individual level, for example education, occupations, and membership in political parties. Online Appendix Table G29 reports the LPM results. The findings, in particular for network migration and detainment, are in line with our previous estimates obtained without taking into account unobserved heterogeneity at the individual level. This observation makes us confident that the overall consistency of our baseline estimation is unaffected by the aforementioned data limitation. We also notice that the approach based on individual fixed effects is extremely demanding from the data. Indeed, the identifying variations come only from the subset of migrants as their spells transit from 0 to 1 in the year of their migration; by contrast, spells of stayers are constantly equal to 0 in the panel dimension and cannot be used for the purpose of identification in the within dimension. Moreover, such high dimensions of fixed effects are a challenge for the non-linear estimation of our structural model. For these reasons, our preferred empirical model abstracts from including individual fixed effects.

FP sample. So far, our estimations have considered the sample of individuals with full spells (F sample). We now extend the estimation to the FP sample (for a complete description of how it is built, see Section 3.2.2). We first consider a restricted version of the FP sample by adding (to the F sample) individuals who can be matched with the 1939 Census but have no migration year or deportation information in the Resident List (row 6 of Table 1). We also use a larger version of the FP sample which also adds individuals who died before 1939 and the ones who were expelled to Poland in 1938 (row 7 of Table 1). Importantly, the two versions of the FP sample are only well defined for the period of 1933–1938, since we do not know what happened to people found in the 1939 census after the census too place.

Estimation results are reported in Table 7. Column 1 re-estimates the baseline specification (column 1 of Table 6, panel A) using the F sample but restricting the period to the years 1933–1938. In line with the reduced-form findings reported in Table 3, the coefficients on detainment and migration of network members are greater when we focus on migration decision taken prior to "*Kristallnacht*". We also see an increase in the coefficient on inclusive utility. Those three changes are pointing in the same direction: The observable incentives to out-migrate have a stronger impact in the years when the urgency of leaving was less obvious. In the particular case of the Inclusive Utility, the coefficient gets closer to 1, which indicates a simple conditional logit decision structure, where Germany is one destination among others. In all columns, this coefficient is actually not significantly different from 1 at the 1% level. In economic terms, this means that the nested structure, where Germany is a special destination country, becomes specially relevant after 1938.

In column 2, we use the restricted version of the FP sample (both for the decision makers and for the distant peers). We re-estimate the triadic gravity regressions (representing the lower model choices) on the FP sample and recompute the inclusive utility. Those are large changes, both in terms of the size of sample (now 60% larger in terms of decision makers—all stayers) and content of the RHS variables. The coefficients of interest all increase, but remain of comparable magnitude. In column 4, we use the expanded version of the FP sample. The coefficients of interest hardly change in terms of size and statistical significance. Overall, the relative stability of results, in spite of such a large change in samples size, is another indication that a selection bias is unlikely to contaminate our empirical analysis.

Table 7: Robustness – Outmigration Decisions on Extended Samples: 1933–1938

	(1)	(2)	(3)
Dependent variable:	Migration Decision		
Sample:	F sample	Restricted FP sample	Extended FP sample
Detainment of network members	1.207***	1.829***	1.883***
	(0.250)	(0.334)	(0.338)
Migration of network members	1.135***	1.238***	1.265***
	(0.123)	(0.173)	(0.188)
Inclusive Utility	0.848***	1.141***	1.083***
	(0.068)	(0.063)	(0.063)
Year FE	✓	✓	✓
Observations	716,862	1,178,629	1,254,111
Pseudo R-squared	0.10	0.08	0.08
Mean of dependent var.	0.041	0.027	0.025
SD of dependent var.	0.199	0.161	0.157
Number of decision makers	130,454	207,950	220,672

NOTE: Logit regressions with the dependent variable being an indicator for migration in year t. The unit of observation is an individual in year t. The sample consists of individuals who were at least 16 years old in 1933. All regressions control for age, age squared, gender, personal detainment, a dummy for whether the individual was born outside Germany, and the partial Inclusive Utility of individual i. Standard errors clustered at the city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7 Counterfactual History

7.1 Exact Hat Algebra

In our counterfactual analysis, we use techniques initiated by the trade literature (Dekle *et al.* (2007) being the seminal contribution, see Costinot and Rodriguez-Clare (2014) for a general presentation), often referred to as Exact Hat Algebra (EHA), which are particularly appropriate in our context. The fundamental idea is to use the CES structure of the model to express proportional changes (denoted by the hat notation) of migration flows resulting from policy changes as a function of the observed levels of the same flows and a very parsimonious set of structural parameters. There are several advantages of this approach. First it computes the counterfactual change directly, rather than having to solve for the model "in levels" twice, reducing computation time by roughly half. Second and most important, because it does not require to solve the model, it is extremely economical in terms of data requirement thanks to the CES structure. For example, with CES demand structure, routinely used in quantitative trade or macro models, the market share of a variety is a sufficient statistic for all relevant variables that will drive its change (price, physical quality, appeal, etc.). In our case, since we work with a CES equation for migration shares (as in the recent literature surveyed by Redding and Rossi-Hansberg, 2017), those shares capture all characteristics that are not affected by the policy change, *including those that are otherwise unobservable*. Third, the EHA approach starts from actual data patterns (migration shares in our case) and replicates the true state of affairs under the status quo. We view this as a crucial advantage in quantitative economic history work such as ours.

The alternative approach solves the model using the observable variables and estimated parameters, once under the true levels of the policy, and once under its scenarized level. As a result, it yields two levels of the migration shares to be compared (hence the denomination of Difference in Expected Value – DEV – given by Head and Mayer, 2019). However, even in cases where the fit of model to data is good, some of the predicted shares under the status quo can be quite far from observed ones. For instance, some countries might attract no migrants despite the model predicting it should. If the policy analyzed is a tightening of the migration restrictions, the counterfactual could predict a fall in a flow that in reality never existed. Worse, since in our model a change in the flows for a given year spills over to later years through network migration, errors will contaminate later years. There is however one major disadvantage of EHA compared to DEV: Because of the CES structure, the model cannot predict a zero unless migration costs are infinite. Therefore, a zero flow in the data will stay a zero flow in the counterfactual scenarios (and conversely), whereas in reality, a policy might induce some positive flows to emerge, or cause some flows to die out. In our view, the benefits of EHA highlighted above dominate this flaw.

All computational details of the exact hat algebra are relegated to Online Appendix Section H.1. The structural elasticities required to run the counterfactual simulations are recovered from the estimation results of the lower-model (column 5 of Table 4), the non-instrumented version of the upper-model (column 1 of Table 6, panel A), and the table showing the determinants of country attractiveness (column 1 of Table 5). These structural parameters take the following values:

$$\lambda_1 = 0.409;\ \lambda_2 = 0.306;\ \alpha_1 = 1.662;\ \alpha_2 = 1.619;\ \beta = 0.155;\ \gamma = 0.197.$$

Counterfactuals are run on a population at risk composed of individuals from our baseline estimation sample (the F sample). Because the hat algebra is fed with information from the lower-level model, we drop all migrants who are not part of its estimation sample. We end up with a population at risk of 167,108 adults: the 62,969 migrants (corresponding to row 5 of Online Appendix Table A3) and the 104,139 individuals who did not migrate and were deported (with a known deportation date, corresponding to row 5″ of Online Appendix Table A1).

7.2 Policy Simulation

We now turn to simulate several counterfactual policies. The counterfactual scenarios are implemented in 1936, a few months after the enactment of the Nuremberg Laws of September 1935, which institutionalized Jewish persecution and made it visible to the international community. We consider the following scenarios: (1) unilateral opening of U.S. borders; (2) non-closing of Palestine in 1936; (3) removing work restrictions; (4) subsidizing transportation; and (5) early perception of the threat. Counterfactuals 1 to 3 are modeled as changes in the attractiveness of the relevant destination countries in the lower-level equation (10).[47] Counterfactual 4 corresponds to a change in bilateral migration cost in the same equation. Counterfactual 5 is engineered as a change in the perception of detainment in the upper-level equation (12). For all these interventions, our efficiency metric is the additional number of migrants accumulated by 1941 and the number of lives saved as a result. Our model generates two margins of adjustment. The first is the intertemporal margin: People migrate out of Germany *earlier* than what they would have done absent the policy change.[48] The second is the extensive margin whereby the total number of migrants increases. Only the second margin contributes to the goal of those policies, which is to save lives and to decrease the number of Jews still in Germany in 1941.

In Table 8, we present the results of our simulations, each row representing a scenario. Columns 1 and 2 are based on the non-instrumented values of parameters, while columns 3 and 4 use the instrumented ones. For each scenario, we report in columns 1 and 3 the counterfactual changes in cumulative migration out of Germany over 1933–1941. We express those changes in percentage point deviations from total migration observed in the data (the status quo): 62,969 migrants. Columns 2 and 4 provide the results of our simulations expressed in terms of people who got deported and ended up murdered. This quantification proceeds in two steps. First, we compute the counterfactual number of deportees, accounting not only for those who stayed in Germany until 1941, but also for those who migrated to countries that were occupied by Germany after 1938 and got deported later on. Based on the Resident List, we compute in Online Appendix Table G.7 the deportation risk for all destination countries. Unsurprisingly, migrants who went to countries neighboring Germany faced a high deportation risk. For example, 54% and 36% of migrants in Netherlands and Belgium were later deported, respectively. Second, we compute mortality among the deportees by using a survival rate of 11%. This figure comes from the Resident List and is within the range of existing estimates reported by other sources (see Online Appendix Table G3). In Table 8, mortality under the status quo amounts to 98,505 out of the sample of 167,108 adults on which the counterfactuals are run.[49]

We now turn to discuss each scenario in detail. We comment on the quantifications based on the

[47]Our counterfactuals use the hat notation to denote the proportional change in the relevant variables, for instance $\widehat{a}_d \equiv a'_d/a_d$ in Online Appendix equation (H.6) is the ratio of counterfactual estimated attractiveness of country d over that in the status quo (recall that the notation for log attractiveness of d used in Section 6.1 is A_d , hence $a_d \equiv \exp(A_d)$).

[48]Because the initial population at risk under consideration is fixed (i.e., Jewish adults living in Germany in 1933), an intervention in 1936, if successful, reduces the remaining population at risk in 1937. It is therefore possible for the counterfactual outflows in 1937 and following years to be lower than the observed ones. This is true even if the dynamics of our network spillovers raise the *probability* of outmigrating in 1937.

[49]More precisely, this number is computed as

$$\text{mortality} = 0.89 \times \left(\text{popatrisk}_{1933} - \sum_d (1 - \eta_d)\text{stock}_{d,1941} \right), \tag{17}$$

where $\text{popatrisk}_{1933} = 167,108$, stock_d is the cumulative number of individuals in the Resident List who have outmigrated between 1933 and 1941 in destination d and η_d is the share of migrants to destination d who were ultimately deported.

non-instrumented values of the parameters while describing the scenarios, leaving the discussion of those based on the instrumented values to the sensitivity analysis paragraph.

Table 8: Outmigration and Mortality in the Counterfactual Scenarios: 1933–1941

	(1)	(2)	(3)	(4)
Set of parameters:	Non-instrumented		Instrumented	
Simulated outcome:	Migration	Mortality in deportation	Migration	Mortality in deportation
Status Quo (obs. nb of Jewish adults)	62969	98505	62969	98505
<u>Scenarios from 1936 onwards</u> (% change)				
(1) 5,000 additional migrants to U.S. in 1936	8.5	-5	13.5	-7.7
(2) Non-Closing of Palestine	7.3	-4.4	10.5	-6.1
(3) Removing work restrictions	12.2	-5.9	20.3	-10.1
(4.a) Travel subsidy: subsidy to U.S. only	4.9	-2.9	6.7	-3.8
(4.b) Travel subsidy: subsidy to all port festinations	9.7	-5.7	13.8	-7.9
(5) Post-Nuremberg perception of threat	2.7	-1.3	4.2	-2.2

NOTE: This table displays the changes in cumulative Jewish migration over 1933-1941 and in mortality across the different counterfactual scenarios. All changes are expressed in percentage point deviations from status quo. All interventions are implemented from 1936 onwards (post Nuremberg laws). The population at risk under consideration is composed of 167,108 adults living in Germany in 1933 (the F sample, corresponding to the combination of row 5 of Online Appendix Table A3 and row 5" of Online Appendix Table A1). The mortality figures implied by different scenarios are computed using first the observed deportation rate in the countries where the first migration occurred, combined with the survival rate of 11% that is revealed by the Resident List.

Scenario 1: Unilateral opening of U.S. borders. We first consider a unilateral opening of borders, focusing on the United States. While the United States had set a quota in 1924 allowing 27 thousand migrants from Germany per year, in the early 1930s the quota was not filled. However, after *Kristallnacht*, migration to the United States from Germany surged and more than 300 thousand applicants were waiting for a visa (Breitman, 2013). Political attempts to open borders, such as the bipartisan Wagner–Rogers Bill in the U.S. Congress, which proposed to allow 10 thousand children per year to come to the U.S. in 1939 and 1940, were rejected. The policy intervention we consider follows this proposal and increases the inflow of Jewish immigrants to the United States by 5,000 people in 1936. Contrary to the other counterfactuals presented below, we do not aim here at detailing how exactly this policy scenario could be implemented. To evaluate the effects of such a scenario, we do not need to take a stance on the measures taken to attract more migrants. We implement this policy experiment by exogenously increasing U.S. attractiveness, a_{USA}, in our model. Inverting the migration equations (10) and (12), we simply reveal the change in attractiveness of the United States in 1936 needed to generate the additional inflow. The implied increase in U.S. attractiveness is very large: $\widehat{a}_{USA,1936} = 2.41$. In the subsequent years, attractiveness remains unchanged ($\widehat{a}_{USA,t>1936} = 1$). As shown in Table 8 (row 1), 8.5% more migrants (5,352 individuals) would have left Germany by 1941 in this scenario compared to what is observed in the status quo; and mortality would have decreased by 5% (4,925 saved lives).

Scenario 2: Non-closing of Palestine after 1936. The next scenario we study is the non-closing of Palestine in 1936. A record number of Jewish immigrants arrived in Palestine in 1935. The increased inflow of Jews led to a revolt by the Palestinian Arab community against the British administration to stop immigration. As a result, the British Mandate reduced significantly the allocation of immigration certificates from 1936 onwards (Nicosia, 2000; Hacohen, 2001). Our empirical results are in line with these historical narratives: We observe a threefold reduction in the estimated attractiveness of Palestine between 1935 and 1940 (see Online Appendix Table G14). What if Palestine had stayed open and a similarly attractive place to migrate to? To study this question, we assign the attractiveness of Palestine (PAL hereafter) in 1935 to the subsequent years $t > 1935$ such that $\widehat{a}_{PAL,t} = a_{PAL,1935}/a_{PAL,t}$. Under this scenario (row 2), 7.3% additional adults would have migrated out of Germany (4,596 migrants); and mortality would have decreased by 4.4% (4,334 saved lives).

Scenario 3: Removing work restrictions. Next, we study how policies that limit the economic opportunities of refugees affect emigration. In the post Great Depression years when economic conditions were harsh globally, many countries restricted the access of migrants to their local labor markets. In the United Kingdom, for example, employment restrictions were severe: While migrants were generally allowed to work, the employer had to prove that no British person could do the job (Löwenthal and Oppenheimer, 1938; London, 2003). The British policies implied that most refugees could not work, or they worked in low-skill jobs, in particular, as domestic servants.

For the counterfactual analysis, we collected information on labor market policies in place during the 1930s. For each destination and year, we code the access of immigrants to the labor market on a scale from 1 to 4, where 1 indicates "no restrictions," 2 "work allowed with permit," 3 "work allowed, but permit difficult to obtain," and 4 indicates "no access to the labor market." In Online Appendix Table H1, using an auxiliary regression, we estimate the elasticity of destinations' attractiveness to their labor market restrictions. The estimated elasticity is negative, which validates that tighter restrictions are indeed associated with lower attractiveness and flows of refugees. We then ask, what if countries had removed employment restrictions after the Nuremberg Laws (i.e., from 1936 onwards), instead of making access to their labor market difficult? Using the estimated elasticity, we compute the counterfactual changes in attractiveness induced by lifting labor market restrictions in destination countries (i.e., moving all labor market policies to "no restrictions"). Online Appendix Figure H1 reports for each country the observed average work restrictions over the period 1936–1941 and the increase in attractiveness caused by the lifting. The effects can be large: For example, attractiveness of UK and France rise by 59% and 43%, respectively. Finally, we simulate the model to quantify the changes in total migration and mortality. As shown in row 3 of Table 8, compared to the status quo, 12.2% more migrants (7,682 individuals) would have left Germany; and mortality would have been reduced by 5.9% (5,811 saved lives). Under this scenario, the number of additional migrants exceeds subsantially the number of saved lives. The reason for this discrepancy is that the counterfactual removal of work restrictions would increase migration flows to many destination countries, including countries neighboring Germany that were later occupied, where the deportation risk and resulting mortality rates were high. By contrast, in scenarios 1 and 2, the discrepancy is less pronounced as individuals relocate preferentially to the U.S. and Palestine—relatively safe heavens for Jews during the war.

Scenario 4: Travel subsidies. One instrument to help Jewish migrants that was discussed at the Evian conference was the provision of financial assistance. The British Government for example considered to open their colonies to migrants and to establish financial help through loans or subsidies to shipping lines that transported migrants overseas (Hoffmann, 2011; Packer, 2017). Indeed, oversea travel costs were high: Transport costs ranged between 150 and 1320 Reichsmark, between 10% to 100% of the average yearly income per capita in Nazi Germany.[50] To ease the financial burden of emigration, organizations such as the American Jewish Joint Distribution Committee provided financial assistance for visa and travel costs (Kaplan, 2020).

In this counterfactual scenario, we simulate how migration would have reacted to travel subsidies. We collected actual ticket prices in Reichsmark from Löwenthal and Oppenheimer (1938) for 42 boat trips from European ports to overseas (non-European) destination ports. In Online Appendix Table H2 we regress the ticket prices on port-to-port distances to estimate the elasticity of travel cost with respect to distance. Then, we reduce the cost of travel to overseas destinations by 50% (on average about 300 Reichsmark). Using the estimated elasticity, we translate this reduction in costs into the equivalent reductions in distance from Germany. Finally, we run the exact hat algebra using parameter values $\tilde{\rho}_1 = 0.141$ and $\tilde{\rho}_2 = 0.038$ that are recovered from the time-invariant distance coefficients in column 4 of Table 4. We consider two subsidy scenarios: a unilateral subsidy to the U.S. and a subsidy that reduces the distance to all oversea ports at the same time — a measure that could have been the outcome of a coordinated policy effort. Row 4.a of Table 8 shows that a 50% reduction of ticket prices to the U.S. in 1936 results in an increase in total migration by 4.9% (3,085 migrants) and a decrease in mortality by 2.9% (2,857 saved lives). According to row 4.b of Table, a subsidy of 50% of the ticket prices to all overseas destinations would have increased migration by 9.7% (6,108 migrants) and reduced mortality by 5.7% (5,615 saved lives).

Scenario 5: Early perception of the threat. The last scenario we simulate focuses on the perception of the threat within Germany. If people had more accurately estimated of the threat early on, would outmigration have increased dramatically? Conceptually there are two main ways to operationalize an improvement in the perception of the threat. On the one hand, an accurate perception of the threat could mean that all individuals have perfect knowledge about the state of the danger in each year. On the other hand, a more accurate estimate of the danger could indicate that individuals have perfect foresight of the events that will happen in the future.

We follow the second approach and simulate outmigration in a world where Jews would have already known in 1936 (after the Nuremberg Laws) the rates of persecution that would take place during and after *Kristallnacht*. That is, we assign the 95th percentile value of cumulative detainment (38%) in 1939, i.e., one year after *Kristallnacht*, to all adults in 1936. As displayed in row 5, knowledge of post-*Kristallnacht* detainment would have increased outmigration: Compared to

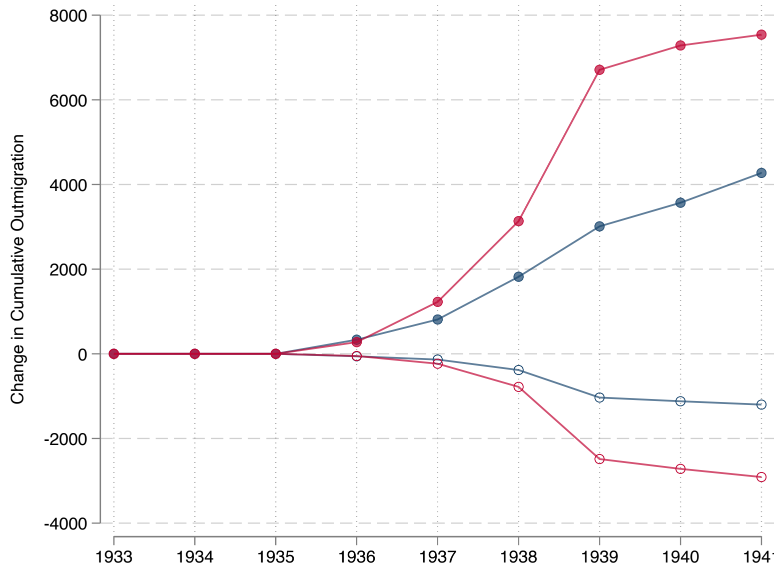
[50] Ticket prices are from Löwenthal and Oppenheimer (1938). Estimates of GDP per capita are from Spoerer (2005).

the status quo, 1,700 additional migrants would have left Germany (2.7% increase) and 1,280 additional lives would have been saved (1.3%). However, compared to policies lowering migration frictions or increasing destinations' attractiveness, the impact of an early assessment of the threat on outmigration would have been systematically smaller.

Migration creation, migration diversion. In Figure 5, we investigate substitution patterns in migration choice. To facilitate interpretation, we consider two scenarios where the migration policy is implemented *unilaterally*: non-closing of Palestine and travel subsidies to the U.S.. In each scenario, the curve with filled circles depicts migration from Germany to the country implementing the policy (i.e., Palestine (red) and U.S. (blue), respectively) and the curve with hollow circles represents migration from Germany to the rest of the world. After the date of intervention in 1936, we observe migration creation in implementing countries (filled-circle curves are upward sloping) and migration diversion in the rest of the world (hollow-circle curves are downward sloping). In both cases, migration creation offsets migration diversion: Total migration out of Germany increases. Yet, it is clear that the policy impact on total outmigration would have been larger in absence of diversion. Such substitution patterns have important policy implications in the context of humanitarian crises and refugee relief. When the degree of substitution is low, unilateral opening of borders increases the total refugee flows out of the country at risk. However, when it is high, unilateral opening of borders mostly redirects flows of refugees who would have fled *anyway*. Hence, a quantitative assessment of the substitution patterns is key for identifying the most efficient interventions aimed at increasing refugee outflows and saving lives.

Spillover decomposition. In Table 9, we decompose the overall migration spillovers into their two components, the exodus and diaspora channels. The contribution of each channel is quantified by re-running all counterfactual simulations after shutting down the alternative channel: zeroing α_1 and α_2 (diaspora channel) to isolate the exodus channel and zeroing β (exodus channel) to isolate the diaspora channel in our migration equations (10) and (12). Column 1 recalls the benchmark simulation figures. Columns 2 and 3 report the isolated effects of exodus and diaspora channels, respectively. Column 4 removes all spillover effects and captures the direct impact of each intervention on migration, absent the spillovers. Comparing the effects in columns 2 and 3 with those in column 4 reveals that both the diaspora and exodus channels magnify the direct impact of interventions: Although their respective contributions do vary across scenarios, both dynamic spillovers enhance the efficiency of refugee policies. For example, in the absence of spillovers, the unilateral opening of U.S. borders (row 1, column 4) would have increased cumulative migration in 1941 by only 2,518 migrants (4%). Here, the intertemporal margin of adjustment is large: Out of the additional 5,000 migrants who flee to the U.S. in 1936, 2,482 (= 5,000 - 2,518) would have fled anyway before 1941. When the exodus and diaspora channels are both active (row 1, column 1), total migration would have increased by 5,352 individuals (8.5%). Therefore, the two migration spillovers contribute to the extensive margin and to the goal of saving lives and decreasing the number of people living in Germany at risk of persecution.

Figure 5: Substitution Patterns



NOTE: This figure displays cumulative migration of Jews to the country implementing the policy unilaterally (filled circles) and to the Rest of the World (hollow circles) under scenarios 2 and 4a (non-closing of Palestine in red, and subsidized travel to the U.S. in blue). All interventions are implemented from 1936 onwards (after Nuremberg Laws). The population at risk under consideration is composed of 167,108 adults living in Germany in 1933 (the F sample, corresponding to the combination of row 5 of Online Appendix Table A3 and row 5" of Online Appendix Table A1).

Sensitivity analysis. As a sensitivity test, we run our counterfactual simulations with alternative parameter values, $\lambda_1 = 0.377$, $\beta = 0.703$, $\gamma = 0.184$, that are recovered from the instrumented version of the upper model (Table 6, Panel B, column 1). The last two columns of Table 8 summarize the results. The main difference with the previous set of parameters is that the estimated value of β, the parameter driving the exodus effect, is larger in the instrumented version. The parameter λ_1 is also slightly lower in the instrumented version. Since the elasticity of outmigration with respect to inclusive utility is λ_2/λ_1 (and since instrumentation does not affect the estimated value of λ_2), this will also strengthen the overall effects of network spillovers. Not surprisingly, the results implied by the instrumented model, therefore, exhibit larger consequences in all five policy scenarios.

In a second robustness check, we re-run all the counterfactual simulations on a larger population at risk, including the 83,959 adults we identified in the 1939 Census (row 6 of Table 1). Because they have incomplete spells ending in 1938, we conduct the simulations on the restricted 1933–1938 period. This restriction prevents us from computing counterfactual deportation and mortality as Jews were still allowed to migrate until late 1941. Results are displayed in Table 10. For benchmarking purposes, in column 1, we replicate the simulations reported in column 1 of Table 8 for the restricted period 1933–1938. The population at risk is unchanged—167,108 adults— but parameters are recovered from the estimation restricted to the same period (column 1 of Table 7). The results show that, in all scenarios, the impact of the policy in 1938 is larger than its impact in 1941 (specially for the last scenario). There are two simple reasons. First, all parameters are larger when estimated over the 1933-1938 period. Second, counterfactuals are expressed in relative terms and

Table 9: Outmigration in the Counterfactual Scenarios – Spillover Decomposition

	(1)	(2)	(3)	(4)
Simulated outcome:	Migration 1933–1941			
Spillovers:	All	Exodus Only	Diaspora Only	None
Scenarios from 1936 onwards (% change)				
(1) 5,000 additional migrants to U.S. in 1936	8.5	4.6	7.6	4
(2) Non-Closing of Palestine	7.3	6.2	6.9	5.8
(3) Removing work restrictions	12.2	10.8	11	9.7
(4.a) Travel subsidy: subsidy to U.S. only	4.9	3.5	4.6	3.3
(4.b) Travel subsidy: subsidy to all port destinations	9.7	8.3	9.1	7.8
(5) Post-Nuremberg perception of threat	2.7	2.2	2.4	2

NOTE: This table displays the changes in cumulative Jewish migration over 1933-1941 across the different counterfactual scenarios. All changes are expressed in percentage point deviations from status quo. All interventions are implemented from 1936 onwards (post Nuremberg Laws). The population at risk under consideration is composed of 167,108 adults living in Germany in 1933 (the F sample, corresponding to the combination of row 5 of Online Appendix Table A3 and row 5″ of Online Appendix Table A1). Column 1 collects the set of benchmark results, column 2 shuts down the diaspora channel (by setting structural parameters α_1 and α_2 to 0), column 3 shuts down the exodus channel (by setting structural parameter β to 0) and column 4 shuts down both channels.

the status quo cumulative migration is lower in 1938 (35,428 migrants) than it is in 1941 (62,969 migrants).

Then, we run our simulations on the extended population at risk of 251,067 adults using the parameters recovered from column 2 in Table 7. We do not include individuals who died in Germany or were expelled to Poland (row 7 of Table 1) to avoid attrition when computing the exact hat algebra. Results are displayed in column 2 of Table 10. Counterfactuals based on the extended sample yield larger quantitative effects than the ones obtained on the F sample. This is the natural consequence of the set of structural parameters used, which are all larger when estimated on the FP sample.

8 Conclusion

In this paper, we study the importance of social networks and immigration restrictions in destination countries for Jewish migration out of Germany during the period 1933 to 1941. Using individual-level data on Jewish residents of Germany, we find that network externalities played a first-order role in outmigration decisions. In particular, we document evidence for two novel channels of how networks can affect emigration in situations of violence: First, our results show that networks aggregate information about the extent of persecution, which affects outmigration incentives positively (threat effect). Second, we estimate a structural model of emigration and quantify the impact on outmigration of observing peers fleeing from Germany (exodus effect). This exodus effect is of significant magnitude and operates besides the standard diaspora effect of migration

Table 10: Outmigration in the Counterfactual Scenarios on Extended Sample: 1933–1938

	(1)	(2)
Simulated outcome:	Migration 1933–1938	
Sample:	F sample	FP sample
Status Quo (obs. nb of Jewish adults)	35428	35428
Scenarios from 1936 onwards (% change)		
(1) 5,000 additional migrants to U.S. in 1936	16.9	22.1
(2) Non-Closing of Palestine	8.1	14.4
(3) Removing work restrictions	24.5	39.7
(4.a) Travel subsidy: subsidy to U.S. only	4.9	8.1
(4.b) Travel subsidy: subsidy to all port pestinations	10.2	16.5
(5) Post-Nuremberg perception of threat	15.4	24.5

NOTE: This table displays the changes in cumulative Jewish migration over 1933-1938 across the different counterfactual scenarios. All changes are expressed in percentage point deviations from status quo. All interventions are implemented from 1936 onwards (post Nuremberg Laws). In column 1, the population at risk under consideration is composed of 167,108 adults living in Germany in 1933 (the F sample, corresponding to the combination of row 5 of Online Appendix Table A3 and row 5″ of Online Appendix Table A1). In column 2, the population at risk under consideration is composed of 251067 adults living in Germany in 1933 (the FP sample, corresponding to the combination of row 5 of Online Appendix Table A3, row 5″ of Online Appendix Table A1, and row 6 of Table 1).

networks. Our results suggest that in situations of conflict when emigration becomes massive the exodus effect is crucial.

Our paper also develops quantitative tools for evaluating how asylum policies affect the volume and composition of refugee flows and for assessing counterfactual scenarios. More precisely, we simulate six policy experiments and report what would have been their consequences in terms of saved lives. While we derive our results from a period of persecution and displacement that happened 80 years ago, our findings can also speak to modern refugee crises. For academics and policy makers alike, it is important to understand how migration decisions are made in situations of conflict. Our paper quantifies the importance of coordination failures in the allocation of refugees across destinations. The interaction between those coordination failures and the presence of social spillovers can have large scale impacts that should be taken into account when designing policies.

References

ABOWD, J. M., KRAMARZ, F. and MARGOLIS, D. N. (1999). High Wage Workers and High Wage Firms. *Econometrica*, **67** (2), 251–333.

ACEMOGLU, D., HASSAN, T. A. and ROBINSON, J. A. (2011). Social Structure and Development: A Legacy of the Holocaust in Russia. *Quarterly Journal of Economics*, **126** (2), 895–946.

ADENA, M., ENIKOLOPOV, R., PETROVA, M., SANTAROSA, V. and ZHURAVSKAYA, E. (2015). Radio and the Rise of the Nazis in Prewar Germany. *The Quarterly Journal of Economics*, **130** (4), 1885–1939.

AKBULUT-YUKSEL, M. and YUKSEL, M. (2015). The Long-Term Direct and External Effects of Jewish Expulsions in Nazi Germany. *American Economic Journal: Economic Policy*, **7** (3), 58–85.

ALICKE, K.-D. (2014). Encyclopedia of Jewish Communities in the German Language. https://www.jüdische-gemeinden.de/index.php/home.

ANDERSON, R. W., JOHNSON, N. D. and KOYAMA, M. (2017). Jewish Persecutions and Weather Shocks: 1100-1800. *The Economic Journal*, **128** (602), 924–958.

ANDERSON, S., DE PALMA, A. and THISSE, J. (1992). *Discrete choice theory of product differentiation*. MIT Press.

ANDREWS, M. J., GILL, L., SCHANK, T. and UPWARD, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **171** (3), 673–697.

ANGELUCCI, C., MERAGLIA, S. and VOIGTLÄNDER, N. (2021). How Merchant Towns Shaped Parliaments: From the Norman Conquest of England to the Great Reform Act. *Unpublished manuscript*.

BARTROP, P. R. and DICKERMAN, M. (2017). *The Holocaust: An Encyclopedia and Document Collection*. ABC-CLIO.

BECKER, S. and PASCALI, L. (2019). Religion, Division of Labor and Conflict: Anti-Semitism in German Regions over 600 Years. *American Economic Review*, **109** (5), 1764–1804.

BECKER, S. O. and FERRARA, A. (2019). Consequences of forced migration: A survey of recent findings. *Labour Economics*, **59**, 1–16.

—, LINDENTHAL, V., MUKAND, S. and WALDINGER, F. (2021). Persecution and Escape: Professional Networks and High-Skilled Emigration from Nazi Germany. *CAGE Working Paper 542*.

BEINE, M., BERTOLI, S. and FERNÁNDEZ-HUERTAS MORAGA, J. (2016). A practitioners' guide to gravity models of international migration. *The World Economy*, **39** (4), 496–512.

—, DOCQUIER, F. and ÖZDEN, Ç. (2011). Diasporas. *Journal of Development Economics*, **95** (1), 30–41.

BERTOLI, S. and MORAGA, J. F.-H. (2013). Multilateral resistance to migration. *Journal of Development Economics*, **102**, 79–100.

— and — (2015). The size of the cliff at the border. *Regional Science and Urban Economics*, **51**, 1–6.

BLUM, M. and REI, C. (2018). Escaping Europe: health and human capital of Holocaust refugees. *European Review of Economic History*, **22** (1), 1–27.

BOHRA-MISHRA, P. and MASSEY, D. S. (2011). Individual Decisions to Migrate During Civil Conflict. *Demography*, **48** (2), 401–424.

BRAMOULLÉ, Y., DJEBBARI, H. and FORTIN, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, **150** (1), 41–55.

BREITMAN, R. (2013). *FDR and the Jews*. Harvard University Press.

BRYAN, G. and MORTEN, M. (2019). The aggregate productivity effects of internal migration: Evidence from Indonesia. *Journal of Political Economy*, **127** (5), 2229–2268.

CALIENDO, L., PARRO, F., OPROMOLLA, L. D. and SFORZA, A. (2020). *Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement*. Mimeo.

CAMERON, A. C. and TRIVEDI, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge University Press.

CAPRETTINI, B. and VOTH, H.-J. (2021). New Deal, New Patriots: How 1930s Welfare Spending Boosted the War Effort during WW II. *Unpublished manuscript*.

CARD, D., HEINING, J. and KLINE, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, **128** (3), 967–1015.

CHIN, A. and CORTES, K. E. (2015). Chapter 12 - The Refugee/Asylum Seeker. In B. R. Chiswick and P. W. Miller (eds.), *Handbook of the Economics of International Migration*, *Handbook of the Economics of International Migration*, vol. 1, North-Holland, pp. 585 –658.

COMOLA, M. and MENDOLA, M. (2015). Formation of migrant networks. *The Scandinavian Journal of Economics*, **117** (2), 592–618.

COSTINOT, A. and RODRIGUEZ-CLARE, A. (2014). Trade Theory with Numbers: Quantifying the Consequences of Globalization. In E. Helpman, G. Gopinath and K. Rogoff (eds.), *Handbook of International Economics*, vol. 4, *4*, Elsevier, pp. 197–261.

DEKLE, R., EATON, J. and KORTUM, S. (2007). Unbalanced Trade. *American Economic Review*, **97** (2), 351–355.

DOEW (1988). *„Anschluß" 1938. Eine Dokumentation*, Wien: Dokumentationsarchiv des österreichischen Widerstandes, chap. 8. Erste Maßnahmen zur Institutionalisierung des Antisemitismus.

EATON, J. and KORTUM, S. (2002). Technology, Geography, and Trade. *Econometrica*, **70** (5), 1741–1779.

ENGEL, S. and IBÁÑEZ, A. M. (2007). Displacement due to violence in Colombia: A household-level analysis. *Economic development and cultural change*, **55** (2), 335–365.

FAJGELBAUM, P. D., MORALES, E., SUÁREZ SERRATO, J. C. and ZIDAR, O. (2018). State Taxes and Spatial Misallocation. *The Review of Economic Studies*, **86** (1), 333–376.

FALLY, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, **97** (1), 76–85.

FISCHER, C. S. (1982). *To dwell among friends: Personal networks in town and city*. University of chicago Press.

FRIEDMAN, S. S. (2017). *No haven for the oppressed: United States policy toward Jewish refugees, 1938-1945*. Wayne State University Press.

GIULIETTI, C., WAHBA, J. and ZENOU, Y. (2018). Strong versus weak ties in migration. *European Economic Review*, **104**, 111–137.

GROSFELD, I., RODNYANSKY, A. and ZHURAVSKAYA, E. (2013). Persistent Antimarket Culture: A Legacy of the Pale of Settlement after the Holocaust. *American Economic Journal: Economic Policy*, **5** (3), 189–226.

—, SAKALLI, S. O. and ZHURAVSKAYA, E. (2020). Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire. *The Review of Economic Studies*, **87** (1), 289–342.

GRUNER, W. (2011). *Deutsches Reich 1933–1937*, vol. 1. Oldenbourg Verlag.

— and PEARCE, C. (2019). *German Reich 1933–1937*. De Gruyter.

GUIMARAES, P., FIGUEIRDO, O. and WOODWARD, D. (2003). A tractable approach to the firm location decision problem. *The Review of Economics and Statistics*, **85** (1), 201–204.

HACOHEN, D. (2001). British immigration policy to Palestine in the 1930s: Implications for youth aliyah. *Middle Eastern Studies*, **37** (4), 206–218.

HAINMUELLER, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, **20** (1), 25–46.

HEAD, K. and MAYER, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In E. Helpman, G. Gopinath and K. Rogoff (eds.), *Handbook of International Economics*, vol. 4, Elsevier, pp. 131–195.

— and — (2019). Brands in Motion: How frictions shape multinational production. *American Economic Review*, **109** (9), 3073–3124.

HEUSLER, A. and SINN, A. (2015). *Die Erfahrung des Exils: Vertreibung, Emigration und Neuanfang. Ein Münchner Lesebuch*. Studien zur Jüdischen Geschichte und Kultur in Bayern, De Gruyter.

HOFFMANN, P. (2011). *Carl Goerdeler and the Jewish Question, 1933–1942*. Cambridge University Press.

HUBER, K., LINDENTHAL, V. and WALDINGER, F. (2021). Discrimination, Managers, and Firm Performance: Evidence from 'Aryanizations' in Nazi Germany. *Journal of Political Economy*, **129** (9), 2455–2503.

IBÁÑEZ, A. M. and VÉLEZ, C. E. (2008). Civil conflict and forced migration: The micro determinants and welfare losses of displacement in Colombia. *World Development*, **36** (4), 659–676.

JEDWAB, R., JOHNSON, N. and KOYAMA, M. (2019). Negative Shocks and Mass Persecutions: Evidence from the Black Death. *Journal of Economic Growth*, **24** (4), 345–395.

JOCHMANS, K. and WEIDNER, M. (2019). Fixed-Effect Regressions on Network Data. *Econometrica*, **87** (5), 1543–1560.

KAPLAN, M. (2020). *Hitler's Jewish Refugees: Hope and Anxiety in Portugal*. Yale University Press.

KAPLAN, M. A. (1999). *Between dignity and despair: Jewish life in Nazi Germany*. Oxford University Press.

— (2005). As Germans and as Jews in Imperial Germany. In M. A. Kaplan (ed.), *Jewish Daily Life in Germany, 1618-1945*, Oxford University Press.

LARSEN, J., AXHAUSEN, K. W. and URRY, J. (2006). Geographies of social networks: meetings, travel and communications. *Mobilities*, **1** (2), 261–283.

LONDON, L. (2003). *Whitehall and the Jews, 1933-1948: British immigration policy, Jewish refugees and the Holocaust*. Cambridge University Press.

Löwenthal, E. G. and Oppenheimer, H. (1938). *Philo-Atlas: Handbuch für die Jüdische Auswanderung*, vol. 3. Philo, Jüdischer Buchverlag.

Matthäus, J. and Roseman, M. (2010). *Jewish Responses to Persecution, 1933–1938 (Documenting Life and Destruction: Holocaust Sources in Context, vol. 1)*. Lanham, MD: AltaMira Press.

Maurer, T. (2005). From Everyday Life to a State of Emergency: Jews in Weimar and Nazi Germany. In M. A. Kaplan (ed.), *Jewish Daily Life in Germany, 1618-1945*, Oxford University Press.

McKenzie, D. and Rapoport, H. (2010). Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks. *The Review of Economics and Statistics*, **92** (4), 811–821.

Mendelsohn, E. (1999). People of the City: Jews and the Urban Challenge. In *Studies in Contemporary Jewry, Volume XV*, Oxford University Press.

Monte, F., Redding, S. J. and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, **108** (12), 3855–90.

Moser, P., Voena, A. and Waldinger, F. (2014). German Jewish Emigres and US Invention. *The American Economic Review*, **104** (10), 3222–3255.

Mulligan, C. B. and Rubinstein, Y. (2008). Selection, Investment, and Women's Relative Wages Over Time. *The Quarterly Journal of Economics*, **123** (3), 1061–1110.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, **118** (2), 549–599.

Nicosia, F. R. (2000). *The Third Reich and the Palestine Question*. Routledge.

— and Scrase, D. (2013). *Jewish life in Nazi Germany: dilemmas and responses*. Berghahn Books.

Niederland, D. (1993). Back into the Lion's Jaws: A note on Jewish Return Migration to Nazi Germany (1933-1938). In *Studies in Contemporary Jewry, Volume IX*, Oxford University Press.

Ortega, F. and Peri, G. (2013). The effect of income and immigration policies on international migration. *Migration Studies*, **1** (1), 47–74.

Packer, D. (2017). *Britain and rescue: Government policy and the Jewish refugees 1942-1943*. Ph.D. thesis, Northumbria University.

Paravisini, D., Rappoport, V., Schnabl, P. and Wolfenzon, D. (2015). Dissecting the Effect of Credit Supply on Trade: Evidence from Matched Credit-Export Data. *The Review of Economic Studies*, **82** (1), 333–359.

Redding, S. J. (2020). *Trade and Geography*. Working Paper 27821, National Bureau of Economic Research.

— and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, **9**, 21–58.

Reichsbürgergesetz (1935). *Deutsches Reichsgesetzblatt*, Jahrgang 1935. https://alex.onb.ac.at/cgi-content/alex?aid=dra&datum=1935&page=1288&size=45.

Ritschl, A. (2019). Financial destruction: confiscatory taxation of Jewish property and income in Nazi Germany. *Unpublished manuscript*.

ROSENSTOCK, W. (1956). Exodus 1933–1939: A Survey of Jewish Emigration from Germany. *The Leo Baeck Institute Yearbook*, **1** (1), 373–390.

SANTOS SILVA, J. and TENREYRO, S. (2006). The log of gravity. *The Review of Economics and Statistics*, **88** (4), 641–658.

SATYANATH, S., VOIGTLÄNDER, N. and VOTH, H.-J. (2017). Bowling for fascism: Social capital and the rise of the Nazi Party. *Journal of Political Economy*, **125** (2), 478–526.

SMALL, K. A. and ROSEN, H. S. (1981). Applied Welfare Economics with Discrete Choice Models. *Econometrica*, **49** (1), 105–130.

SPITZER, Y. (2015). Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States, 1881–1914.

SPOERER, M. (2005). Demontage eines Mythos? Zu der Kontroverse über das nationalsozialistische" Wirtschaftswunder". *Geschichte und Gesellschaft*, **31** (3), 415–438.

STATISTISCHES REICHSAMT (1936). Die Bevölkerung des Deutschen Reichs nach den Ergebnissen der Volkszählung 1933. Heft 5: Die Glaubensjuden im Deutschen Reich. In *Statistisk des Deutschen Reichs 1933*, vol. 451,5, Verlag für Sozialpolitik, Wirtschaft und Statistik, Paul Schmidt, Berlin 1936.

STIFTUNG JÜDISCHES MUSEUM BERLIN (ed.) (2006). *Heimat und Exil: Emigration der Deutschen Juden nach 1933*. Jüdischer Verlag im Suhrkamp Verlag.

STRAUSS, H. A. (1980). Jewish Emigration from Germany: Nazi Policies and Jewish Responses (I). *The Leo Baeck Institute Year Book*, **25** (1), 313–361.

— (1981). Jewish Emigration from Germany: Nazi Policies and Jewish Responses (II). *The Leo Baeck Institute Year Book*, **26** (1), 343–409.

TOMBE, T. and ZHU, X. (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review*, **109** (5), 1843–72.

TRAIN, K. E. (2003). *Discrete choice methods with simulation*. Cambridge university press.

UNHCR (2021). *Global Trends: Forced Displacement in 2020*. United Nations High Commissioner for Refugees.

VIRY, G. (2012). Residential mobility and the spatial dispersion of personal networks: Effects on social support. *Social networks*, **34** (1), 59–72.

VOIGTLÄNDER, N. and VOTH, H.-J. (2012). Persecution Perpetuated: the Medieval Origins of Anti-Semitic Violence in Nazi Germany. *The Quarterly Journal of Economics*, **127** (3), 1339–1392.

— and VOTH, H.-J. (2020). Highway to Hitler. *Unpublished manuscript*.

WALDINGER, F. (2010). Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany. *Journal of Political Economy*, **118** (4), 787–831.

— (2011). Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *The Review of Economic Studies*, **79** (2), 838–861.

WÜNSCHMANN, K. (2010). Cementing the Enemy Category: Arrest and Imprisonment of German Jews in Nazi Concentration Camps, 1933-8/9. *Journal of contemporary history*, **45** (3), 576–600.

ZIMMERMANN, N. M. (2013). The list of Jewish Residents in the German Reich 1933- 1945, federal Archives Germany.