DISCUSSION PAPER SERIES

DP15533

THE REFUGEE'S DILEMMA: EVIDENCE FROM JEWISH MIGRATION OUT OF NAZI GERMANY

Johannes Buggle, Thierry Mayer, Seyhun Sakalli and Mathias Thoenig

DEVELOPMENT ECONOMICS

ECONOMIC HISTORY

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

THE REFUGEE'S DILEMMA: EVIDENCE FROM JEWISH MIGRATION OUT OF NAZI GERMANY

Johannes Buggle, Thierry Mayer, Seyhun Sakalli and Mathias Thoenig

Discussion Paper DP15533
Published 08 December 2020
Submitted 07 December 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre’s research programmes:

- Development Economics
- Economic History
- International Trade and Regional Economics
- Public Economics

THE REFUGEE'S DILEMMA: EVIDENCE FROM JEWISH MIGRATION OUT OF NAZI GERMANY

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 almost the entire Jewish community living 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 in order 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 (in 1935) would have led to a 28% increase in Jewish migration out of Germany.

JEL Classification: F22, N40, F50, D74

Keywords: Refugees, migration policy, Counterfactual History, antisemitism, Nazi Germany

Johannes Buggle - johannes.buggle@unil.ch
University of Lausanne

Thierry Mayer - thierry.mayer@sciencespo.fr
Sciences Po

Seyhun Sakalli - seyhun.sakalli@kcl.ac.uk
King’s College London

Mathias Thoenig - mathias.thoenig@unil.ch
Université de Lausanne and CEPR

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

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

December 7, 2020

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 almost the entire Jewish community living 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 in order 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 (in 1935) would have led to a 28% increase in Jewish migration out of Germany.

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

JEL Codes: F22, N40, F50, D74

*Helpful comments from Alberto Bisin, 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, EIEF, EHA conference in Atlanta, ENS Lyon, EPCS conference in Jerusalem, Fribourg, Geneva, 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, 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.

†Department of Economics, University of Lausanne.

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

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

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

1 Introduction

Violence is a major driver of migration: In 2019, 79.5 million people were forcibly displaced because of wars, civil conflicts, mass-killings, and other forms of persecution (UNHCR, 2020). 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 imposes, 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 in developing countries (the recent wave of Syrian emigration towards Europe being a prominent example) and the discussions about restricting inward migration in developed countries (both in the U.S. and Europe).

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 year when migration was banned). By the end of 1938, more than two-thirds of the Jewish community was still located in Germany despite years of persecution and emigration being encouraged by Nazi authorities. This puzzling fact has attracted a lot of attention from historians who have contrasted two main channels, namely (i) migration frictions and (ii) the underestimation of the actual threat by a community that was well integrated in the German society.[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 in order 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 statisti-

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

cal 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 extensively documented in the migration literature (McKenzie and Rapoport (2010) and Beine et al. (2011) for early contributions). Conditional on migrating, an expanding social network in a destination country facilitates future moves to this destination by lowering all kinds of 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. We 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 and socialization of children at the local level.[3] These communities were spread all across the German territory, were 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 currently 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 migrating population 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 who often live in the same places in ethnically fractionalized societies tend to be exposed to the same unobserved correlated effects (e.g. localized violence and economic deprivation). Our data and context provide a unique setting to tackle these issues. First, Jewish persecution and migration out of Germany has been highly scrutinized by historians. It was recently the topic of an intense data collection effort with the purpose of compensating survivors and nourishing collective memory, which resulted in the dataset we use in this study. Second, we can exploit information at a high level of spatial resolution to control for a large battery of fixed effects that absorb many unobserved correlated effects. Third, cross-city mobility of the German population was very high in the aftermath of the collapse of the German Empire

[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 were public corporations, collected taxes and organized local Jewish life by financing religious and secular institutions, such as synagogues and the more than 5,000 Jewish associations (Gruner and Pearce, 2019).

in 1918. Hence, when Hitler came into power, many peers, friends, and relatives were living in different cities of residence and consequently exposed to different migration incentives at the local level. Our measure of social network builds upon this fact and our focus on distant peers allows us to further filter out potential correlated effects.

Our main dataset offers a very wide coverage of the inter-war Jewish community, both accounting for the location inside Germany and eventual migration destination. However, the information is limited regarding individual characteristics. In particular, the dataset at hand lacks information on education, income, and wealth. This prevents us from assessing the effect of financial constraints on migration decisions. Our causal analysis of how networks affected migration should be immune to the fact that we cannot control for these unobserved individual characteristics because they are likely to be uncorrelated with the sources of identifying variation we use: 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 should capture most of the relevant but omitted individual characteristics. Finally, due to confidentiality purpose, the dataset provides no information on household memberships. This limits our capacity to investigate how strong and weak social ties differentially affect migration prospects.

We begin our empirical analysis with a reduced form estimation of the outmigration decision determinants. 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 more 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. In order 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 straightforward estimation with a lower model explaining destination choices that yields a gravity equation of city-to-country migration, and an upper model explaining 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 peer past migration impacts positively and substantially the likelihood of emigration. A one-standard-deviation increase in the past migration of network members, i.e., the exodus effect, increases the emigration probability by 0.8 percentage points (16% 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 probability of emigration by 2.3 percentage points (44% of the sample mean).

Equipped with those structural estimates, and building on recent advances in the analysis of policy scenarios in trade (often referred to as Exact Hat Algebra), we use the model to conduct a number of counterfactual policy scenarios with a special emphasis on their historical relevance and their concrete policy implementation. The simulation results show very large 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 (in 1935) would have led to a 28% increase in Jewish migration out of Germany (with respect to the observed one). We quantify the migration multiplier, a compact way for assessing the extent of social network spillovers on migration, to be around 0.45 additional refugee per year. Our findings also document that the diaspora and exodus effects are both at work in the data with quantitatively close magnitudes. Finally, our quantification suggests that migration frictions in the destination countries contributed more to the observed 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 describes the data, while 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 in more depth 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 economics literature that investigates the determinants of migration, in particular the push factors of migration. Violent conflict and natural disasters are among the most important push factors for refugees and asylum seekers (Chin and Cortes (2015) and Becker and Ferrara (2019) provide excellent reviews of the literature on forced migration). Recent studies, e.g., Cattaneo and Peri (2016) and Missirian and Schlenker (2017), identify a positive relationship between increasing temperatures and an increase in migration outflows and asylum applications from countries affected by such climatic shocks. The closest paper to ours in this vein of research is Spitzer (2015), which investigates the effect 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 United States between 1881 and 1914. Compared to Spitzer (2015), this study focuses on individual decision of Jews to migrate during the conflict and not on migration decisions taken after the conflict and aggregated at the district level. Moreover, we consider not only the diaspora networks but also the networks of Jews within Germany. Bohra-Mishra and Massey (2011) examine the effect of exposure to violence on individual decision to migrate during the Nepalese civil conflict. However, the authors do not have access to the observed migration decision of *all* individuals affected by the conflict, and do not consider the channel of informational spillovers and social learning as a determinant of migration decision. Besides push factors of migration, a large literature has researched 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). Our results similarly document the important role played by diaspora networks in the context of Jewish out-migration from Nazi Germany.

This paper also relates to a literature on the economic determinants and impacts of persecution of Jews both in historical times and during the Holocaust.[4] Anderson et al. (2016) and Grosfeld et al. (2020) highlight 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. (2018) 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. To our best knowledge, the only other study that looks at the migration of Jews during the Holocaust is Blum and Rei (2018), who document a positive self-selection of refugee Jews in terms of health and human capital in a sample of migrants that travelled from Lisbon to New York between 1940 and 1942.

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; 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.

Overall, our study contributes to these four strands of the literature by documenting the im-

[4]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 Maurer et al. (2005); Kaplan (1999); 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 qualitative 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 uncovers, 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.

portance of social networks for the migration decisions of a well-integrated population for which the perception of the danger was imperfect. Moreover, the individual-level analysis with multiple destination countries allows us to disentangle the different roles of the network within the country of origin from the diaspora networks. Finally, the evidence on the detainment of peers sheds light on persecution as a driver of the spatial patterns of migration.

3 Historical Background and Data

3.1 Historical Background

Jewish Life in Germany. When Hitler took power in 1933, about 520 thousand Jews were living in Germany. They were well integrated economically and culturally, 425 thousand of them were of German origin, about 100 thousand of foreign nationality.[5] 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.[6] Only one-fifth of the Jewish population resided in small villages. The Jewish community was spatially concentrated even in large towns: In Berlin, for example, it concentrated largely in only six out of 20 districts (Maurer et al., 2005). Germany's Jews worked primarily in commerce and trade, finance, civil service, and other service occupations (law, medicine). They were over-represented in higher education institutions (Maurer et al., 2005; Windolf, 2011).

Social relationships within the Jewish community (*"Gemeinde"*) were the cornerstone of Jewish life.[7] Jews interacted intensively with their extended family, Jewish friends, and community members. The *"Gemeinden"* 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 early twentieth century, the synagogue remained a crucial center of social interactions.[8] Loyalty towards the Jewish community was high. Even Jews that were not practicing their faith participated in community meetings 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. To sum, socialization within the community and the extended family created strong

[5]We draw on Maurer et al. (2005) that provide a detailed description of Jewish life in Germany. See also Appendix Section B for additional details of the historical background.

[6]About 55% of the entire Jewish population lived in the top 10 cities in 1933 (Maurer et al., 2005). The distribution of people in our data matches this number perfectly.

[7]While social relationships were located primarily within the Jewish community, many Jews shared the sentiments of patriotism that characterized the German society in the beginning of the twentieth century. About 100,000 Jewish men fought in the German Army at the French or Russian front during World War I, out of which one-third received bravery medal (Totten and Feinberg, 2016).

[8]Some 23,000 Jews converted in the German Empire from Judaism to Christianity. The motives for conversion were often non-religious, as most converts tried to escape discrimination. Despite not adhering to Judaism anymore, the social relationships of converts with their Jewish friends and family remained intact (Maurer et al., 2005).

social ties among Germany's Jews.

Persecution. Immediately after the Nazis rose to power in January 1933, anti-Jewish legislation, state-led antisemitic actions, and violence began.[9] The Nazi government sought to push Jews out of the country, by taking away legal rights and economic opportunities and by isolating them from social life. Just weeks after the new government was elected, a nation-wide boycott of Jewish business took place that 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 unwanted individuals. It was officially framed as being necessary to protect the detained from the "righteous" wrath of the German population. Individuals could be taken into custody by the police (Gestapo, SA, and SS) 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, they 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 strongly 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 antisemitic sentiments using our dataset in Appendix Section C.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 the 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 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.

When the Nazis took power, most of the Jewish population that strongly identified with the German culture and nation remained in the belief 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 Appendix Section B.2, strongly suggests that within the Jewish community, information about incidences of persecution was frequently exchanged. These personal and indirect experiences with antisemitic events provided important information about the extent of persecution.

[9] See Appendix Section B.5 for a description of the main antisemitic events.

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.[10] 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.[11] Thus, migrating implied a future in which living conditions, status, and social relationships were highly uncertain. Anecdotal evidence, presented in Appendix Section B.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 allowed the entry of 27,370 migrants per year from Germany and Austria, independent 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 few countries, such as Shanghai, had very little 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 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.[12] As the world's doors gradually closed and Nazi terror intensified, it became increasingly difficult for Jews to leave Germany.

[10]From May 1934 onwards, migrants had to pay a flight tax ("Reichsfluchtsteuer") of 25 percent if their yearly income exceeded 20,000 Marks (equivalent of 5,000$ in 1934), or if they possessed assets worth 50,000 Marks (equivalent of 12,500$ in 1934). On top of that, after the November pogroms of 1938, Jewish emigrants were also required to pay an emigration levy for assets above 1,000 Mark that they wished to transfer of between one to ten percent ("Auswandererabgabe"). 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 on migrants of 77%.

[11]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 Israel, 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).

[12]Opinion polls conducted in 1938 and 1939 in the US, shown in Appendix Figure B1, illustrate that a majority of Americans was against hosting Jews persecuted in Germany.

3.2 Data on Migration and Deportation of Germany's Jews

We base our empirical investigation on individual-level information on migration, detainment, and deportation of Jews living in Germany in the 1930s. We draw on "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).[13] 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. To this end, in a decade-long data collection effort, the Archives consulted more than 1,000 different sources and over 2.5 million references to carefully reconstruct information on Jews living within the borders of Germany on December 31st, 1937.[14]

The "Resident List" provides biographic information, including first and last name, birth date, birthplace, gender, and place of residence, of about 484 thousand individuals.[15] Particularly valuable for the purpose of our study, it records the migration and/or deportation history of about 285 thousand German Jews: timing of migration movements, the first and second destination countries, and/or the deportation date and place.[16] Moreover, it includes information on whether and when a person was detained ("Schutzhaft"). This rich and unique dataset allows us to study a large portion of the Jewish population of Germany during 1933 and 1941. It enables us to trace the migration timing and destination of German Jews, to identify their social networks within Germany, and to link their migration decisions to their social networks. To the best of our knowledge, we are the first to quantitatively analyze this dataset. Appendix Table C1 describes the data. Panels A and B summarize the biographic characteristics of Jews in the full sample and for a sample of individuals for whom migration or deportation information is available, respectively. Also, Panel B summarizes the available information on migration, deportation, and detainment history of Jews.

To identify potential peers in social networks and compute the pull and push factors of migration that peers face, we geolocated the cities of birth and residence of Jewish individuals. (We describe in detail how we define social networks and identify potential peers in the next section.) Overall, our sample comprises 285 thousand individuals coming from 4,251 birth cities and living in 4,025 cities of residence. Appendix Figure A1 displays the size of the Jewish communities in cities of birth and cities of residence.

[13] Additional data that we use is described in the text and the Appendix.

[14] The Archives drew on a multitude of sources: Information on deportations are taken from the deportation lists created by the Nazi perpetrators and registers of concentration camps. Information on migration movements comes from several main sources: (i) the list of individuals of the Reich Association of Jews in Germany; (ii) membership lists of Jewish parishes; (iii) records including resident files, civil registries, and local chronicles from about 1,000 German municipal archives contacted by the Federal Archives; (iv) foreign archives, Jewish communities, and institutions that possessed information about Jewish emigrants from Germany.

[15] The original "Resident List" contains several thousand additional observations; the 484 thousand individuals are those for which information on the date of birth, birthplace, and city of residence is available. The data on first and last names is available only for sub-set of individuals that are either born before 1919, or dead for at least 10 years, as the German Archival law lifts the privacy protection of these persons.

[16] We investigate the distribution of Jews across cities in our sample by comparing it to city-level data from the official Census of 1933 in Appendix Section A.6. We find that the spatial distribution of Jews across cities in our sample resembles closely that of Jewish communities constructed from city-level data from the official Census of 1933.

Descriptives of Migrants and Deportees. We start by describing several facts about the characteristics of migrants and the aggregate migration flows for the sample of 285 thousand individuals for which information on migration and deportation is available. These descriptive statistics give a first idea of the broad patterns of Jewish migration and can be regarded as a further validity check of some of the well-known facts that the historical literature has described. (i) Compared to non-migrants, migrants were on average younger (by almost 10 years) and more likely to be male (14 percentage points), see Appendix Table C2. (ii) Only few individuals migrated in the early periods, while after 1941 it was almost impossible to emigrate (Appendix Figures C1 and C2). (iii) Jews initially migrated to neighboring countries, such as France and Netherlands, in a false sense of security. About 16% of all emigrants were later deported. 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 (Appendix Table C2). 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 Appendix Table C9).[17] (iv) Before the November pogroms in 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 Appendix Figure C4). (v) Deportation was a death sentence: The average year of death for deported Jews is 1942. As shown in Appendix Table C3, 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, and more aggregate sources of information (in particular Rosenstock, 1956).

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 a rare event in Nazi Germany, 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 incentive to migrate out. On the one hand, this is the result of an expanding network in each destination country that facilitates future emigration to this specific country. Migrant networks can affect the

[17] See Appendix Figure C3 for the total migration by destination country over the entire period from 1933 to 1945.

choice of the destination and migration decisions. 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. In addition, 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.

Taken together, migration spillovers can be decomposed into diaspora and exodus effects. To fix ideas, think of a person called Aaron fleeing Germany to the United States: the migration of Aaron not only increases the future likelihood that his friend David also moves to the U.S. (diaspora effect), but also increases the likelihood that David moves to a country other than the US, for example Palestine (exodus effect). This example also illustrates how our data that includes information on destination choice can be used to discriminate between the diaspora and exodus effects. Indeed, 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. To measure social networks in our data, *we assume that social ties exist between individuals who were born in the same city and are of comparable age*. This is 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 (see first row, Table 1). 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 show that our results are robust and stable when excluding from the estimation sample cities that are particularly large (e.g. Berlin or cities with more than 5 thousand Jews).

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.[18] To address these so-called correlated effects, we restrict the construction of social networks to *Distant Peers* only. We code as distant peers of an individual i all the individuals: (a) covered in our "Resident List" dataset (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); (d) living in a *different city of residence* than i, at least 5 km away from the city of birth. 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) co-move.

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, originating from different cities of birth, have different distant peers. 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.[19] Together, local push and pull factors create spatial variation in outmigration of distant peers across cities. In Section 6.3, we describe in detail how we construct the measures of network pull and push factors that we use in an instrumental framework. Figure 1 visualizes the variations in local persecution (i.e. detainment) and outmigration across cities in 1938.

Table 1 summarizes the characteristics of distant peers' network for the sample of adult Jews living in Germany in 1933. Furthermore, we exclude individuals (roughly one fifth of the sample) whose distant-peer network is below 5 people. The average size of Jewish communities in birth cities is about 3,700 individuals, while the median size is considerably smaller with only 209 individuals (Berlin explaining most of the right-skewness of the size distribution). The mean size of a distant-peer network comprises 152 individuals (median = 43), and peers are living on average about 275 kms apart from each other. We then compute cumulated migration and persecution rates among distant peers (from 1933 up to year $t-1$). Empirically, we measure persecution by incidences of detainment ("Schutzhaft"). As detainment was a rare event, the average cumulative detainment rates were low and range from about 0.1% in 1933 to 5.4% in 1941. In the beginning of the period cumulative network migration were similarly low, with on average only 5.8% of distant peers that have migrated by 1933. The share of distant peers that left Germany increased to 41.2% until the end of 1941.

Limitations of the Data and Network Construction. Data and network construction come with certain limitations that we discuss in Appendix Section A.6. The section describes how we deal

[18]For example, imagine a city in which the Jewish community is economically prosperous. This could influence simultaneously outmigration incentives and the likelihood of victimization, making it difficult to causally identify peer effects. Individuals that live in the same city not only experience similar degrees of local persecution, but other local contextual factors, such as similar costs of migrating out as a result of city's geographic location or local connections with foreign countries. For example, imagine a city close to the Swiss border. In this city, the immigration possibilities of all residents would simultaneously decrease when Switzerland decides to stop Jewish immigration in 1938.

[19]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 Appendix Table C4, 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.

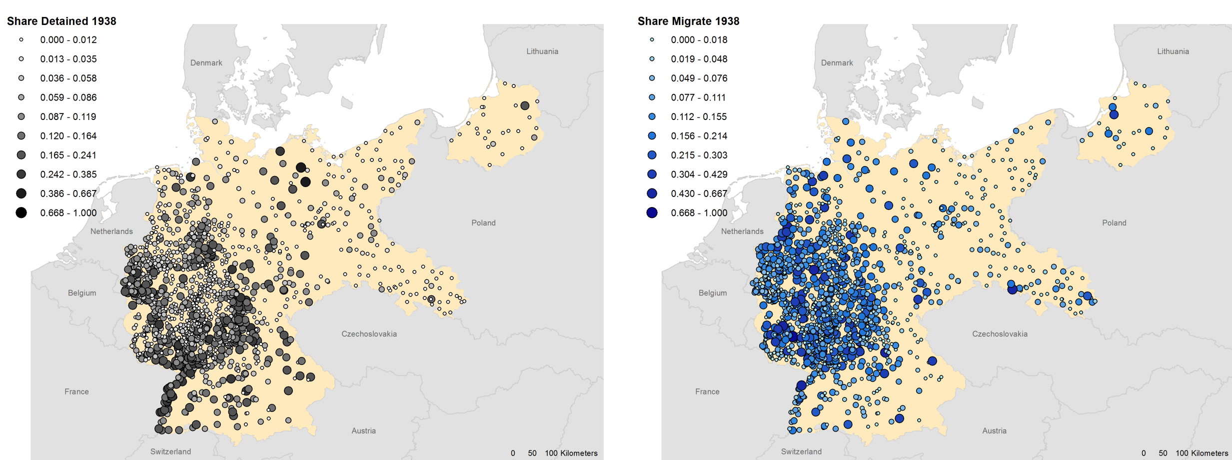
Table 1: Network Descriptives

	Mean	Median	SD	Min	Max	Obs.
Size of Jewish community	3,680	209	7,945	10	24667	201,293
Share of individuals with 5+ distant peers	0.829	1	0.377	0	1	201,293
Size of distant-peer network	153	43	209	5	731	166,929
Distance between individuals and distant peers	273	283	157	7	7,817	166,591
Network detainment rate by 1933	0.001	0	0.007	0	0.200	166,929
Network detainment rate by 1941	0.054	0.047	0.057	0	0.833	166,929
Network migration rate by 1933	0.058	0.047	0.065	0	1	166,929
Network migration rate by 1941	0.412	0.400	0.199	0	1	166,929

NOTE: This table displays network summary statistics. The sample are individuals who were adults in 1933.

with missing information on migration and deportation (representativeness of the sample), investigates the characteristics of distant peers with respect to those of decision makers (selection) and addresses the possibility of spatial spillovers between distant peers and decision makers (correlated effects).

Figure 1: Detainment and Migration across Cities of Residence in 1938



NOTE: This figure displays the variation in the share of detained in 1938 (left figure), and the share of emigrants in 1938 (right) across cities of residence. For better visibility, cities with less than 10 residents are excluded. The map represents Germany on December 31st, 1937.

5 Outmigration and Social Networks: Preliminary Analysis

In this section, we conduct a preliminary "reduced-form" analysis of the data. This first pass on the data has the advantage of relying on a simple econometric framework. We put relatively more emphasis on the threat effect in migration decisions, namely how social networks, by aggregating information on the extent of persecution, affects migration incentives. By contrast, 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.

A theory-free empirical model. The structure of our data naturally calls for a discrete time duration model that we combine with a model of social interactions. We transform the individual migration information into single-spell data, 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 1 in the year t in which individual i migrates. 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 simply 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.[20] We study migration decisions from 1933 until 1941, as 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 $r(i)$, born in city $b(i)$, with network $n(i)$, at time t:

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

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 compatible with our data structure, the duration model, and the structural model we outline in the next section, LPM has the advantage of accommodating high-dimensional fixed effects and a rich clustering structure and offers a transparent interpretation of the coefficients of interaction terms. With the goal of providing a simple and easily interpretable first pass on the data, we therefore present only LPM estimation results in the main text (Tables 2 and 3); and relegate non-linear estimations, which yield comparable results when expressed in terms of marginal effects, to Appendix Table C7.

Our first main variable of interest is $\texttt{Mignet}_{n(i)t}$ that measures the post-1933 share of distant peers who already left Germany at 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. 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

[20] 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. The return migration stopped, when in 1935 the German authorities started to place returning migrants in training/concentration camps.

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. In most specifications, the vector $\texttt{FE}_{r(i)t}$ corresponds to city-of-residence × 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 1,400 cities of residence, and over 12,000 city of residence × year combinations, the estimation of fixed effects is extremely demanding from the data.

The estimation sample is composed of all individuals who are adults in 1933 (at least 18 years old) and who are living in Germany. The time period ranges from 1933 to 1941. 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 × 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 only 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 2 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 × 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), while those born outside Germany – who were on average poorer (Kaplan, 1999) – are less likely to outmigrate. 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 cumulated 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 (0.15) translates into a 0.85 percentage point increase in the migration probability (16% of the sample mean). Regarding the threat effect, a one-standard-deviation increase in peers' past persecution (0.04) increases the migration probability by 0.15 percentage

Table 2: Outmigration Decision

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	Migration Decision						
						Pre-1938	Post-1938
Migration of network members	0.060***	0.057***	0.056***	0.057***	0.056***	0.071***	0.019***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
Detainment in city of residence	0.052**						
	(0.021)						
Detainment of network members	0.046***	0.037***	0.072***	0.042***	0.078***	0.077***	0.004
	(0.011)	(0.010)	(0.013)	(0.010)	(0.013)	(0.016)	(0.012)
× Jewish Name Index (> *Median*)			-0.063***		-0.063***		
			(0.013)		(0.013)		
× Ever detained				-0.073**	-0.083**		
				(0.037)	(0.037)		
Jewish Name Index (> *Median*)			-0.004***		-0.004***		
			(0.000)		(0.000)		
Ever-detained	0.028***	0.021***	0.021***	0.023***	0.024***	-0.037***	0.049***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Age	-0.044***	-0.040***	-0.041***	-0.041***	-0.041***	-0.044***	-0.049***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.009)
Age squared	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Male	0.026***	0.025***	0.025***	0.025***	0.025***	0.024***	0.029***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Born outside Germany	-0.003**	-0.002**	-0.002**	-0.002**	-0.002**	-0.001	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Year FE	✓						
CoR × Year FE		✓	✓	✓	✓	✓	✓
Observations	1,025,497	1,024,976	984,414	1,024,976	984,414	754,468	270,508
R-squared	0.06	0.10	0.10	0.10	0.10	0.07	0.13
Mean of dependent var.	0.052	0.052	0.052	0.052	0.052	0.039	0.088
SD of dependent var.	0.222	0.222	0.222	0.222	0.222	0.194	0.283

NOTE: LPM estimation with the dependent variable being an indicator for migration. The unit of observation is an individual in year t. The sample consist of individuals who were adults 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 × year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

points (2.8% 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 (40% 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 update their beliefs more when observing 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 who are well integrated to the German society.

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 2, 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.[21] We interact network detainment with an indicator for whether the Jewish Name Index is above the median of the distribution. The estimated coefficient of this interaction, displayed in column 3, is negative and statistically significant. The negative coefficient suggests that individuals with a traditionally Jewish (German) first name are less (more) likely to respond to their peers' victimization. This is in line with anecdotal evidence that highly integrated Jews who identified strongly with the German society felt relatively safer (e.g., Nicosia and Scrase, 2013; Heusler and Sinn, 2015). The main effect of a more Jewish sounding name is also negative, as individuals with a higher JNI were less educated (see Appendix Table C6), 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.[22] 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 Table 3, we consider characteristics of residence cities interacted with peers' persecution. In

[21] See the Appendix Section A for details on the construction of the the Jewish Name Index. In Appendix Table C6 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. Appendix Table C5 lists the top and bottom 20 first names in terms of Jewish Name Index (JNI) by gender.

[22] The marginal effect of detainment among network members is negative only for 6 individuals who were detained themselves.

Table 3: Learning: City Characteristics

	(1)	(2)	(3)
Dependent variable:	Migration Decision		
Migration of network members	0.057***	0.056***	0.057***
	(0.007)	(0.007)	(0.007)
Detainment of network members	0.038***	0.052***	0.041***
	(0.010)	(0.015)	(0.011)
× Medieval Pogrom		-0.036*	
		(0.020)	
× NSDAP Votes 1928			-0.477*
			(0.247)
× Share Jews 1933	-0.009**	-0.011***	-0.010**
	(0.004)	(0.004)	(0.004)
Observations	966,799	934,826	963,510
R-squared	0.09	0.09	0.09
Mean of dependent var.	0.052	0.052	0.052
SD of dependent var.	0.222	0.222	0.222

NOTE: LPM estimation with the dependent variable being an indicator for migration. The unit of observation is an individual in year t. The sample consist of individuals who were adults in 1933. Detainment of network members measures the cumulative share of network members that have been detained until year t. All continuous variables in interactions are demeaned. All regressions control for age, age squared, gender, personal detainment, and a dummy for whether the individual was born outside Germany, the interaction between network detainment with (log) city population in 1933, with the (log) distance to the German border, as well as city-of-residence × year fixed effects. Standard errors clustered at the city-of-birth × year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

column 1, we document that detainment among network members has a smaller effect for individuals living in cities with a larger local Jewish community. This negative interaction suggests that there was some competition between informational sources: In localities where information provided by local Jewish networks is available, private information from distant peers is less important. In columns 2 and 3, we find that detainment among network members interacts negatively with city-level antisemitism: individuals in cities that experienced a Medieval pogrom or in cities with above median vote shares for the NSDAP in the 1928 elections, responded less to their peers' persecution. We again interpret these interaction terms as a surprise effect. In cities where historical and recent antisemitism are ingrained in the collective memory of the local community, information about the danger coming from distant peers is less relevant.

In addition, in Appendix Table C8, 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 host 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 20). 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.[23] 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 host 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 in year t by an individual i when locating in country $d \in D$. 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,

[23]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.

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 $B_m = \{\text{USA, GBR, FRA, ...}\}$. 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.

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 attractiveness of destination country d (e.g. economic prospects) — a component that is estimated in our empirical analysis with destination $\times$ year fixed effects.[24] 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 modelled 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, which we denote 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

[24] 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.

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 already left Germany at 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), $\texttt{Detainment}_{n(i)t}$, corresponds to the share of peers who were detained. 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 idiosyncratic preference term:

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

With this distribution, 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 ($W_{i\text{DEU}t}$) reduce 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. First, 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.[25] An important theoretical requirement is $\lambda_2 \leq \lambda_1$ 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 host 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.[26] 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. This is true for both upper level variables $W_{i\text{DEU}t}$ and the expected maximum utility of migration $\lambda_2 I_{it}$.[27]

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

[26] 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 .

[27] 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.

Estimating 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, then included 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 × 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 components: $\ln \tau_{idt} \equiv \ln \tau_{rdt} + \ln \tau_{bdt}$ 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.[28] 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 cumulated flows of individuals from r who

[28] 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. Appendix Section C.7 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.

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,[29] and the migration frictions with geodesic distances, such that $\ln \tau = \rho \ln \texttt{dist}$.[30] The resulting estimating equation is:

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

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; it captures the attractiveness of each host country ($\text{FE}_{dt} = A_{dt}/\lambda_2$).[31] 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 (opposite of) the inclusive utility, 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 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).[32] Accounting for the zeroes is particularly important in the

[29] 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 32.

[30] 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.

[31] 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 Appendix Section D.1.

[32] The combination of cities of residence and cities of birth with the 32 destination countries result in a total of 907,458 $rbdt$ cells, for which there has been at least one migrant from rb in year t to any destination. If nobody migrates from rb

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).[33]

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 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

in t, the left-hand side variable, the share of migrants to d is undefined. Out of those, 14,904 (or 1,6%) of cells experienced a positive migration flow to country d in year t, for the remaining we impute a zero-migration flow.

[33] 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).

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.[34]

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.

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.[35] The standard errors are clustered at the city-of-residence $\times$ city-of-birth $\times$

[34] This problem of zeroes 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 10,431 individuals, and the mean of Jewish communities in 1933 in cities of birth, pop_{b,t_0}, is 492. In the end, 43% 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 19% for non-zero diaspora cells from the city of birth b in year t.

[35] 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.

year level. The sample includes individuals that are adults in 1933 (at least 18 years old) and that are living in Germany. 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 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 tp PPML regressions, first on positive flows, and then on the entire sample including *rbdt* cells with zero migration flows. Columns 5 and 6 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 0.094 (0.015) 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

Table 4: Triadic Gravity Estimation Linear Distances (Lower Stage)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:	Share Migrants (log)	Share Migrants (log)	Share Migrants > 0	Share Migrants	Share Migrants > 0	Share Migrants
	All Cities				Top 100 Cities	
	OLS	OLS	Poisson	Poisson	Poisson	Poisson
log (Cumulative # $\text{Migrants}_{\text{rdt}}$+1)	0.087***					
	(0.009)					
log (Cumulative # $\text{Migrants}_{\text{bdt}}$+1)	0.177***					
	(0.013)					
$\text{Diaspora}_{\text{rdt}}$ = Cumulative Share of $\text{Migrants}_{\text{rdt}}$		0.269	0.322*	5.294***	1.399**	6.712***
		(0.240)	(0.194)	(0.212)	(0.570)	(0.777)
$\text{Diaspora}_{\text{bdt}}$ = Cumulative Share of $\text{Migrants}_{\text{bdt}}$		2.265***	1.425***	5.416***	4.081***	9.584***
		(0.280)	(0.218)	(0.221)	(0.661)	(1.040)
(log) $\text{Distance}_{\text{rd}}$	-0.066***	-0.094***	-0.076***	-0.473***	-0.064***	-0.385***
	(0.013)	(0.013)	(0.011)	(0.013)	(0.018)	(0.026)
(log) $\text{Distance}_{\text{bd}}$	-0.001	-0.015*	-0.015**	-0.134***	-0.034**	-0.087***
	(0.009)	(0.009)	(0.008)	(0.011)	(0.014)	(0.025)
Observations	14,904	14,904	14,904	907,458	7,634	155,998
R-squared	0.82	0.81				
Pseudo R-squared			0.75	0.18	0.75	0.17
Avg. Elasticity $\text{Diaspora}_{\text{rdt}}$	0.087	0.006	0.007	0.065	0.021	0.062
Avg. Elasticity $\text{Diaspora}_{\text{bdt}}$	0.177	0.058	0.037	0.126	0.071	0.105
Mean of dependent var.	-1.464	-1.464	0.317	0.032	0.240	0.033
SD of dependent var.	0.984	0.984	0.190	0.168	0.186	0.158

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 adults in 1933. Diaspora_{rdt} (Diaspora_{bdt}) is defined as the cumulative share of migrants from a residence city (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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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 zeroes in columns 4 and 6.

Turning to our main variable 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 city-of-residence diaspora network is more noisy columns 2 and 3. In our preferred specification, PPML with zero flows, both diaspora networks are largely significant. 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 .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. However, the adjustment distorts the distribution of the underlying variable. This is why 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 substantially lower, but still significantly positive. The results of column 4 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 0.65%. The effect of networks from city of birth is larger with a elasticity close to 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 zeroes 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 zeroes in the choice set increases connectivity, since the absence of migrants in a particular destination provides a valid comparison point to positive flows.[36] 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.

Appendix Table C12 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 zeroes, this table would constitute of columns uniformly filled with the total number of *rb* that year. The

[36] 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.

consequence of including the zeroes can also be seen in Appendix Figure C8 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 corresponding to column 3 of Table 4, while panel (b) presents results from column 4. The version including the zeroes corresponds to a better fit with less variance as expected from the insights of the labor literature. Including the zeroes therefore is our preferred specification.

Columns 5 and 6 of Table 4 reduce the sample to the set of birth and residence cities that are 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 zeroes (column 6 compared to 4). Furthermore, the elasticities reported in the bottom of the table are very close between these columns 6 and 4. This stability reinforces our decision to consider the specification of a PPML regression with zeroes included as our preferred one.

Time-varying distance effects. While Table 4 controls for the distances between the residence and birth cities and the destination, Table 5 adds interaction terms between distances and year dummies to evaluate the evolution of the effect of distance over time. Results are very similar, with the pattern of distance effects showing an upward trend: Figure C7 shows that distance matters less over time and becomes very noisy after the start of WWII hostilities. Since Table 5 allows for more flexibility in the impact of migration frictions, we adopt the whole set of coefficients and parameters obtained through the estimation of column 4 as our baseline for computation of the inclusive utility.

Revealing the attractiveness of countries. From the gravity estimation we also obtain the estimated fixed effects FE_{dt}, which measure the attractiveness of destination d in year t.[37] The bottom panel of Appendix Table C9 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 the top panel of the same table.[38]

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 6, we therefore show results of an auxiliary regression of the estimated fixed effects FE_{dt} on the log income per capita of the countries.[39] From equation (10), the structural interpre-

[37]The fixed effects FE_{dt} are based on column 4 of Table 5.

[38]The last row of appendix Table C9 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 .54 to .86.

[39]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.

Table 5: Triadic Gravity Estimation (Lower Stage): Time-varying Distance Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables:	Share Migrants (log)	Share Migrants (log)	Share Migrants > 0	Share Migrants	Share Migrants > 0	Share Migrants
	All Cities				Top 100 Cities	
	OLS	OLS	Poisson	Poisson	Poisson	Poisson
log (Cumulative # $\text{Migrants}_{\text{rdt}}$+1)	0.087***					
	(0.009)					
log (Cumulative # $\text{Migrants}_{\text{bdt}}$+1)	0.182***					
	(0.013)					
$\text{Diaspora}_{\text{rdt}}$ = Cumulative Share of $\text{Migrants}_{\text{rdt}}$		0.251	0.339*	5.459***	1.455**	7.360***
		(0.241)	(0.195)	(0.216)	(0.578)	(0.821)
$\text{Diaspora}_{\text{bdt}}$ = Cumulative Share of $\text{Migrants}_{\text{bdt}}$		2.382***	1.504***	5.526***	4.527***	10.324***
		(0.283)	(0.219)	(0.222)	(0.677)	(1.071)
Observations	14,904	14,904	14,904	907,458	7,634	156,484
R-squared	0.83	0.81				
Pseudo R-squared			0.75	0.18	0.75	0.17
Avg. Elasticity $\text{Diaspora}_{\text{rdt}}$	0.087	0.005	0.007	0.067	0.022	0.068
Avg. Elasticity $\text{Diaspora}_{\text{bdt}}$	0.182	0.061	0.039	0.129	0.078	0.113
Mean of dependent var.	-1.464	-1.464	0.317	0.032	0.240	0.033
SD of dependent var.	0.984	0.984	0.190	0.168	0.186	0.158

NOTE: The unit of observation is a city-of-residence × city-of-birth × country in year t. The sample consists of individuals who were adults in 1933. Diaspora_{rdt} (Diaspora_{bdt}) is defined as the cumulative share of migrants from a residence city (city of birth) to the destination country. All regressions control for country × year fixed effects, city-of-residence × city-of-birth × year fixed effect, as well as distance from city-of-residence to d × year fixed effects, and distance from city-of-birth to d × year fixed effects. Standard errors clustered at the city-of-residence × city-of-birth × year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

tation of the coefficient is $1/\lambda_2$, which we estimate to be 3.59 in column 1.[40] 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 US 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 co-

[40] 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.

Table 6: Determinants of Country Attractiveness

	(1)	(2)	(3)	(4)
Dependent variable:	Country Attractiveness (FE_{dt})			
(log) GDP per capita	3.594*	3.811**	3.099*	2.723*
	(1.824)	(1.685)	(1.528)	(1.353)
(log) GDP per capita × Post 1938		-1.210**	-0.991*	-1.016**
		(0.489)	(0.492)	(0.492)
(log) Distance × Post 1938			0.425**	0.236
			(0.187)	(0.194)
German Occupation				-1.541***
				(0.484)
Observations	238	238	231	229
R-squared	0.60	0.62	0.64	0.65
Mean of dependent var.	-2.68	-2.68	-2.64	-2.65
SD of dependent var.	2.17	2.17	2.18	2.19

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

efficients 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 4 of Table 5 with formula given in equation (13).

Panel A of Table 7 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 with identical estimation methods in column 5, panel A of Table 7 and in column 2 of Table 2).

The coefficient on inclusive utility is interpreted as a marginal effect in columns 2–5. The method of estimation leaves the effect quite unaffected. However, 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 $\lambda_2/\lambda_1 \simeq 0.68$, confirming that our nested logit structure is compatible with revealed preferences. Finally, combined with our lower-level estimate of $\lambda_2 = 1/3.59 = 0.28$, we can reveal $\lambda_1 = 0.28/0.68 = 0.41$. The magnitude of the estimated coefficients implied by column 5 are large: a one-standard-deviation increase in the migration of network members (0.15), i.e., the exodus effect, increases the outmigration probability by 0.81 percentage points, or 16% of the sample mean. A one-standard-deviation increase in the inclusive utility (1.10), which encompasses not only the diaspora effect but also destinations' attractiveness and migration frictions, increases the probability of outmigration by 2.3 percentage points, or 44% 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 column 1 in Table 2). 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 in Germany up to year $t-1$ to generate the second shifter:

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

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).[41]

Panel B of Table 7 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 Appendix Table C14 and confirm the statistical power of both instruments.[42]

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 (16.34) and column 5 (22.46) 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 $\lambda_1 \simeq 0.28/0.81 = 0.35$. 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 7, Panel A. We briefly summarize the checks we perform here and discuss them in detail in Appendix Section C.11. In Appendix Table C15, we explore how missing information on

[41] 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.

[42] We present the intention-to-treat (or reduced-form) results where the two endogenous network effects are replaced by their shifters in Appendix Table C13. 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.

Table 7: Outmigration Decision with Inclusive Utility

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Migration Decision				
Panel A: OLS	Logit coef.	Logit dy/dx	Cloglog dy/dx	LPM coef.	LPM coef.
Detainment of network members	0.427***	0.020***	0.019***	0.040***	0.036***
	(0.145)	(0.007)	(0.007)	(0.011)	(0.010)
Migration of network members	0.209***	0.010***	0.006*	0.056***	0.054***
	(0.078)	(0.004)	(0.003)	(0.007)	(0.007)
Inclusive Utility	0.681***	0.031***	0.030***	0.037***	0.021***
	(0.042)	(0.002)	(0.002)	(0.003)	(0.004)
R-squared				0.06	0.10
Pseudo R-squared	0.12	0.12			
Panel B: Instrumented	Logit coef.	Logit dy/dx	Cloglog dy/dx	LPM coef.	LPM coef.
Detainment of network members	0.538***	0.025***	0.023***	0.047***	0.041***
	(0.164)	(0.007)	(0.007)	(0.013)	(0.011)
Migration of network members	2.584***	0.118***	0.102**	0.190**	0.168**
	(0.905)	(0.041)	(0.040)	(0.085)	(0.073)
Inclusive Utility	0.807***	0.037***	0.034***	0.039***	0.075***
	(0.127)	(0.006)	(0.006)	(0.009)	(0.027)
R-squared				0.02	0.01
Pseudo R-squared	0.12	0.12	.		
Year FE	✓	✓	✓	✓	
CoR x Year FE					✓
Observations	1,025,497	1,025,497	1,025,497	1,025,497	1,024,976
Mean of dependent var.	0.052	0.052	0.052	0.052	0.052
SD of dependent var.	0.222	0.222	0.222	0.222	0.222
F-Stat				16.01	22.13

NOTE: The dependent variable is an indicator for migration. The unit of observation is an individual in year t. The sample consists of individuals who were adults 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 Appendix Table C7. Marginal effects reported in columns 2 and 3. Standard errors clustered at the city-of-birth $\times$ year in parentheses. Panel B: Columns 1, 2 and 3 use control function approach, while columns 4 and 5 estimate two-stage least squares. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

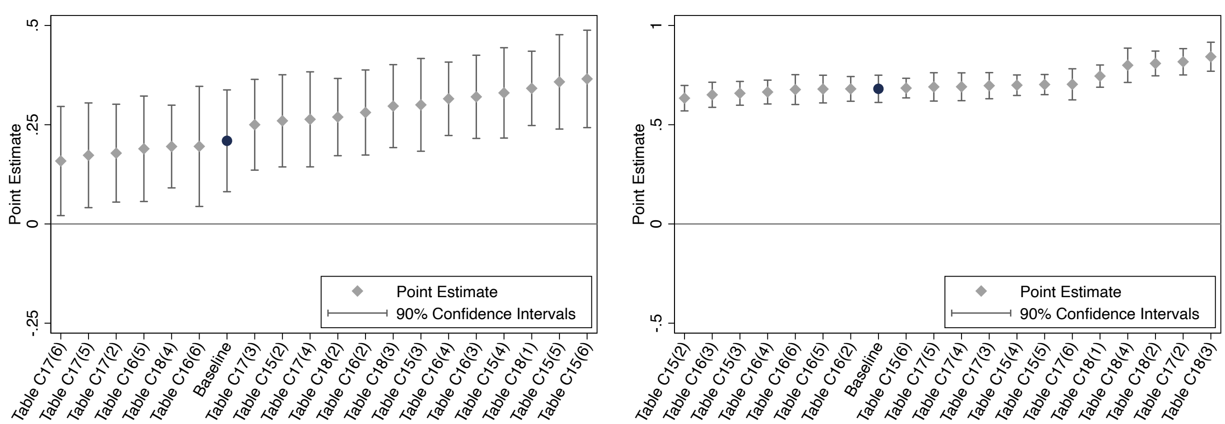
migration and deportation status for some distant peers affects the estimated effects of migration spillovers. In Appendix Tables C16 and C17, we check the sensitivity of our estimates to dropping from the sample cities of residence and birth in the lower and upper tails of the distribution size. In Appendix Table C18, we focus on a sample of decision makers who are not living anymore in their city of birth (they are internal movers like their distant peers), and impose a minimum distance between the cities of residence of decision makers and their distant peers.

Figure 2 reports the set of point estimates obtained from these 19 robustness checks. Each light gray diamond represents a point estimate along with 90% confidence intervals from a different robustness check; the dark blue circles represent our baseline point estimates that we benchmark on. Overall, the estimated effects are robust and stable across specifications. Our preferred specification from column 1 of Table 7 yields point estimates that are located slightly below the median of the distribution of estimates.

Last, we perform a placebo test. We shuffle observed networks across decision makers. To each decision maker, we randomly assign the migration rate of network members and inclusive utility, as well as the detainment rate of network members, which are observed for another individual. We perform this random assignment process 1,000 times and re-estimate the specification reported in column 1 of Table 7, Panel A. Figure C9 presents the coefficients on migration rate of network members (panel (a)) and inclusive utility (panel (b)) obtained from this exercise. The blue lines present the distribution of coefficients obtained from random assignments of peer networks to decision makers, whereas the red vertical lines represent our benchmark estimates of the coefficients. This figure shows that the likelihood of obtaining the benchmark estimates by chance is less than one in a thousand.

Figure 2: 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 obtained from the robustness checks presented in Appendix Section C.11. Each diamond and circle sign represent point estimates obtained from a different regression. Under each of them, we indicate from which Appendix Table (C15–C18) and column (1–6) the point estimate originates.

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. On the one hand, wealthier individuals could pay more easily for the costs associated with emigration, such as visa and travel costs. On the other hand, the Nazi government sought to benefit from Jewish emigration by levying several taxes to expropriate the wealthiest migrants (see footnote 10). Similarly, a high level of education did not necessarily translate into better migration prospects as some professional skills were less portable abroad (e.g. lawyers) than others (e.g. scientists, farmers).

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. 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 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. Appendix Table C19 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, in particular when city-of-residence $\times$ year fixed effects are included. Moreover, such high dimensions of fixed effects are hardly compatible with the non-linear estimation of our structural model. For this reason, our preferred empirical model abstracts from including individual fixed effects.

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 as a function of the observed true levels of the same flows, together with the policy change and a very parsimonious set of structural parameter. 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, dividing computing time by roughly two. Second and most important, because it does not require to solve the model, it is extremely economical in terms of data requirement. This is where the approach relies critically on the CES structure. With CES demand 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*. The third and related advantage of the EHA approach is that it starts from actual data patterns (migration shares in our case). It therefore 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. This way of doing things 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 case where the fit of model to data is good, some of the predicted shares can be quite far from real ones. For instance, some countries might attract no migrant despite the model predicting it should. If the policy is a tightening of the migration policy, the counterfactual will 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 a diaspora effect, the error will contaminate later years. There is however one major disadvantage of EHA: 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, there are some flows that become positive or some flows that die out. In our view, the benefits highlighted above dominate this flaw.

All the computational details of the exact hat algebra are relegated to the Online Appendix D.1. The structural elasticities needed for running the counterfactual simulations are retrieved from the estimation results of the lower-model (Table 4, column 4), the instrumented version of the upper-model (Table 7, Panel B, column 1), and the table showing the determinants of country attractiveness (column 5 of Table 6). These structural parameters take the following values:

$$\lambda_1 = 0.345; \lambda_2 = 0.278; \alpha_1 = 1.47; \alpha_2 = 1.51; \beta = 0.891; \gamma = 0.185; \tilde{\rho}_1 = 0.131; \tilde{\rho}_2 = 0.037.$$

As a sensitivity test, we also run simulations with the alternative parameter values, $\lambda_1 = 0.408$, $\beta = 0.085$, $\gamma = 0.174$, that are recovered from the non-instrumented version of the upper model (Table 7, Panel A, column 1).

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; (5) Early perception of the threat. Counterfactuals 1 to 3 are modeled as relevant changes in attractiveness in destination countries in the lower-level equation (8). Our counterfactuals use the hat notation to denote the proportional change in the relevant variables, for instance $\widehat{a}_d \equiv a'_d / a_d$, is the ratio of counterfactual over initial attractiveness of country d. Counterfactual 4 corresponds to a change in bilateral migration cost in the same equation. Counterfactual 5 is engineered as a changing perception of detainment in the upper-level equation (11). We discuss each scenario in detail below, with a special emphasis on historical relevance and concrete policy implementation. Note that the exposition of the first scenario is a bit lengthier, not because we view it as more prominent, but because we use it as an illustrative case of the various mechanisms at play.

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 already 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 to the United States (Breitman, 2013). Political attempts to open borders, such as the Wagner–Rogers Bill in the U.S. Congress that proposed to allow 10 thousand children per year to come to 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 scenarios, we do not aim here at detailing the concrete implementation of this policy. Hence, in our simulation procedure, we simply reverse engineer the change in attractiveness of the United States in 1936 such as to generate this additional inflow: This implies to set $\widehat{a}_{USA,1936} = 2.2318$. In years $t > 1936$, attractiveness remains unchanged ($\widehat{a}_{USA,t} = 1$). The required change in attractiveness is large and we have little to say on the policy tools that could have been used to achieve it.

Table 8 reports the observed and counterfactual migration flows that result from the additional 5,000 migrants going to the U.S. in 1936. In the data for this same year, we observe a total of 2,082 Jewish migrants to the U.S. This number is increased to a total of 7,082 migrants in 1936. By the end of the period in 1941, we observe a total of 61 thousand migrants who had left Germany, while the total migration out of Germany in the counterfactual simulation is significantly larger and sums up to more than 74 thousand emigrants. This implies an increase of about 13 thousand additional migrants out of Germany to all destinations during the period 1936 to 1941. Thus, we find a multiplier of 2.7, or that each additional refugee to the United States in 1936 would have increased emigration by 0.45 refugees per year across all destinations.

In Table 9, we decompose the overall migration spillover into its two components, the exodus and diaspora effects, by re-running counterfactual simulations after shutting down each effect sequentially (i.e. zeroing their respective coefficients (α_1, α_2) and β). We see that they both contribute with a quantitatively similar magnitude. Each component alone increases total migration from an observed 61 thousand to about 69 thousand (exodus) and 67 thousand (diaspora) in each counterfactual simulation. The total number of migrants to the United States doubles from 12.5 thousand

Table 8: Increasing migration to the U.S. in 1936 by 5,000 individuals

	1936	End of 1941
Observed		
Pop. in Germany	160,847	113,537
Total Migration	18,922	61,065
Migration US	2,082	12,654
Migration Rest of World	16,840	48,411
Counterfactual		
Pop. in Germany	160,847	100,020
Total Migration	23,434	74,582
Migration US	7,082	25,292
Migration Rest of World	16,352	49,290
Outmigration Multiplier		
All post-1936 Years	0.91	2.71
Average by Year	0.91	0.45

NOTE: This table displays the observed and counterfactual migration flows resulting from a one-shot increase of 5,000 migrants to the United States in 1936.

in the observed data to 25 thousand in the counterfactual scenario. As can be seen from columns 2 and 3, the diaspora effect plays a larger effect in increasing migration to the United States: Following the policy intervention, the size of the diaspora in the U.S. increases immediately and then this pulls even more migrants to the U.S. in the following years. As expected, the exodus effect plays a larger role for migration to countries outside of the United States.

This experiment of a one-shot increase in migrants to the United States has important policy implications. The counterfactual analysis implies that a unilateral opening of borders does not just reallocate emigrants from other countries to the United States, migrants who would have fled anyway out of Germany (a pure substitution effect), but it increases total migration out of Germany and therefore has the potential to save lives.

Table 9: Spillover Decomposition

		Counterfactuals		
Total Outmigration 1941:	All spillovers	Exodus Only	Diaspora Only	No spillover
All Destinations	74,582	69,161	67,411	63,905
US	25,292	18,683	21,860	17,102
Rest of World	49,290	50,478	45,551	46,803

NOTE: This table displays the observed and counterfactual migration and the decomposition of migration spillovers for all destinations, the United States, and the Rest of the World, resulting from a one-shot increase of 5,000 migrants to the United States in 1936.

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 Arabic community against the British administration with the goal 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 as we observe a threefold reduction in the estimated attractiveness of Palestine after 1935. What if Palestine had stayed open and a similarly attractive place to migrate to? To study this question we impute the attractiveness of Palestine in 1935 for the subsequent years $t > 1935$ such that $\hat{a}_{PAL,t} = a_{PAL,1935}/a_{PAL,t}$. As shown in Table 10, about 8 thousand additional migrants would have left Germany in this counterfactual scenario (14% increase compared to the status quo).

Table 10: Summary of Counterfactual Scenarios

Scenario from 1936 onwards	Jewish Outmigration Germany 1933–1941			
	Instrumented		Non-instrumented	
	Total	% Change	Total	% Change
Status Quo	61,065	-	61,065	-
1. 5,000 additional migrants to U.S. in 1936	74,582	+22.1%	67,041	+9.8%
2. Non-Closing of Palestine	69,571	+13.9%	65,287	+6.9%
3. Removing work restrictions	78,266	+28.2%	67,770	+11%
4.a. Travel Subsidy: Subsidy to U.S. only	67,010	+9.7%	64,212	+5.2%
4.b. Travel Subsidy: Subsidy to all Port Destinations	72,926	+19.4%	66,992	+9.7%
5. Post-Nuremberg Perception of Threat	66,830	+9.4%	63,124	+3.4%

NOTE: This table displays cumulated migration of Jews out of Germany in 1941 in the different counterfactual scenarios. The initial population under consideration is composed of 174,603 adults. All interventions are implemented from 1936 onwards (after Nuremberg laws).

Scenario 3: Removing work restrictions. Next, we study how policies that limit the economic opportunities of migrants 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". Using an auxiliary regression, where the estimated attractiveness of a country-year is explained by restrictions, we validate in Appendix Table C20 that a high level of restrictions indeed lowered the flows of refugees. We then ask, what if countries removed their employment restrictions on employment after the Nuremberg Laws (i.e. from 1936 onward), instead of making access to the labor market difficult? Using the auxiliary regression, we compute the counterfactual attractiveness of destination countries that results from removing labor market restrictions (i.e. moving all labor market policies to 1). The quantitative effect on total migration that results from the implied increases in attractiveness is large: As shown in Table 10, compared to the status quo, about 17 thousand additional migrants would have left Germany in this counterfactual scenario, a 28% increase.

Scenario 4: Travel subsidies. One instrument to help Jewish migrants that was discussed around 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, ticket costs for oversea travel were substantial: Transport costs ranged between 150 and 1320 Reichsmark, which amounted to between 10% to 100% percent of the average yearly income per capita in Nazi Germany.[43] To ease the financial burden of emigrating, organizations such as the American Jewish Joint Distribution Committee also 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 from Löwenthal and Oppenheimer (1938) for 42 boat trips from Europe to overseas destinations with which we can calculate the travel cost by boat in Reichsmark per km. Appendix Table C21 displays the regressions of trip prices on distances between ports. We then reduce ticket prices to each destination by 50% (on average about 300 Reichsmark) and calculate the equivalent reduction in distance from Germany. We consider first a unilateral subsidy to the U.S. only, and second a coordinated subsidy that reduces the distance to all oversea port destinations at the same time as it could have been the outcome of a coordinated policy effort.

As shown in Table 10, a 50% reduction of prices to the U.S. in 1936 results in an increase in total migration by about 6 thousand migrants to a total of 67 thousand. A subsidy of 50% of the ticket prices to all overseas destinations has an effect that is even larger, as almost 73 thousand individuals would have migrated out in this scenario (implying an increase of 19% compared to the status quo).

In Table 11, we enter into the details of comparing between the effect of a unilateral subsidy for travel to the US, to a subsidy implemented in a coordinated fashion to many countries. We observe that while total migration to the rest of the world increases substantially under the coordinated policy, the total migration to the U.S. resulting from either policy is very similar. This implies that there is very little substitution from the U.S. to other countries in the coordinated policy scenario, likely because the U.S. is among the most attractive destinations.

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

Table 11: Unilateral versus Coordinated Policy

Scenario	Year	Migration to U.S.	Migration Rest of the World
4.a. Travel Subsidy: Subsidy to U.S. only	1941	18,783	48,227
4.b. Travel Subsidy: Subsidy to all Port Destinations	1941	18,938	53,988

NOTE: This table displays cumulated migration of Jews to the U.S. and the Rest of the World in 1941 as a result of subsidizing travel. The initial population under consideration is composed of 174,603 adults.

Scenario 5: Early perception of the threat. The last scenario we simulate focuses on the perception of the threat within Germany. Historians frequently describe that Jewish residents of Germany underestimated the danger that the Nazi regime posed to them. If people had had a more accurate estimate of the threat early on, would outmigration have increased dramatically? Conceptually there are several ways how an improvement in the perception of the threat can be operationalized, all of which having issues. First, an accurate perception of the threat could imply that all individuals have perfect knowledge about the state of the danger in each year. Alternatively, a more accurate estimate of the danger could indicate that individuals have perfect foresight of the events that will happen in the future.

In this counterfactual simulation we follow the second approach and simulate outmigration in a world in which people would have already known in 1936 (after the Nuremberg Laws) the rates of persecution after Kristallnacht. That is, we assign the 95th percentile value of cumulative detainment (38%) in 1939, i.e., one year after Kristallnacht, to people in 1936. As displayed in Table 10, knowledge of post-Kristallnacht detainment would have increased outmigration substantially: Compared to the status quo, about 5 thousand additional migrants would have left Germany until 1941. However, in comparison with policies lowering migration frictions or increasing destinations' attractiveness, the impact on outmigration of such an early assessment of the threat and future events would have been systematically smaller.

Sensitivity Analysis. The last two columns of Table 10 summarize the results of our counterfactual simulations obtained with alternative parameter values, i.e. estimates from the non-instrumented version of the upper model (Table 7, Panel A, column 1). The main difference is the value taken by β, the parameter driving the exodus effect, which is much larger in the instrumented version. Not surprisingly, the alternative set of results therefore exhibit smaller consequences in all five policy scenarios. The multiplier effect of the initial change is mostly driven by β, which measures the feedback of outmigration decisions in one's network. The total effects of the low β version are 2 to 3 times smaller and constitute a natural lower bound of our counterfactual results. Those remain however sizable, ranging from 2 to 6 thousand more Jewish migrants fleeing Germany during the period of the Nazi regime. A reinforcing channel is that the elasticity with respect to outmigration incentives ($1/\lambda_1$) is smaller in the non-instrumented set of results, dampening the response to any change in the determinants, including the ones that do not involve network effects directly (as is the case for scenario 5).

8 Conclusion

In this paper, we study the importance of social learning 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 networks. Our results suggest that in situations of conflict when emigration becomes massive and rapid the exodus effect is crucial.

Our paper also contributes to developing quantitative tools for evaluating the impact of asylum policies on the volume and composition of refugee flows and on the counterfactual number of casualties that could have been avoided. 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. Current episodes of mass emigration from conflict areas, such as the flight of Syrians during the civil war, seem to share several characteristics with the situation of Jewish refugees pre-WWII: A certain lack of solidarity of destination countries, backed by public opinions that share strong anti-migrant sentiments, and a failure of the international community to coordinate on the acceptance of refugees. For academics and policy makers alike, it is important to understand how migration decisions are made in situations of conflict. Our paper also emphasizes that social spillovers can create multiplicative effects which should be considered when designing policies.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

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

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2015). Radio and the Rise of the Nazis in Prewar Germany. *The Quarterly Journal of Economics 130*(4), 1885–1939.

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

Anderson, R. W., N. D. Johnson, and M. Koyama (2016). Jewish Persecutions and Weather Shocks: 1100-1800. *The Economic Journal*, 1–35.

Anderson, S., A. De Palma, and J. Thisse (1992). *Discrete choice theory of product differentiation*. MIT Press.

Andrews, M. J., L. Gill, T. Schank, and R. Upward (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.

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

Becker, S. and L. Pascali (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 A. Ferrara (2019). Consequences of forced migration: A survey of recent findings. *Labour Economics 59*, 1–16.

Beine, M., S. Bertoli, and J. Fernández-Huertas Moraga (2016). A practitioners' guide to gravity models of international migration. *The World Economy 39*(4), 496–512.

Beine, M., F. Docquier, and Ç. Özden (2011). Diasporas. *Journal of Development Economics 95*(1), 30–41.

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

Bertoli, S. and J. F.-H. Moraga (2015). The size of the cliff at the border. *Regional Science and Urban Economics 51*, 1–6.

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

Bohra-Mishra, P. and D. S. Massey (2011, May). Individual Decisions to Migrate During Civil Conflict. *Demography 48*(2), 401–424.

Bramoullé, Y., H. Djebbari, and B. Fortin (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 M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy 127*(5), 2229–2268.

Caliendo, L., F. Parro, L. D. Opromolla, and A. Sforza (2020). Goods and factor market integration: A quantitative assessment of the eu enlargement. Mimeo.

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

Card, D., J. Heining, and P. Kline (2013, 05). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Cattaneo, C. and G. Peri (2016). The migration response to increasing temperatures. *Journal of Development Economics 122*, 127 – 146.

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

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

Costinot, A. and A. Rodriguez-Clare (2014). Trade Theory with Numbers: Quantifying the Consequences of Globalization. In E. Helpman, G. Gopinath, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4, Chapter 4, pp. 197–261. Elsevier.

Dekle, R., J. Eaton, and S. Kortum (2007). Unbalanced Trade. *American Economic Review 97*(2), 351–355.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Fajgelbaum, P. D., E. Morales, J. C. Suárez Serrato, and O. Zidar (2018, 09). 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.

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

Giulietti, C., J. Wahba, and Y. Zenou (2018). Strong versus weak ties in migration. *European Economic Review 104*, 111–137.

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

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

Gruner, W. and C. Pearce (2019). *German Reich 1933–1937*. De Gruyter.

Guimaraes, P., O. Figueirdo, and D. Woodward (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.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In E. Helpman, G. Gopinath, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4, pp. 131–195. Elsevier.

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

Heusler, A. and A. Sinn (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., V. Lindenthal, and F. Waldinger (2018). Discrimination, Managers, and Firm Performance: Evidence from ’Aryanizations’ in Nazi Germany.

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

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

Jünger, D. (2016). *Jahre der Ungewissheit: Emigrationspläne deutscher Juden 1933–1938*, Volume 24. Vandenhoeck & Ruprecht.

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.

Klemperer, V. (2013). *I Shall Bear Witness: The Diaries Of Victor Klemperer 1933-41*. Orion.

Kruse, A. and E. Schmitt (2013). *Wir haben uns als Deutsche gefühlt: Lebensrückblick und Lebenssituation jüdischer Emigranten und Lagerhäftlinge*. Springer-Verlag.

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

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

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

Maurer, T., M. A. Kaplan, S. M. Lowenstein, and R. Liberles (2005). *Jewish Daily Life in Germany, 1618-1945*. Oxford University Press on Demand.

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

Missirian, A. and W. Schlenker (2017). Asylum applications respond to temperature fluctuations. *Science 358*(6370), 1610–1614.

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

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

Mulligan, C. B. and Y. Rubinstein (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.

Nicosia, F. R. and D. Scrase (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 G. Peri (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., V. Rappoport, P. Schnabl, and D. Wolfenzon (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, September). Trade and geography. Working Paper 27821, National Bureau of Economic Research.

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

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

Roberg, K. W. (2009). *A Visa or Your Life!* AuthorHouse.

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

Small, K. A. and H. S. Rosen (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*, 415–438.

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.

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

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

Totten, S. and S. Feinberg (2016). *Essentials of Holocaust education: Fundamental issues and approaches*. Routledge.

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

UNHCR (2020). *Global Trends: Forced Displacement in 2019*. United Nations High Commissioner for Refugees.

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

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.

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

Windolf, P. (2011). The German-Jewish Economic Elite (1900 to 1930). *Zeitschrift für Unternehmensgeschichte 56*(2), 135–162.

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.

Online Appendix for

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

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

December 7, 2020

This Online Appendix accompanies the paper *The Refugee's Dilemma: Evidence from Jewish Migration out of Nazi Germany*. Section A describes the data. Section B provides further details on the historical background. Section C contains additional empirical results. Section D exposes the computational details of the Exact Hat Algebra and displays additional material related to the counterfactual simulations.

Contents

*Department of Economics, University of Lausanne.
†Department of Economics, Sciences Po, Paris; Banque de France; CEPII; and CEPR.
‡King's Business School, King's College London.
§Department of Economics, University of Lausanne; and CEPR.

A Data and Variable Description

This section describes the data sources and the variables used in the empirical analysis.

A.1 Individual Data on Emigration and Deportation

To measure migration, deportation, and detainment, we draw on "the list of Jewish residents in the German Reich 1933–1945" (hereafter, the "Resident List") that was compiled by the German Federal Archives. Please refer to Zimmermann (2013a) for additional information. In 2004, the Federal Government of Germany commissioned the Federal Archives to construct a scientifically sound and complete list of an estimated 600,000 Jewish residents who had lived in Germany between 1933 and 1945.

To this end, in a more than a decade-long data collection effort, the Archives carefully reconstructed information on Jews living in Germany during 1933 and 1945, defined by its border of 31 December 1937. In this process, the Federal Archives consulted more than 1,000 different sources and over 2.5 million references. To reconstruct deportations, the Archives drew mainly on the deportation lists created by the Nazi perpetrators, registers of concentration camps and other Gestapo documents. Information on migration movements comes from several main sources: (i) the list of individuals of the Reich Association of Jews in Germany; (ii) membership lists of Jewish parishes; (iii) records including municipal resident files, civil registries, chronicles, local histories and other publications from about 1,000 German municipal archives that were contacted by the Federal Archives; (iv) foreign archives, Jewish communities, and institutions that possessed information about Jewish emigrants from Germany (in particular emigrants to Brazil, Sweden, South Africa and Switzerland). As of 2013, the data is still being updated with new information on a constant basis (Zimmermann, 2013a). We use the version that was provided to us by the Federal Archives on March 25, 2019.

Regarding the content of the "Resident List", it provides biographic information, including the first and last name, birth date, birthplace, gender, and place of residence, of about 484 thousand individuals. It is important to note that the original "Resident List" contains several thousand additional observations; the 484 thousand individuals are those for which information on the date of birth, birthplace, and city of residence is available, which is crucial information for the purpose of our study. The data on first and last names is available only for a sub-set of individuals that were either born before 1919, or dead for at least 10 years, as the German Archival law lifts the privacy protection of these persons.

The "Resident List" records the migration and/or deportation history of about 285 thousand German Jews, including the timing of migration movements, the first and possible second destination countries, and/or the deportation date and place. Moreover, it includes information on whether and when a person was detained (*"Schutzhaft"*).

In the following, we describe the main variables used in the individual-level analysis and in the gravity specifications.

i. Main variables of the individual-level analysis:

- **Age:** Measures the age of individual i in year t.
- **Male:** A dummy variable taking on the value 1 for individuals that are male, and 0 for individuals that are female.
- **Born outside Germany:** A dummy variable taking on the value 1 for individuals that were born outside Germany as defined by its borders of December 31st, 1937.
- **Year of emigration / Year of deportation:** Measure the year of the first known emigration / deportation.
- **Emigration destination:** Indicates the first destination of emigration.
- **Ever-detained:** A dummy variable taking on the value 1 for individuals that were detained before or in year t.
- **Detainment in city of residence:** Measures the cumulative share of co-residents in the city of residence that have been detained until year t.
- **Detainment of network members:** Measures the cumulative share of network members that have been detained until year t.
- **Migration of network members:** Measures the cumulative share of network members that emigrated until year $t-1$.
- **Network Push:** For an individual i, it is defined as the average past detainment share in the cities of residence $r(j)$ of her distant peers j up to year $t-1$:

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

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

- **Network Pull:** Defined as the average partial inclusive utility across distant peers up to year $t-1$. 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 in Germany up to year $t-1$:

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

- **Jewish Name Index (JNI):**

  The Jewish Name Index (JNI) measures the Jewishness of the first name of an individual and ranges from 0 to 100. It 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.

To analyze the Jewishness of first names, we follow several papers in the literature, in particular (Fryer Jr and Levitt, 2004; Fouka, 2019). We cannot rely on our sample of German Jews to assess the frequency of names among Jews versus non-Jews, as we do not observe the distribution of first names among ethnic Germans. We therefore draw on the U.S. censuses of 1910, 1920, and 1930, as it allows to identify the first name *and* the language spoken by a person, and thus to get an estimate of an individual's identity. The census micro-data is provided by the Integrated Public Use Microdata Series (IPUMS) project of the Population Center at the University of Minnesota (Ruggles et al., 2019).

We identify all Yiddish speakers (who are Jewish), and German speakers in these samples. For each first name n, we compute its Jewishness by dividing the share of name n among Yiddish speakers by the sum of share of name n among Jewish speakers and that among German speakers. Formally:

$$\text{JNI}_{\text{US Census}} = \frac{\text{Share of name } n \text{ among Jiddish speakers}}{\text{Share of name } n \text{ among Jiddish speakers + Share of name } n \text{ among Germans}} \times 100$$

The JNI therefore shows how disproportionately a name is given by Yiddish speakers compared to German speakers. The index ranges between 0 and 100. If a name is given only by Yiddish (German) speakers, the index will take the value of 100 (0). We validate in Table C6 of Section C.3 that the Jewishness of a first name as defined by our index correlates strongly with Jewish ancestry based on a sample of German Jews observed in the supplementary cards of the German Census of 1939 (Zimmermann, 2013b).

ii. Main variables of the gravity estimation:

- **Destination d:** A destination d is an emigration country that belongs in at least one year t between 1933-1941 to the set of countries that make up 95% of total migration in a given year. The set of destinations contains 35 such countries, which we restrict further to 32 countries that belong to the largest connected set. A destination is defined as the first emigration destination after leaving Germany.

- **$\textbf{ShareMigrants}_{\textbf{rbdt}}$** : is defined as the share of migrants from a combination of residence city r and birth city b to destination country d in year t.

- **$\textbf{Diaspora}_{\textbf{rdt}}$** : is defined as the cumulative share of migrants from a residence city r to destination country d up to year t (not included). It is computed by dividing the stock of migrants from residence city r in destination country d in year $t-1$ by the initial Jewish population in the residence city in 1933.

$$\texttt{Diasp}_{rdt} \equiv \frac{\text{stock}_{rd,t-1}}{\text{pop}_{r,t_0}} \tag{1}$$

- **Diaspora$_{\mathbf{bdt}}$** : is defined as the cumulative share of migrants from a birth city b to destination country d up to year t. It is computed by dividing the stock of migrants from birth city b in destination country d in year $t-1$ by the initial Jewish population in the birth city in 1933.

$$\texttt{Diasp}_{bdt} \equiv \frac{\text{stock}_{bd,t-1}}{\text{pop}_{b,t_0}} \tag{2}$$

- **Distance$_{\mathbf{rd}}$** : is defined as the distance from residence city r to the closest point along the border of destination d defined as of February 28th, 1938. The shapefiles that represent country boundaries come from Groot (2010). Own calculation using ArcGIS®.

- **Distance$_{\mathbf{bd}}$** : is defined as the distance from a birth city b to the closest point along the border of destination d defined as of February 28th, 1938. The shapefiles that represent country boundaries come from Groot (2010). Own calculation using ArcGIS®.

A.2 Individual Data on Jewish Ancestry

- In Section C.3 we use individual level information on the number of **Jewish ancestors** of individuals, as well as on the completion of higher education. This data come from the supplementary cards of the 1939 German Census.

A.3 City-Level Data

- City-level information on **antisemitic attitudes** (Medieval pogroms, NSDAP votes in 1928, and synagogue attacks in 1938) come from Voigtländer and Voth (2012).

- **Socio-economic data** at the level of the city to measure population, the share of Jewish population, the share of Protestants, as well as the economic structure of towns is based on the German censuses of 1925 and 1933 and provided by Hänisch (1989).

- The strength of the **radio signal** in 1933 is taken from Adena et al. (2015).

A.4 Destination-Level Data

- **Income per capita** in destination d and year t is taken from the Maddison Project Database 2018 (Inklaar et al., 2018). Income per capita measures real GDP per capita in 2011 US$ using multiple benchmarks (`cgdppc`).

- The **distance** between Germany and emigration destinations is taken from the CEPII GeoDist database (Mayer and Zignago, 2011). We use the variable `dist` which measures the simple distance based on the most populated cities, in km.

- **German occupation** takes on the value 1 in years in which a destination is occupied by Nazi Germany. This information is taken from Stiftung Jüdisches Museum Berlin (2006).

- For the simulation of work restrictions in Section D.2, we coded the extent of **labor restrictions** in destination countries using historical information from a multitude of sources, in particular country-wide and comparative studies of emigration in the 1930s (e.g., Lohfeld and Hochstadt, 2006; Meyer, 1998; Jackisch, 1994; Strauss, 1971; Stiftung Jüdisches Museum Berlin, 2006; Wegner, 2013).

  Labor restrictions measure the extent of work prohibitions on an ordinal scale ranging from 1 to 4:

  - 1 classifies destinations with "No restrictions";
  - 2 classifies destinations where "Work allowed with permit";
  - 3 classifies destinations where "Work allowed, but permit difficult to obtain";
  - 4 classifies destinations where "No access to the labor market".

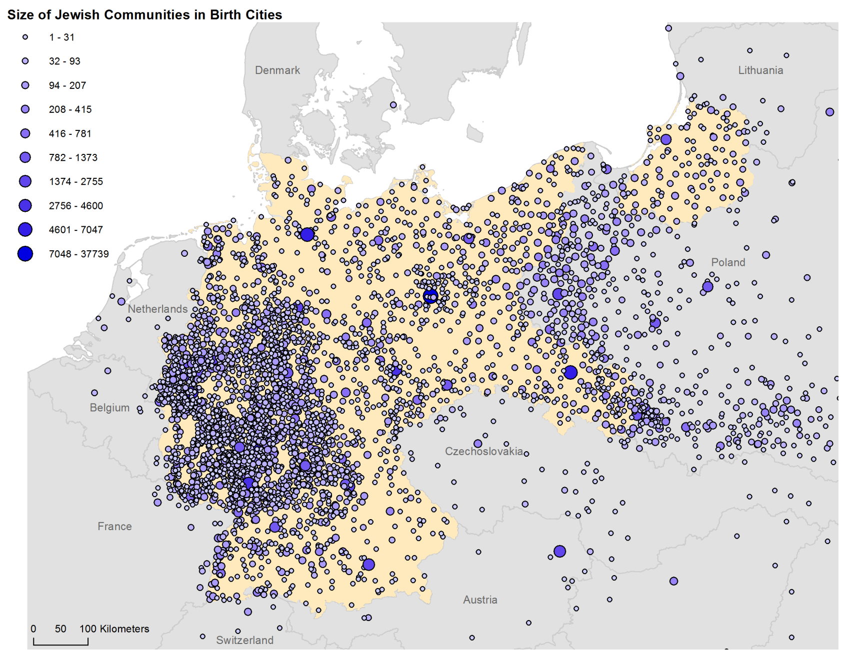
- In Section D.2, we use information on the **prices of maritime travel** to calculate the cost of travel per km and to simulate the consequences of a travel subsidy. To this end, we collected the ticket prices (in Reichsmark) of boat trips between European harbors and harbors overseas for a total of 42 trips. The ticket prices, as well as the travel time in days, are taken from the Philo-Atlas, a handbook for Jewish emigrants (Löwenthal and Oppenheimer, 1938).

A.5 Spatial Distribution and Size of Jewish Communities in 1933

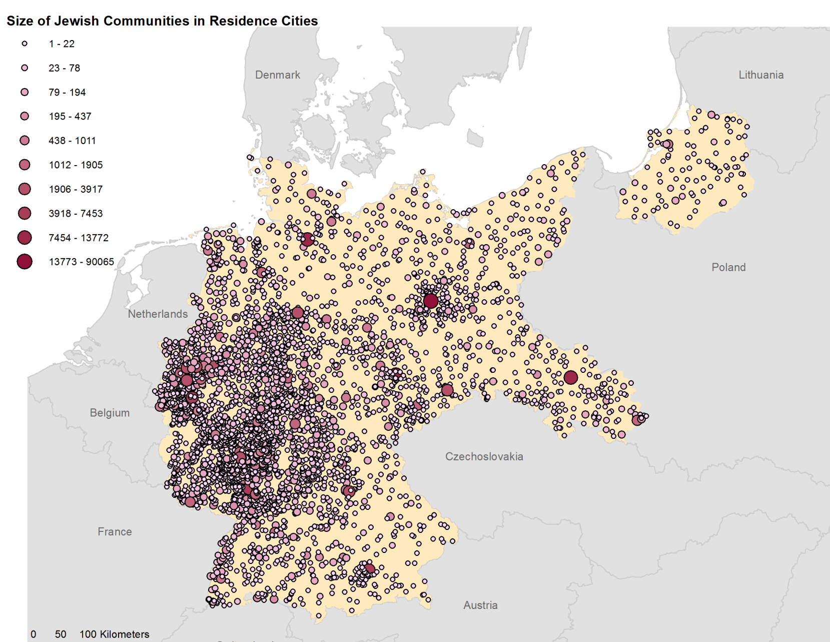
Figure A1 displays the size of Jewish communities across cities of birth (upper figure), and across cities of residence (lower figure), based on individual-level data from the “Resident List".

Figure A1: Size of Jewish Communities in 1933

(a) Birth Cities



(b) Residence Cities



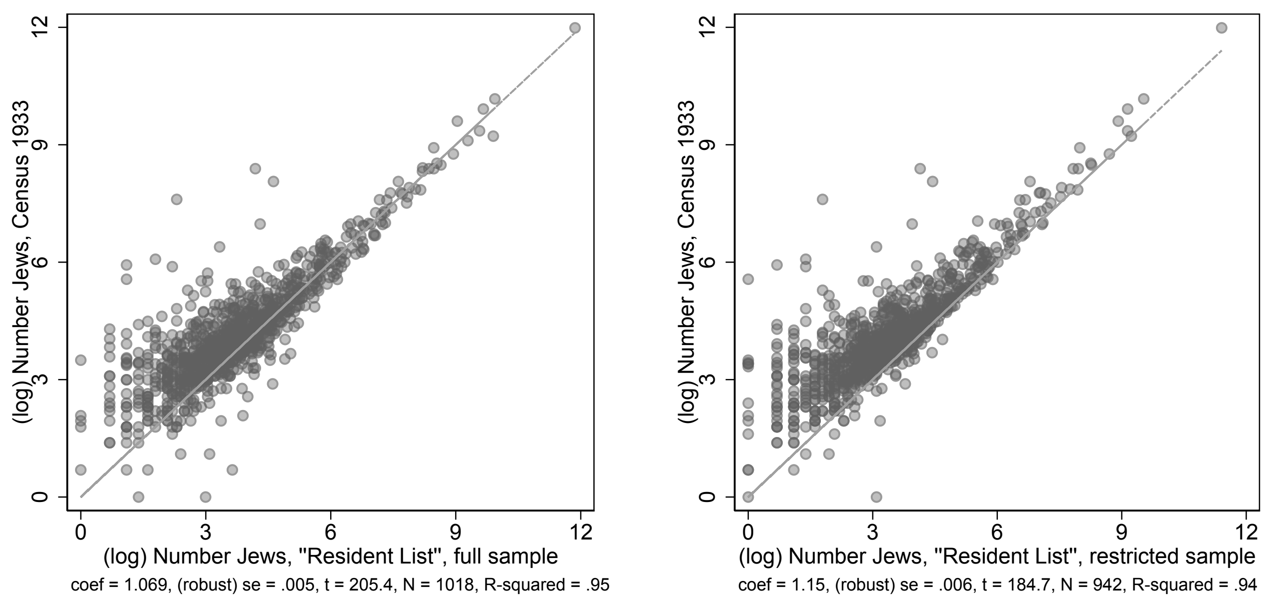
NOTE: Boundaries of Germany as of February 28th, 1938. Author’s own calculation.

A.6 Limitations of the Data and Network Construction.

Since our data set on migrants and deportees was collected retrospectively, it comes with limitations that are discussed in the following:

Missing information. First, the dataset misses information on migration and deportation for several thousand individuals. Information on deportation and deportation exists for 59% (or 285 thousand individuals) out of 484 thousand individuals in the "Resident List". For the remaining, we cannot tell whether they migrated or were deported (because the information is not traceable), or whether they stayed in German and survived (because they were hidden or because they were married to a non-Jewish spouse which protected them). Estimates suggest that at most 15,000 individuals were hiding in the underground. Of the 5,000 individuals hiding in Berlin (3% of the Jewish population of Berlin in 1933), an estimated 1,500 survived (Schreiber, 2005).

Figure A2: Correlation of Jewish City Populations across Different Sources



(a) Full sample, N = 484k

(b) Sample with mig./dep. information, N = 285k

NOTE: The gray line represents the 45-degree line.

We validate the distribution of Jews across cities in our sample by comparing it to city-level data from the official Census of 1933. This is a first check on sample selection and shows that our sample represents very well the distribution of Jewish communities across cities. Figure A2 plots the log number of Jews in the 1933 Census against the log number of Jews in the "Resident List": for the full sample of 484k individuals (Figure (a)) and using the sample of 285k individuals, for which migration and deportation information is available (Figure (b)). Variation in (log) number of Jewish inhabitants per city in the "Residence List" in the full sample explains 95% of the variation in the (log) number of Jews per city as reported in the Census with a point estimate of 1.069 (Figure (a)). The correlation is especially strong for larger cities. Deviation from the 45-degree line is more pronounced for smaller cities. Therefore, we present the robustness of our estimates excluding smaller cities from the sample. We observe a similar correlation pattern when looking at the restricted sample (Figure (b)).

Sample selection due to missing information could pose a threat to our identification strategy. To address the concern that sample selection might bias our findings, we compute for each individual the share of her distant peers with missing migration and deportation information and check the robustness of our findings to excluding decision makers with a high share of distant peers with missing information. This robustness exercise is presented in Table C15 of Section C.11.

Movers. Moreover, the focus on distant peers born in the same city and of similar age comes with potential pitfalls that require a careful investigation. Distant peers are a selected sample as they are internal movers, i.e., individuals who moved from their city of birth to another city. In Table A1, we compare individual characteristics of movers and non-movers and find that the former are different along several dimensions: they are older and more likely to be female, which makes them less likely to migrate out.

Table A1: Individual Characteristics of Internal Movers

	Mover	Non-Mover	Diff.	Std. Error	Obs.
Birth Year	1890.39	1904.79	14.40***	0.07	285,143
Male	0.46	0.50	0.04***	0.00	283,283
Born outside Germany	0.30	0.01	-0.29***	0.00	262,880
Migrated	0.46	0.57	0.11***	0.00	285,143
Deported	0.60	0.54	-0.06***	0.00	285,143
Year of Emigration	1937.44	1937.35	-0.09***	0.02	99,247

NOTE: Sample composed of 285,143 individuals for which information on migration/deportation is available.

However, when we compare the characteristics of cities in which movers and non-movers live, we find that they are largely similar, see Table A2. For example, they are economically similar, experienced comparable levels of antisemitism before 1933, and are as distant to the German border. The only statistically significant difference between movers and non-movers is that movers originate from smaller cities of birth. This is reassuring, as it indicates that movers did not select into cities with specific push and pull factors that we exploit for identification.

Spatial spillovers. Furthermore, there are potential spatial spillovers if distant peers are living in proximity to decision makers. To address this concern, we restrict our analysis to internal movers as decision makers and impose minimal distance between the city of birth and residence of decision makers in the robustness work of Section C.11. Restricting the network to cohorts of similar

Table A2: City Characteristics of Internal Movers

Panel A, Demographics:	(1) (log) Pop 1933 CoR	(2) (log) Pop 1933 CoB	(3) Share Jews 1933 CoR	(4) Share Jews 1933 CoB
Mover = 1	0.220	-2.448***	-0.136	-0.003
	(0.860)	(0.659)	(0.381)	(0.003)
Observations	237,231	174,750	234,703	167,452
R-squared	0.00	0.19	0.00	0.00
Mean of dependent var.	13.024	11.857	2.619	0.026
SD of dependent var.	2.330	2.793	2.015	0.027

Panel B, Economics:	(1) % Unemployed 1933 CoR	(2) % Blue Collar 1933 CoR	(3) % Self Empl. 1933 CoR	(4) % Retail/Trade 1933 CoR
Mover = 1	0.546	0.463	-0.379	0.734
	(3.051)	(2.093)	(0.676)	(2.535)
Observations	237,687	237,687	237,687	237,687
R-squared	0.00	0.00	0.00	0.00
Mean of dependent var.	28.543	49.408	19.513	34.587
SD of dependent var.	8.613	7.984	3.283	9.619

Panel C, Other:	(1) Medieval Pogrom CoR	(2) Vote share NSDAP 1928 CoR	(3) (log) Distance German Border CoR	(4) Radio Signal 1933 CoR
Mover = 1	0.049	0.001	-0.002	0.972
	(0.136)	(0.004)	(0.191)	(1.257)
Observations	230,224	237,687	261,719	237,687
R-squared	0.00	0.00	0.00	0.00
Mean of dependent var.	0.367	0.027	4.376	32.954
SD of dependent var.	0.482	0.027	0.989	10.670

NOTE: Standard errors clustered at the city of birth in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Sample composed of 285,143 individuals for which information on migration/deportation is available.

age limit measurement errors given that Jewish communities were on average relatively small and concentrated. However, we perform a number of sensitivity analysis on the size of Jewish community in birth cities. Specifically, we exclude cities with large Jewish communities (e.g. Berlin), or small cities.

B Additional Historical Background

B.1 Jewish Life in Germany

While antisemitic sentiments loomed since the Medieval pogroms in Germany and turned into episodes of violence sporadically, in early 20th century, Jews enjoyed equal rights and were economically and culturally integrated. When Hitler took power in 1933, about 520 thousand Jews were living in Germany, of which 425 thousand were of German origin, and about 100 thousand were of foreign nationality. 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, in particular Berlin ($>$ 150 thousand people), Frankfurt/Main ($>$ 20 thousand people), and Hamburg ($>$ 20 thousand people).[1] Only one-fifth of the Jewish population resided in small villages. The Jewish community was spatially concentrated even in large towns: In Berlin, for example, it concentrated largely in only six out of 20 districts (Maurer et al., 2005).

Historical restrictions on occupational choice made Jews to work in commerce and trade, finance, civil service and other service occupations, such as in law, medical professions, in religious functions or for communal service (Maurer et al., 2005). Most Jews of German origin were members of the middle class, they valued education and were over-represented in higher education institutions (Maurer et al., 2005; Windolf, 2011). In contrast, Jews originating from Eastern Europe were on average economically worse off and belonged to the working class (Kaplan, 1999). Estimates of the wealth of Germany's Jews suggest that they owned a share of wealth slightly larger than their population share, i.e., Jews were richer than average but not extremely so (Ritschl, 2019).[2]

In Germany, social divisions between religious groups were traditionally pronounced. As a result, social relationships within the Jewish community (*"Gemeinde"*) were the cornerstone of Jewish life.[3] Jews interacted mainly with their extended family, Jewish friends, and community members. The *"Gemeinden"* 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, ranging from youth and student movements, Zionist groups, athletic, and music clubs. While religiosity declined in early 20th century, the synagogue remained a crucial center of social interactions.[4] Loyalty towards the Jewish community was high, and even Jews that were not practicing their faith participated in community meetings and Jewish organizations, and donated to Jewish charity. Besides the community, socialization took place within the extended family, and matchmaking created extensive

[1] About 55% of the entire Jewish population lived in the top 10 cities in 1933 (Maurer et al., 2005). The distribution of people in our data matches this number perfectly.

[2] Nazi ideology and earlier historical studies overstated the extent of wealth owned by the Jewish community in Germany (Ritschl, 2019).

[3] We draw on Maurer et al. (2005) that provide a detailed description of Jewish life in Germany.

[4] Some 23,000 Jews converted in the German Empire from Judaism to Christianity. The motives for conversion were often non-religious, as most converts tried to escape discrimination. Despite not adhering to Judaism anymore, the social relationships of converts with their Jewish friends and family remained intact (Maurer et al., 2005).

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.[5] To sum, socialization within the community and the extended family created strong social ties among Germany's Jews.

B.2 Anecdotal Evidence on Communication of Persecution

When the Nazis took power, most of the Jewish population that strongly identified with the German culture and nation remained in the belief 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 that affected individuals personally, or their friends and family members, created fear and made the danger more apparent.

Anecdotal evidence documents that within the Jewish community, information about incidences of persecution was frequently exchanged, for example during private contacts:

> *"Frau Blumenfeld whispered to me, Dr. Salzburg's second son, a medical student, has been arrested."* (March 1933, Klemperer (2013))

Moreover, the members of the Jewish community communicated using the telephone and letters, so even news from other towns spread:

> *"Grete told us ghastly things about the treatment of Jews in Bad Kuldowa."* (October 1938, Klemperer (2013))

> *"In October Oom received a letter from Aunt Babette telling us that she and all the other Jews in Baiertal, fourteen persons in total, were being deported from their homes."* (Roberg, 2009)

> *"I telephone the Arons. [...] I learned that he and very many others with him had been arrested and taken away; at present we still don't know whether they are in the camp at Weimar or are working on the fortifications in the West as convicts and hostages."* (November 1938, Klemperer (2013))

These personal and indirect experiences with antisemitic events provided important information about the extent of persecution.

B.3 Anecdotal Evidence on Communication of Migration Plans

Within the Jewish community, the emigration question, migration plans and prospects were heavily discussed in personal visits, telephone and letters. Here we present some of the anecdotal evidence:

[5] While social relationships were located primarily within the Jewish community, many Jews shared the sentiments of patriotism that characterized the German society in the beginning of the twentieth century. About 100,000 Jewish men fought in the German Army at the French or Russian front during World War I, out of which one-third received bravery medal (Totten and Feinberg, 2016).

"My mother and her five siblings had always kept in touch with weekly letters. Every one of the five sisters now considered the idea of leaving Germany. [...] Uncle Wilhelm was in touch with family members and friends in Switzerland, Luxembourg, America, Palestine and Australia. His letters, handwritten, were constant sources of information." (Roberg, 2009)

From the diaries of Victor Klemperer (2013):

June, 1933: *"An evening visit by Fräulein Walter [...] She took her political economy examination in Leipzig, is a librarian at the State Library, is facing certain dismissal, wants to got to Palestine."*

October, 1933: *"Holldack [....] has approached Dember to ask whether there are any possibilities for him in Constantinople; he no longer feels safe here."*

October, 1935: *"Georg wrote [...] he is going to emigrate. It will cost him three quarters of what he has saved, but after Nürnberg he does not want to live "under the guillotine"."*

November, 1938: *"Frau Aron advised us in the strongest terms to take immediate steps to emigrate and to sell the house; everything here is lost."*

Personal contacts and information sharing across individuals sometimes impacted the migration decisions of individuals directly:

"Her husband refused to leave. [...] The friend told them to flee to the United States where he was heading. Only then did her husband agree to go. [...] They fled only as far as France" (Kaplan, 1999)

Moreover, contacts in foreign countries facilitated the migration process, as social contacts were often necessary to obtain a visa, and as contacts in destination countries were crucial for settling and for obtaining a job and housing (e.g. Roberg, 2009).

B.4 Immigration Policies

Besides factors at the individual level, immigration policies in potential destination countries affected the possibility of outmigration. 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 the allocation of entry visas depending on qualifications, financial means, resident relatives and friends, age or state of health. Already in 1924, the U.S. had fixed a quota that allowed the entry of 27,370 migrants per year from Germany and Austria, independent of religion (Stiftung Jüdisches Museum Berlin, 2006).[6] For visa requests to the U.S., applicants had to make appointments at the U.S. consulate and were often in need for a contact person already living in the U.S. Similarly, Canadian policies imposed a quota and Jews were classified among the "least desirable" groups of immigrants. Immigration to Palestine, regulated by the British

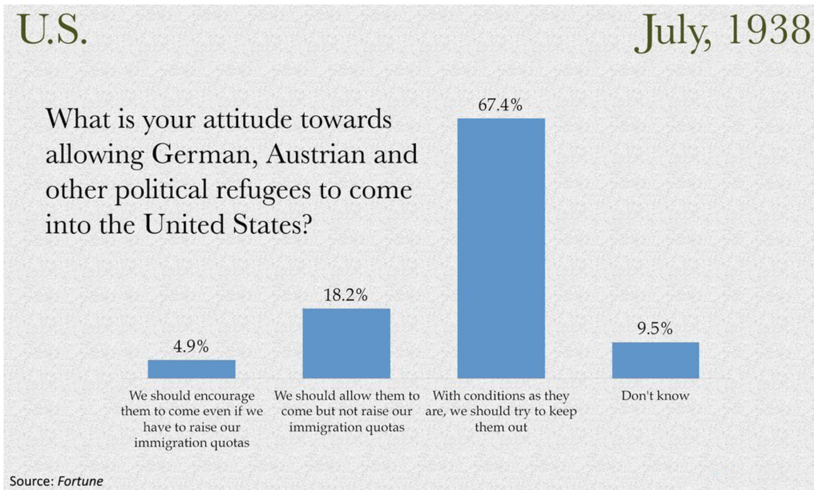
[6]When in May 1939 about 900 Jewish refugees travelled from Germany to the U.S. on the MS St. Louis, they were denied entry as they were not in possession of immigration visas, and therefore had to return to Europe. 250 passengers died later in the concentration camps.

government, required the applicant to have agricultural and domestic skills and to speak Hebrew which was taught only in special Hachschara courses. As the situation for Jews within Germany aggravated during the 1930s, more and more countries closed their doors. From 1938 onwards, Switzerland, for example, demanded from the German government the implementation of the "J-stamp" in Jewish passports, to be able to identify Jewish migrants at the border. Similarly, from 1938 Jews were rejected at the Belgian border and deported to France. Only few countries, such as Shanghai, had very little restrictions and remained open for Jewish refugees until 1941 (Stiftung Jüdisches Museum Berlin, 2006).

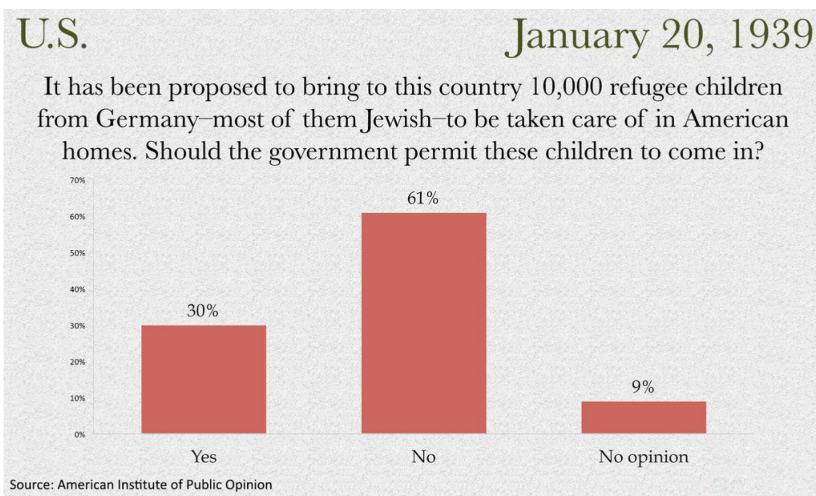
In July 1938, 32 countries met at a conference in Evian, France, to discuss solutions to the Jewish refugee crisis. The conference, however, did not result in an agreement regarding how to allocate the flow of Jewish migrants, as none of the precipitants wanted to commit to accept additional refugees. 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. The delegate of Australia, which took in only 10,000 Jewish migrants during the period 1933—1940, stated: "As we have no real racial problems, we are not desirous of importing one by encouraging any scheme of large-scale foreign migration" (Bartrop and Bartrop, 2018). Polls conducted in 1938 and 1939, shown in Figure B1, illustrate that a majority of Americans was against hosting Jews persecuted in Germany. Even after the November pogroms in 1938, only one-third of the surveyed Americans were in favor of policy proposal that would have allowed 10,000 Jewish children from Europe to come to the U.S. In Canada, a Gallup poll conducted in 1943 ranked Jewish migrants as the second least desirable groups of immigrants (Goldberg, 2019). The Prime Minister of Canada, which accepted a mere 3,500 migrants, believed that "This is no time for Canada to act on humanitarian grounds " (Cymet, 2010). As the world's doors gradually closed and Nazi terror intensified, it became increasingly difficult for Jews to leave Germany.

Figure B1: U.S. Opinion Polls on Refugees in 1938–39

(a) Pre Kristallnacht July 1938



(b) Post Kristallnacht January 1939



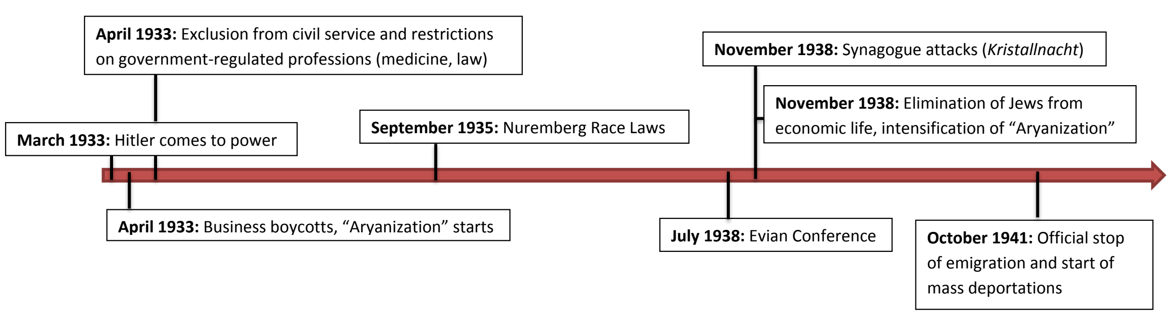
B.5 Timeline of Main Historical Events

Figure B2 contains a timeline of the main events from 1933 to 1941.

- **March 1933:** After Hitler was named Chancellor on January 30, 1933, he seizes complete political power in March 1933.
- **April 1933:** On April 1, 1933, a national boycott against Jewish businesses and other professionals marks the first planned act of Jewish persecution and starts the "Aryanization" of Jewish businesses and assets.

  In the same month, Jews are excluded from civil service positions, and subsequently banned from working in legal, medical, and educational professions ("Law for the Restoration of the Professional Civil Service").
- **July 1933:** The Denaturalization Law allowed to revoke the citizenship of persons naturalized between 1918 and 1933 and targeted in particular Jews originating from the East.
- **September 1935:** The "Nuremberg Laws" fix the definition of Jewishness based on ancestry. Based on this classification, the law excludes Jews from citizenship and forbids the marriage between Jews and Germans.
- **July 1938:** Representatives of 32 countries meet in the French city of Evian to discuss the Jewish refugee crisis.
- **November 1938:** During the Kristallnacht on November 9 hundreds of synagogues are attacked, several thousand businesses destroyed, and thousands of men taken into custody. After the Kristallnacht the "Aryanization" of Jewish businesses and assets intensifies.
- **October 1941:** Emigration from Nazi Germany becomes officially forbidden, and the mass deportations to the concentration camps start.

Figure B2: Timeline of Events



C Additional Empirical Results

C.1 Descriptives

This subsection contains descriptives of the data.

Summary statistics and missings. Table C1 reports summary statistics. Panels A and B summarize the biographic characteristics of Jews in the full sample and for a sample of individuals for whom migration or deportation information is available, respectively. Also, Panel B summarizes the available information on migration, deportation, and detainment history of Jews.

Table C1: Summary Statistics and Missings of the Full Residence List

Variable	N	Missing	Mean	SD	Median
Panel A: Full Sample					
Year of Birth	484,891	0	1896.8	20.7	1896
Male	480,768	4,123	0.49	0.50	0
Born outside Germany	439,145	45,746	0.17	0.38	0
Latitude CoB	439,145	45,746	51.27	1.99	51.34
Longitude CoB	439,145	45,746	12.37	5.89	11.57
Latitude CoR	476,996	7,895	51.44	1.47	51.51
Longitude CoR	476,996	7,895	11.30	3.38	11.54
Emigration and deportation info. available	484,891	0	0.59	0.49	1
Panel B: Sample for which migration or deportation information is available					
Year of Birth	285,143	0	1895.96	19.41	1894
Adult in 1933	285,143	0	0.80	0.40	1
Male	283,283	1,860	0.48	0.50	0
Born outside Germany	262,880	22,263	0.18	0.38	0
Latitude CoB	262,880	22,263	51.21	1.89	51.23
Longitude CoB	262,880	22,263	12.16	5.77	10.9
Latitude CoR	283,589	1,554	51.4	1.42	51.46
Longitude CoR	283,589	1,554	11.04	3.42	10.98
Emigration information available	285,143	0	0.50	0.50	1.00
Year of Emigration	99,247	185,896	1937.4	2.53	1938
Emigration Place information available	138,156	146,987	1	0	1
Year of Deportation	153,293	131,850	1942.01	0.96	1942
Deportation Place information available	164,866	120,277	1	0	1
Detainment Year	17,463	267,680	1939.57	2.32	1938
Year of Death	81,431	203,712	1947.89	12.93	1943

NOTE: This sample does not include individuals that do not have information on their date of birth, nor their birthplace nor their city of residence.

Characteristics of migrants and deported. Table C2 displays mean differences between individuals who migrated between 1933 and 1945, and those who stayed in Germany. Table C3 describes characteristics of individuals that have been deported.

Table C2: Mean Differences Between Migrants and Non-Migrants

	Non-Migrants	Migrants	Diff.	Std. Error	Obs.
Year of Birth	1891.00	1900.89	-9.89***	0.07	285,143
Male	0.40	0.55	-0.14***	0.00	283,283
Deportation	1.00	0.16	0.84***	0.00	285,143
Year of Death	1942.99	1955.81	-12.82***	0.08	81,431

NOTE: Sample comprises 285,143 individuals for which information on migration/deportation is available.

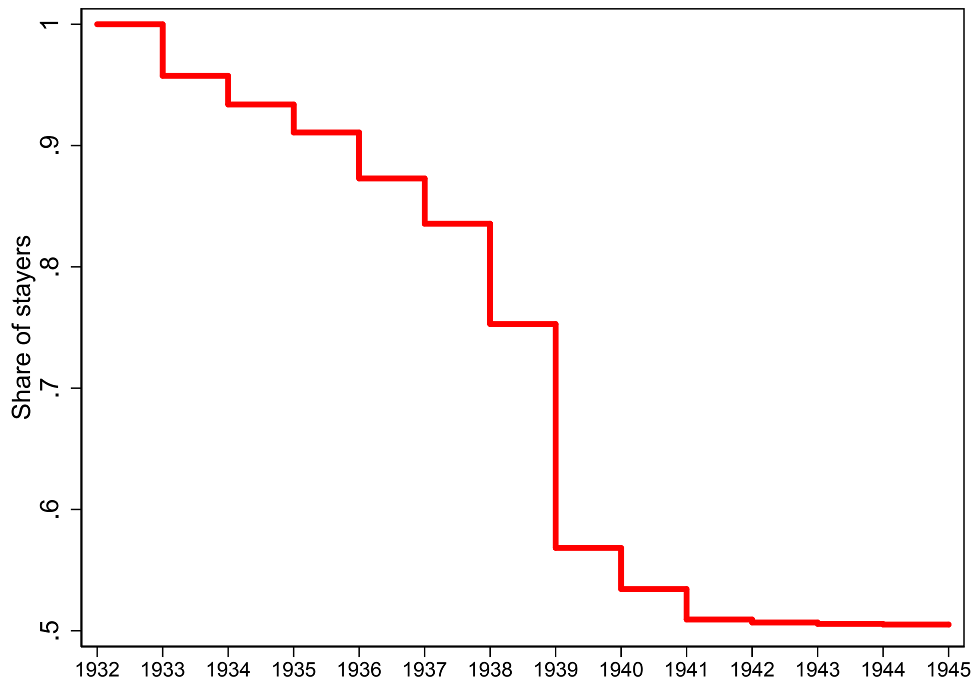
Table C3: Descriptives of Deported Persons

	Mean	Median	SD	Obs.
Year of Deportation	1941.93	1942.00	0.91	132,833
Age at Deportation	51.55	55.00	19.66	132,832
Survived (Death before 1946 or indicated in fate)	0.11	0.00	0.32	70,287
Year of Death	1942.99	1942.00	4.71	50,339

NOTE: Sample composed of 285,143 individuals for which information on migration/deportation is available.

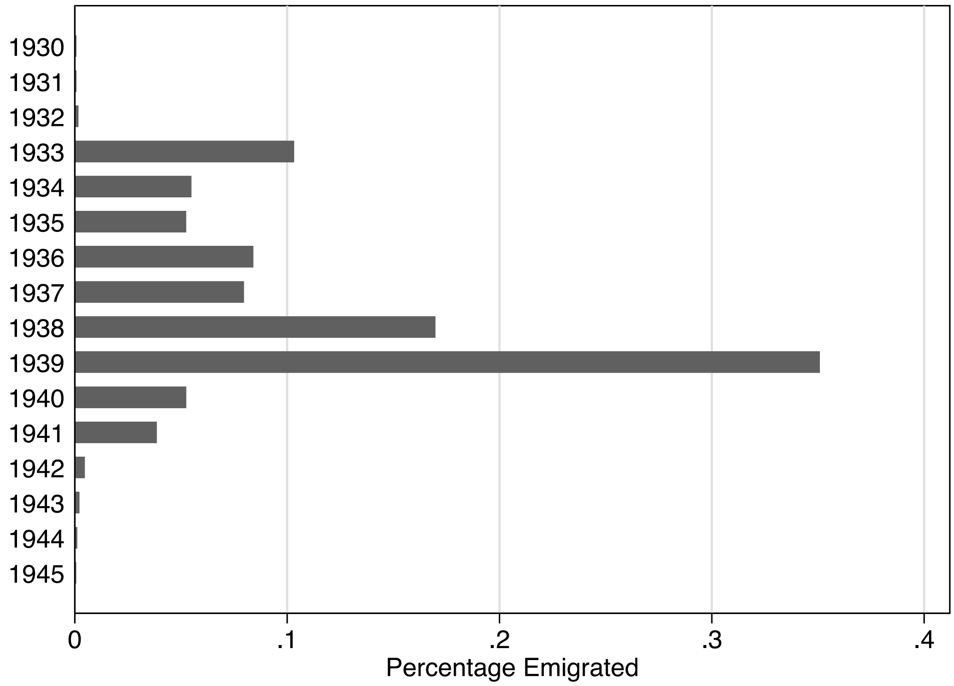
Temporal and spatial patterns of migration. Figure C1 displays the Kaplan-Meier estimator of emigration. It shows that migration intensified during 1938 and 1939, because of the November pogroms of 1938, and afterwards it came to an halt. Figure C2 displays the share of total migration by year. Finally, Figure C3 displays the share of total migration by destination country.

Figure C1: Kaplan - Meier Estimator of Emigration



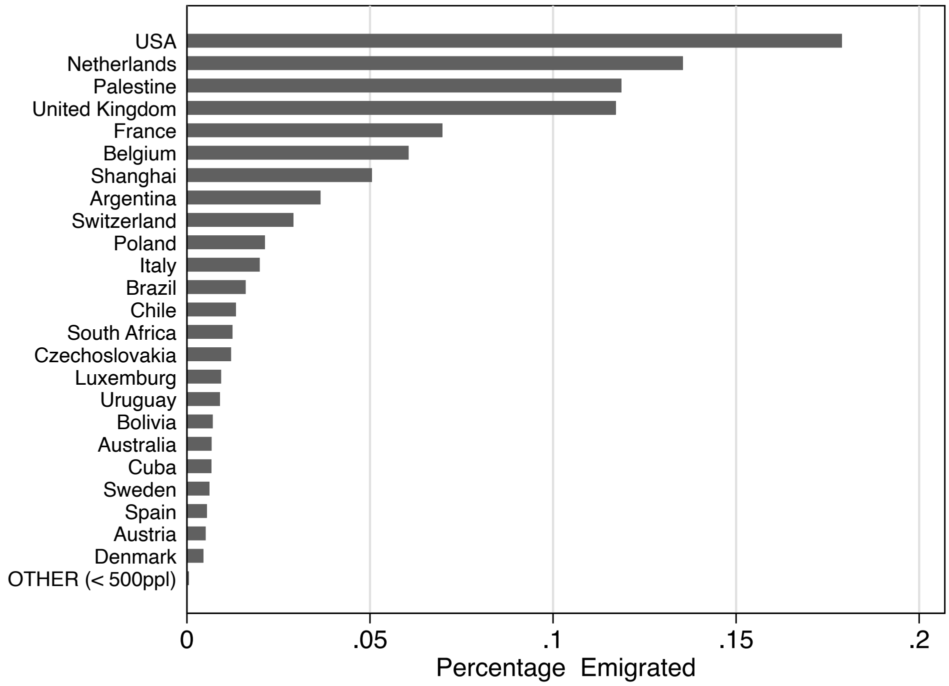
NOTE: Sample composed of 285,143 individuals for which information on migration/deportation is available.

Figure C2: Total Emigration by Year



NOTE: Sample composed of 285,143 individuals for which information on migration/deportation is available.

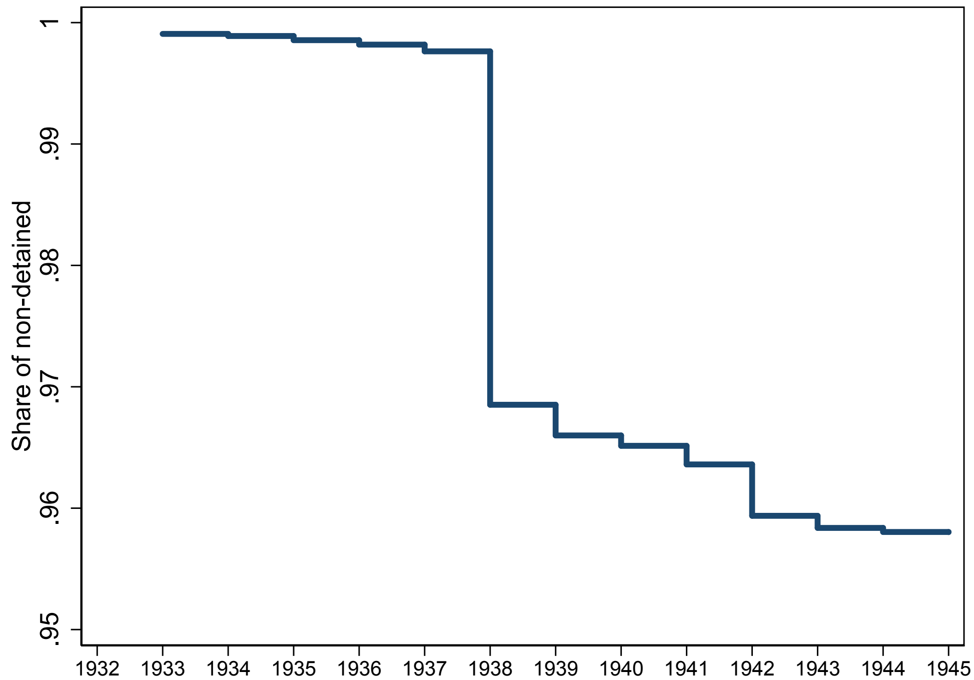
Figure C3: Total Emigration by Destination Countries



NOTE: Sample composed of 285,143 individuals for which information on migration/deportation is available.

Detainment by year. Figure C4 displays the Kaplan-Meier estimator of detainment. It shows that detainment was rare before 1938 and spiked during the November pogroms in 1938.

Figure C4: Kaplan - Meier Estimator of Detainment



NOTE: Sample composed of 285,143 individuals for which information on migration/deportation is available.

C.2 Validating Detainment as Measure of Antisemitism

In Table C4, we validate detainment as a measure of local antisemitism, by regressing incidences of detainment in a city on measures of antisemitism. All regressions control for population, the share of Jews, and the share of Protestants. In column 1, we find that historical antisemitism, measured by the occurrence of a Medieval pogrom in a city, is positively and statistically significantly correlated with the likelihood of detainment before the Kristallnacht. We find a similar result when considering the vote share for the Nazi party in 1928 as predictor of detainment in column 2. In column 3, we only focus on detainment during the year of the Kristallnacht, 1938, and find that cities in which a synagogue was attacked during the November pogroms were more likely to experience at least one incidence of detainment in the same year.

Table C4: Detainment and City-Level Antisemitism

	(1)	(2)	(3)
Dependent variables:	Any Detainment:		
	before 1938	before 1938	in 1938
Medieval Pogrom	0.065*		
	(0.033)		
Vote share NSDAP in 1928		0.913***	
		(0.307)	
Synagogue Attacked in 1938			0.155***
			(0.038)
(log) Population 1933	0.095***	0.099***	0.085***
	(0.010)	(0.009)	(0.010)
Share Jews 1933	0.017***	0.016***	0.026***
	(0.004)	(0.004)	(0.007)
Share Protestant 1925	-0.000	-0.001**	-0.002***
	(0.000)	(0.000)	(0.001)
Observations	862	950	950
R-squared	0.19	0.18	0.13
Mean of dependent var.	0.140	0.136	0.591

NOTE: The dependent variable is an indicator for incidences of detainment before 1938 (columns 1 and 2), and in 1938 (column 3). The unit of observation is a city. Standard errors clustered at the district-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3 Jewish Name Index

Table C5 tabulates the top and bottom 20 first names in terms of Jewish Name Index (JNI) by gender. We focus on a sample of first names that are given to at least 250 individuals in the "Resident List". In Table C6, we validate the JNI as a measure of Jewish identity, by regressing it on the Jewish origin of individuals observed in the 1939 census. The supplementary cards of the 1939 census assessed Jewish origin as defined by the Nuremberg Race Laws by asking about the Jewishness of grandparents both on the father's and mother's side.

For each individual we observe both the first name from which we calculate the Jewish Name Index, as well as the number of Jewish grandparents she has had. In column 1, we regress the JNI on the share of Jewish grandparents, and find a strongly positive and significant correlation. This positive association between Jewish origin and the Jewishness of names persist if we control for a number of individual controls in column 2, and when we add fixed effects for the city of birth in column 3. Instead of using the share of Jewish grandparents, we can also categorize each individual by the number of Jewish grandparents and regress the JNI on dummy variables for each of the four categories of Jewish origin. As column 4 shows, compared to individuals without Jewish grandparents (the omitted category), individuals with one or more Jewish grandparents have a significantly more Jewish sounding name. Moreover, the coefficients are monotonically increasing when moving from individuals with one Jewish grandparent to individuals with four Jewish grandparents. These results strongly support the idea that the Jewishness of a name reflects the strength of the Jewish identity of an individual.

The supplementary cards of the 1939 Census also record if an individual obtained University education or not. We use this information to test whether individuals with a stronger Jewish identity were more or less educated. Column 5 documents a negative association implying that persons with a stronger Jewish identity were less likely to have tertiary education. This can be explained by a higher incidence of traditional Jewish names among Jewish migrants from East Europe. This sub-group of the Jewish community was less integrated and educated than traditional German Jews.

Table C5: Top and bottom 20 first names in terms of JNI by gender

Top 20 names		Bottom 20 names	
Female	Male	Female	Male
Chaja	Eduard	Margarete	Fritz
Cilly	Mendel	Else	Heinrich
Esther	Israel	Hedwig	Heinz
Ester	Abraham	Margot	Friedrich
Bella	Nathan	Luise	Werner
Fanny	Samuel	Karoline	Siegmund
Rachel	Benjamin	Hilde	Franz
Chana	Isaak	Elfride	Gerhard
Sara	Aaron	Franziska	Eugen
Lea	Salomon	Grete	Guenther
Ruth	Isidor	Ernestine	Manfred
Jenny	Harry	Marianne	Helmut
Rose	David	Friederike	Siegbert
Jetta	Moses	Erika	Wolfgang
Lilli	Wolf	Jette	Horst
Jeanette	Elias	Bianka	Curt
Dina	Sally	Liesbeth	Lothar
Vera	Chaim	Thea	Edgar
Ida	Leon	Annemarie	Johannes
Dora	Emanuel	Inge	Egon

NOTE: This table displays the top and bottom 20 names in terms of Jewish Name Index by gender. We focus on a sample first names which are given to at least 250 individuals in the "Resident List".

Table C6: JNI, Ancestry and Socio-Economic Status (1939 census)

	(1)	(2)	(3)	(4)	(5)
Dependent variables:	JNI	JNI	JNI	JNI	Educated
Share of ancestors who were Jewish	17.532***	16.492***	12.933***		
	(0.843)	(0.612)	(0.319)		
Nb. of Jewish ancestors: 1				1.312***	
				(0.288)	
Nb. of Jewish ancestors: 2				3.981***	
				(0.172)	
Nb. of Jewish ancestors: 3				8.380***	
				(0.554)	
Nb. of Jewish ancestors: 4				12.820***	
				(0.337)	
Jewish name index [0,1]					-0.023***
					(0.002)
Individual controls		✓	✓	✓	✓
CoB FE			✓	✓	✓
Observations	377,790	377,603	350,332	353,119	353,119
R-squared	0.07	0.08	0.16	0.16	0.09
Mean of dependent var.	25.495	25.476	25.478	25.413	0.060

NOTE: This table displays the results of regressing the Jewish Name Index (JNI) of an individual on her/his Jewish ancestry as measured by the share of grandparents that are Jewish (columns 1 to 3), or indicators for the number of Jewish grandparents (column 4). Column 5 regresses an indicator for whether the individual has completed higher education on the JNI. All regressions control for the number of ancestors for which information exists. The sample are individuals reported in the 1939 Census. Standard errors clustered at the birth city in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4 Threat Effect

Table C7 documents the results of the threat effect in non-linear specifications (columns 1 and 2), and using LPM in column 3. All regressions control only for year fixed effects, which allows to include a first order local push factor, namely city-of-residence level persecution. City-level detainment shows the expected positive coefficient that is statistically significant. This finding is particularly relevant for our instrumental variable framework (and the statistical power of its first stage) that exploits the link between detainment and migration in peers' city of residence.

Table C7: Network and City-Level Detainment

	(1)	(2)	(3)
Dependent variable:	Migration Decision		
	Logit	Cloglog	LPM
Detainment in city of residence	0.058***	0.058***	0.052**
	(0.013)	(0.012)	(0.021)
Detainment of network members	0.024***	0.024***	0.046***
	(0.007)	(0.007)	(0.011)
Migration of network members	0.016***	0.013***	0.060***
	(0.004)	(0.004)	(0.007)
Age	-0.030***	-0.029***	-0.044***
	(0.003)	(0.003)	(0.003)
Age squared	0.001***	0.001***	0.003***
	(0.000)	(0.000)	(0.000)
Male	0.026***	0.026***	0.026***
	(0.001)	(0.001)	(0.001)
Ever-detained	0.005**	0.005**	0.028***
	(0.002)	(0.002)	(0.005)
Born outside Germany	-0.002*	-0.002*	-0.003**
	(0.001)	(0.001)	(0.001)
Observations	1,025,497	1,025,497	1,025,497
R-squared			0.06
Pseudo R-squared	0.12		
Mean of dependent var.	0.052	0.052	0.052
SD of dependent var.	0.222	0.222	0.222

NOTE: The dependent variable is an indicator for migration. The unit of observation is an individual in year t. The sample consist of individuals who were adults in 1933. Marginal effects reported. Detainment of network members measures the cumulative share of network members that have been detained until year t. All regressions control for year fixed effects. Standard errors clustered at the city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.5 Learning: Network-Peer Characteristics

In Table C8, we investigate how peer characteristics affect the learning process. We start by exploring whether individuals react more to information coming from peers living closer or farther away. In column 1, we interact the detainment among network members with a dummy variable indicating whether the average distance between the residence city of an individual and her peers is greater than the median distance in the sample (275km). The estimated coefficient on this interaction is statistically significant and positive. Moreover, it captures almost the entire effect of the detainment among network members, which suggests that people react to the detainment of their peers only if they live sufficiently far away. This is in line with the argument that private information coming from peers living farther away provide new information about the level of threat in addition to what the individual can observe locally. We find similar results when considering detainment separately for peers that live in distances below the median distance, and above the median distance as documented in column 2. The results suggest that individuals react only to the detainment among their peers who live in greater distance. This differential effect does not seem to arise from a difference in distribution of detainment among peers that are closer and further away. Figure C5 shows that the trends in detainment among network members over time (panel a)), and the distribution at the internal margin, i.e., where detainment rate is above 0.5%, (panel b)) are similar for these two subgroups.

We turn to explore the importance of homophily in behavioral responses to victimization of peers. We consider first names as a social identifier and compute the distance between the first name of an individual and her peers in terms of JNI. We interpret closeness in this measure as a relevant dimension of homophily, as it is likely to represent a similarity across individuals in terms of how secular and integrated their families were to the German society. In column 3, we interact the detainment among network members with an indicator for whether the average distance between the first name of an individual and those of her peers in terms of Jewish Name Index (JNI) is greater than the median distance in the sample (0,29 percentage points). The estimated coefficient on this interaction is statistically significant and negative. To investigate whether people react more to the detainment of their peers with a similar first name as theirs, in column 4, we split individuals' network into peers with a first name above or below the median distance in terms of JNI compared to theirs, and compute the detainment rates for these subsamples. Individuals react only to the 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. Again, these estimated differences in subgroups do not seem to arise from a differential distribution of detainment among peers that are culturally closer and more distant. Figure C6 shows that the trends in detainment among network members over time (panel a)) and the distribution at the internal margin, i.e., where detainment rate is above 0.5%, (panel b)) are similar for these two subgroups. In column 5, we interact the detainment among network members with similar first names with an indicator for whether the average JNI of the first names of network members is above the median in the sample. We estimate a negative and statistically significant coefficient on this interaction term, suggesting that the behavioral response is greater among people with less

Jewish (more German) sounding names, who are more likely to be integrated to the German society. A likely interpretation is that the secular Jews were the most reactive to the victimization of their (secular) peers.

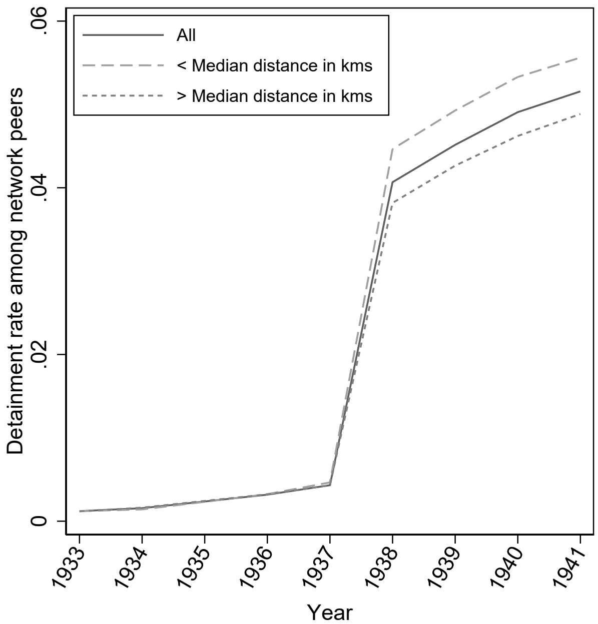
Table C8: Learning: Network-Peer Characteristics

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Migration Decision				
Migration of network members	0.056***	0.071***	0.056***	0.070***	0.069***
	(0.007)	(0.015)	(0.007)	(0.013)	(0.013)
Detainment of network members	0.023**		0.062***		
	(0.011)		(0.014)		
× Distance in kilometers (> *Median*)	0.054**				
	(0.023)				
× Distance in JNI (> *Median*)			-0.041***		
			(0.015)		
Detainment: network members < *Median* distance in kms		0.026			
		(0.016)			
Detainment: network members > *Median* distance in kms		0.059***			
		(0.020)			
Detainment of network members with different JNI				0.013	0.013
				(0.011)	(0.011)
Detainment of network members with similar JNI				0.052***	0.092***
				(0.013)	(0.025)
× Average JNI of network members (> *Median*)					-0.058**
					(0.027)
Observations	1,024,976	645,410	984,413	703,033	703,033
R-squared	0.10	0.10	0.10	0.10	0.10
Mean of dependent var.	0.052	0.056	0.052	0.054	0.054
SD of dependent var.	0.222	0.229	0.222	0.226	0.226

NOTE: LPM estimation with the dependent variable being an indicator for migration. The unit of observation is an individual in year t. The sample consists of individuals who were adults in 1933. Detainment of network members measures the cumulative share of network members that have been detained until year t. All continuous variables in interactions are demeaned. All regressions control for age, age squared, gender, personal detainment, a dummy for whether the individual was born outside Germany, and city-of-residence × year fixed effects. Standard errors clustered at the city-of-birth × year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure C5: Detainment rate among network peers by distance in kilometers

(a) Trends over time



(b) Kernel density distribution

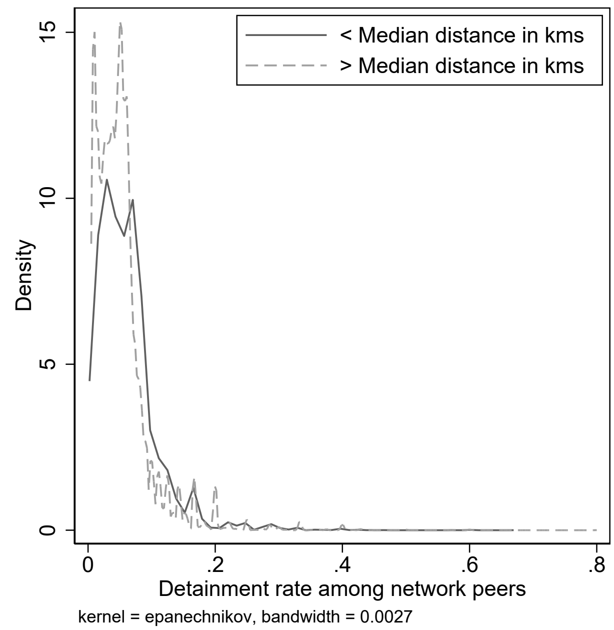
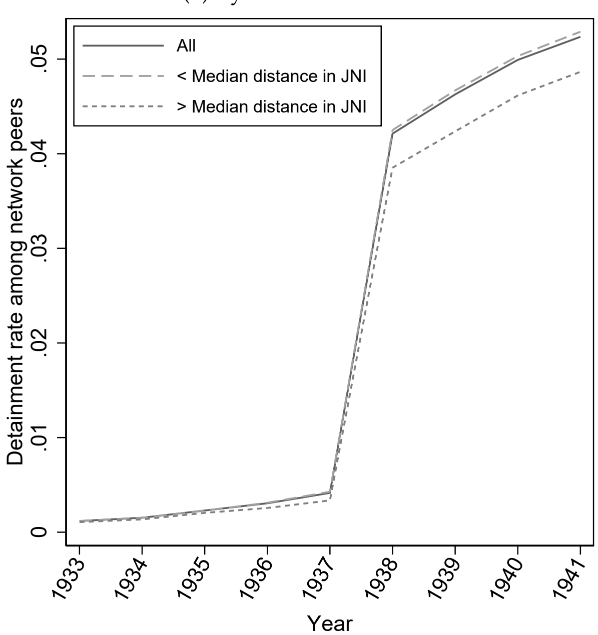
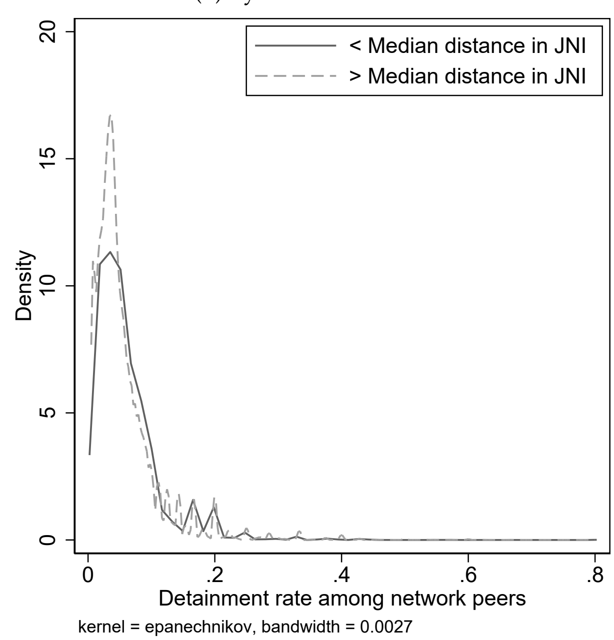


Figure C6: Detainment rate among network peers by distance in Jewish Name Index (JNI)

(a) By distance in kms



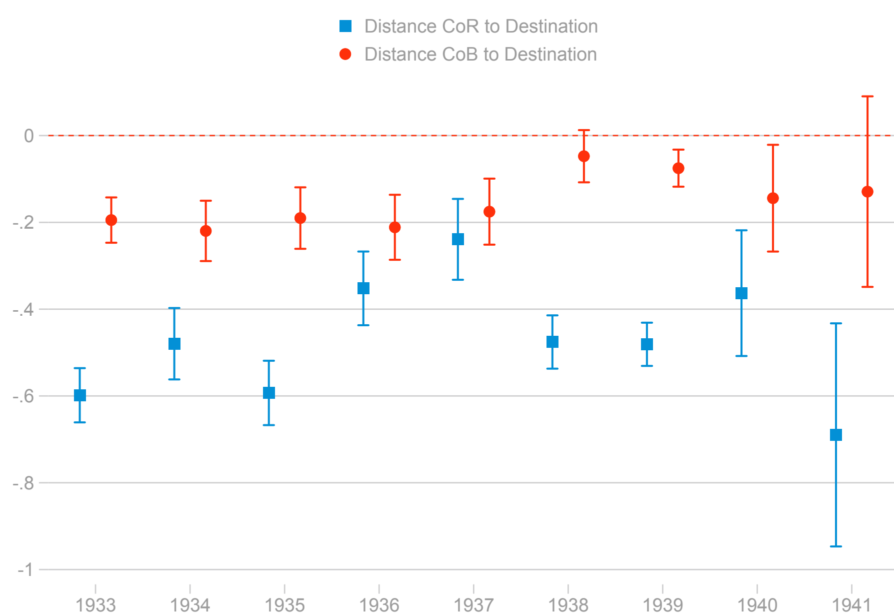
(b) By distance in kms



C.6 Gravity Specification and Country Attractiveness

Distance coefficients by year. Figure C7 displays the time varying effects of the bilateral distance between the cities of residence and destinations (blue) and those between the cities of birth and destination countries (red). It presents the estimated coefficients and the 95% confidence intervals obtained from the gravity specification of column 4, Table 5.

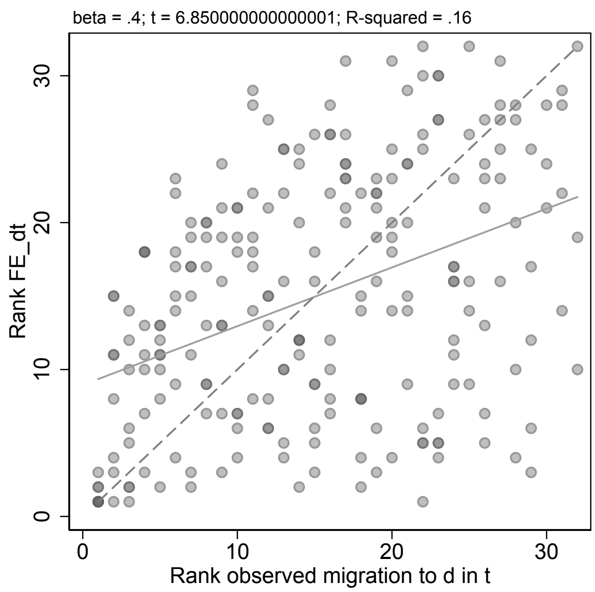
Figure C7: Time-Varying Effects of Distance from Birth and Residence Cities



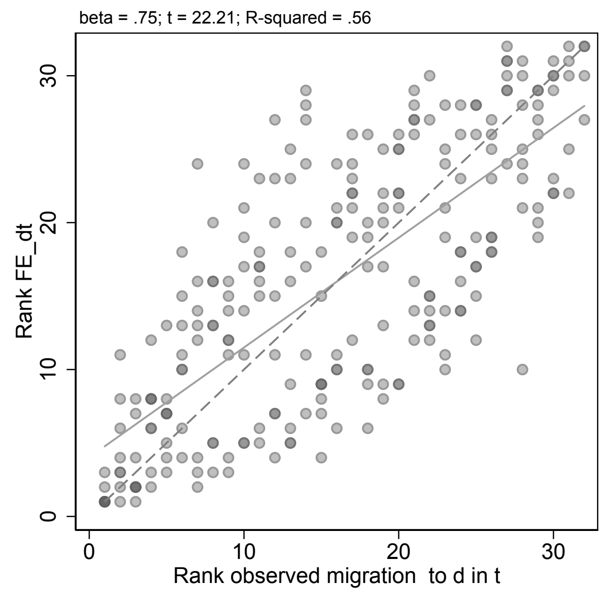
NOTE: Coefficients obtained from column 4 of Table 5.

Observed and estimated country attractiveness Figure C8 displays the correlation between the rank of countries as determined by observed migration (x-axis) and their rank based on the estimated fixed effect FE_{dt} (y-axis) obtained from models without zeros (Figure (a)), and models with zeros (Figure (b)) from the gravity specification of Table 5.

Figure C8: Relationship between observed and estimated country attractiveness (with & without zeros)



(a) Poisson without zeros



(b) Poisson with zeros

NOTE: The dashed line corresponds to the 45-degree line. Darker shading of points indicates several observations with similar x and y values.

In Table C9 we display the top 10 destinations for each year. In Panel A, we show for each year the top 10 destinations in terms of observed migration shares, and in Panel B in terms of the estimated attractiveness of destination countries. Country attractiveness is measured by the estimated fixed effects FE_{dt} based on the estimation of column 4, Table 5, and normalized by the attractiveness of Switzerland in each year. We observe a fairly close relationship between observed attractiveness and estimated attractiveness. In the bottom row of Table C9 we display the correlation between the rank of countries based on observed migration shares and the rank based on the estimated fixed effects.

C.7 First and Second Emigration Destinations

We base our gravity estimation of section 6.2 on the first country of destination. However, for a sub-set of individuals we also observe a second emigration destination. In this sub-section we document the extent of the first and second emigrations, both in terms of how many individuals are concerned, and which destination countries were chosen.

1st and 2nd emigrations by year. Out of 97,795 individuals for which we know the emigration destination, about 10,000 individuals are reported to have emigrated again to a second country, and for 8,374 or about 8% we know the first and second emigration date. In Table C10 we report the migration years to first and second destinations. The cells show the total number of individuals that migrated in the same combination of years. For example, 277 individuals migrated twice in 1933, while 460 individuals moved in 1933 and then again in 1934.

Importantly, we observe that only a small number of individuals moved twice in the same year: 2,404 people migrated twice in the same year, which is equivalent to 2.5% of the total migrants for which we know the destination, and 28.7% of those who migrated twice. This suggests that the first destination in our data represents not just a transitory country, but the outcome of a real choice. Moreover, most of the secondary migration dates fall into the years 1938 - 1940, when individuals that moved to neighboring countries escaped the threat of German occupation and went to farther away destinations.

Table C9: Observed Migration & Estimated Country × Year Fixed Effects Top 10 Destinations

	1933		1934		1935		1936		1937		1938		1939		1940	
Rank	Country	Mig. Share	Country	Mig. Share	Country	Mig. Share	Country	Mig. Share	Country	Mig. Share	Country	Mig. Share	Country	Mig. Share	Country	Mig. Share
1	Palestine	0.22	Palestine	0.32	Palestine	0.29	Palestine	0.21	United States	0.32	United States	0.28	United Kingdom	0.25	United States	0.41
2	France	0.20	Netherlands	0.15	France	0.14	United States	0.20	Netherlands	0.16	Netherlands	0.16	United States	0.12	Shanghai	0.10
3	Netherlands	0.18	United States	0.12	Netherlands	0.12	Netherlands	0.11	Argentina	0.11	Argentina	0.09	Shanghai	0.11	Argentina	0.06
4	Belgium	0.09	France	0.10	United States	0.09	United Kingdom	0.06	Palestine	0.10	United Kingdom	0.08	Netherlands	0.11	Palestine	0.05
5	United Kingdom	0.06	United Kingdom	0.06	United Kingdom	0.06	Argentina	0.06	United Kingdom	0.06	Palestine	0.07	Belgium	0.06	Brazil	0.04
6	Switzerland	0.05	Switzerland	0.03	Luxemburg	0.05	South Africa	0.05	France	0.04	France	0.04	Palestine	0.05	Netherlands	0.04
7	United States	0.04	Belgium	0.03	Argentina	0.04	France	0.05	Switzerland	0.03	Belgium	0.04	Poland	0.03	Bolivia	0.03
8	Czechoslovakia	0.03	Italy	0.03	Switzerland	0.04	Brazil	0.04	Brazil	0.02	Uruguay	0.03	France	0.03	Belgium	0.02
9	Italy	0.02	South Africa	0.02	Italy	0.03	Italy	0.03	Belgium	0.02	Switzerland	0.03	Chile	0.03	Italy	0.02
10	Poland	0.02	Czechoslovakia	0.02	Belgium	0.02	Switzerland	0.03	Italy	0.02	Shanghai	0.03	Argentina	0.03	Switzerland	0.02

	1933		1934		1935		1936		1937		1938		1939		1940	
Rank	Country	FE_{dt}	Country	FE_{dt}	Country	FE_{dt}	Country	FE_{dt}	Country	FE_{dt}	Country	FE_{dt}	Country	FE_{dt}	Country	FE_{dt}
1	Palestine	3.31	Palestine	3.54	Palestine	3.81	United States	3.93	United States	3.62	United States	3.83	Shanghai	3.25	United States	4.19
2	United States	2.12	United States	3.33	United States	3.39	Palestine	3.35	Argentina	2.84	Argentina	3.08	United States	3.15	Argentina	3.25
3	France	1.14	South Africa	1.52	Argentina	3.03	Argentina	3.08	Palestine	1.98	Uruguay	2.18	United Kingdom	2.76	Shanghai	3.13
4	Netherlands	1.08	Argentina	1.43	South Africa	2.10	South Africa	2.88	Netherlands	1.42	Palestine	1.89	Chile	2.35	Bolivia	2.60
5	South Africa	0.98	Brazil	1.17	Brazil	1.56	Brazil	2.29	Brazil	1.21	Shanghai	1.35	Argentina	2.35	Brazil	2.57
6	Brazil	0.75	Netherlands	1.02	Chile	1.48	Uruguay	2.11	Uruguay	0.99	United Kingdom	1.26	Palestine	1.84	Palestine	1.34
7	United Kingdom	0.65	United Kingdom	0.71	United Kingdom	0.94	Netherlands	1.24	South Africa	0.76	Colombia	1.24	Brazil	1.76	South Africa	1.26
8	Belgium	0.55	France	0.53	Netherlands	0.78	United Kingdom	1.21	United Kingdom	0.67	Netherlands	1.23	Uruguay	1.60	Japan	1.15
9	Argentina	0.22	Chile	0.33	Paraguay	0.23	Chile	1.00	Colombia	0.30	South Africa	1.00	Bolivia	1.58	Dominican Republic	1.08
10	Switzerland	-0.00	Mexico	0.02	Luxemburg	0.17	Paraguay	0.98	Chile	0.14	Chile	0.56	South Africa	1.56	Chile	0.92

Correlation Rank Observed Migration and Rank based on FE_{dt}								
ρ	0.75	0.73	0.68	0.83	0.86	0.83	0.77	0.54

NOTE: The table shows for each year the top 10 destinations in terms of observed migration shares in Panel A, and in terms of their estimated attractiveness in Panel B. Country attractiveness is measured by the estimated fixed effects FE_{dt} based on the estimation of Table 5, and normalized by the attractiveness of Switzerland in each year.

Table C10: Emigration to First and Second Destination by Year

Year of 1st Emigration	Year of 2nd Emigration 1933 No.	1934 No.	1935 No.	1936 No.	1937 No.	1938 No.	1939 No.	1940 No.	1941 No.	Total No.
1933	277	460	306	222	146	208	230	116	96	2,061
1934	1	57	83	83	71	97	84	45	31	552
1935	0	0	59	139	87	95	69	52	40	541
1936	1	0	0	165	149	136	135	49	37	672
1937	0	0	1	0	181	200	100	51	28	561
1938	0	0	0	0	1	427	537	319	93	1,377
1939	0	0	0	2	0	8	936	1,082	208	2,236
1940	0	0	0	0	0	1	0	148	67	216
1941	0	0	0	0	0	1	3	0	154	158
Total	279	517	449	611	635	1,173	2,094	1,862	754	8,374

NOTE: For 19 out 8,374, or 0.22%, the year of second emigration precedes the year of first emigration. We cannot know whether this is a mistake in the data, therefore we did not correct it.

1st and 2nd emigration by country. In Table C11 we compute for each country of destination the number of individuals for which it represented the first or second destination. In addition, we also report the net change in the number of migrants that results from the arrival of new migrants for which the country represented the second destination minus the outflow of migrants from the country of first destination. Countries that report a significant negative net change are neighbors of Germany, such as Belgium, France, Italy, and the Netherlands. This implies that more people moved out of the country after it was their initial choice, then were replaced by individuals that chose these countries as secondary destination. Countries with a significant positive net change in secondary migration are far away overseas countries such as Argentina, United States, or Uruguay. This finding is in line with the observation that secondary migration peaks around 1939 when the threat of German occupation in neighboring countries became apparent, leading to a second emigration for some individuals to destinations farther away.

Table C11: First and Second Emigration Destinations

Country	1st Destination IN	1st Destination Out	2nd Destination IN	Net Change
Argentina	4,408	29	1,212	1,183
Austria	408	186	29	-157
Belgium	4,287	1,010	373	-637
Bolivia	727	55	71	16
Brazil	1,557	86	140	54
Chile	1,228	34	94	60
China	167	13	8	-5
Colombia	353	13	16	3
Cuba	726	190	83	-107
Czechoslovakia	1,079	342	41	-301
Denmark	432	104	18	-86
Dominican Republic	121	7	11	4
France	5,580	1,375	755	-620
Hungary	261	18	7	-11
Italy	1,566	878	139	-739
Japan	65	58	8	-50
Luxemburg	761	417	19	-398
Mexico	187	8	28	20
Netherlands	11,652	1,878	286	-1,592
Palestine	10,807	270	901	631
Panama	119	22	8	-14
Paraguay	272	143	63	-80
Poland	1,757	86	40	-46
Portugal	242	86	51	-35
South Africa	1,278	20	88	68
Spain	537	322	56	-266
Sweden	591	71	46	-25
Switzerland	2,431	732	210	-522
United Kingdom	11,877	1,600	815	-785
United States	18,178	93	3,230	3,137
Uruguay	1,151	238	1,251	1,013
Yugoslavia	293	65	11	-54
Total	85,098	10,449	10,108	-341

NOTE: This table shows for each country for how many migrants it was the first, and for how many migrants it was the second destination. It also computes the net change in migration that results from the out-migration of migrants from their first destination, and the arrival of new migrants for which the country represented the second destination.

C.8 Limited Mobility Bias

In Table C12 we document the number of mover cells *rb* by destination *d* and year *t*. A mover cell is defined as a combination of city of residence *r* and city of birth *b* that sends migrants to *at least two* destinations in year *t*. The table also shows in the last row the total number of cells that send to *any*, i.e. at least one, destination in year *t*.

Table C12: Number of Mover Cells by Country-Year

	# mover cells rb									
Country	1933	1934	1935	1936	1937	1938	1939	1940	1941	Average
Argentina	11	10	40	89	127	262	218	34	66	95.22
Austria	31	5	3	12	10	13	12	7		11.63
Belgium	127	31	29	36	32	123	360	22	4	84.89
Bolivia					2	13	158	26	7	41.20
Brazil	26	23	15	55	43	34	179	57	10	49.11
Chile	3	7	6	18	11	25	239	7	2	35.33
China	8	3	3	2	9	98	554	74	24	86.11
Colombia			2	8	10	42	16	5	5	12.57
Cuba			4	2	3	29	114	11	59	31.71
Czechoslovakia	62	25	21	39	24	67	20	7	3	29.78
Denmark	27	11	10	10	8	8	30	3	4	12.33
Dominican Republic						3	13	10	7	8.25
France	253	73	84	65	59	146	238	9	11	104.22
Hungary	5		3	4		9	12	5	3	5.86
Italy	69	40	34	59	30	68	95	25	2	46.89
Japan						4	2	12		6.00
Luxemburg	14	13	48	10	12	24	61	6		23.50
Mexico	3	2		3	10	8	29	11	4	8.75
Netherlands	227	108	83	115	162	305	552	31	5	176.44
Palestine	300	187	160	195	122	207	351	26	6	172.67
Panama			2			9	19	12	2	8.80
Paraguay			5	14	3	22	8	11	3	9.43
Poland	9	7	7	8	4	13	49	2	3	11.33
Portugal	7	3	3	4	3	8	21	12	26	9.67
South Africa	34	23	25	81	34	39	104	15		44.38
Spain	40	16	19	19	11	7	4	10	40	18.44
Sweden	9	5	3	4	13	30	90	6	6	18.44
Switzerland	95	37	44	43	39	118	176	22	5	64.33
United Kingdom	143	75	52	106	73	235	1014	22	10	192.22
United States	89	105	79	207	220	498	708	155	151	245.78
Uruguay	2	3	5	33	22	102	124	11	5	34.11
Yugoslavia	24	7	5	9	7	8	10	13	11	10.44
Cells that send to *any* destination	4,034	2,367	2,328	3,629	3,561	6,717	12,412	2,552	1,818	4,379.78

C.9 Intention-to-Treat Estimates

In this sub-section we present the results from an Intention-to-Treat setup where the two endogenous network effects are replaced by their exogeneous network shifters, i.e., network pull and push factors.

Table C13 presents the estimated effects. The coefficients of the shifters capture the network externalities driven by push and 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 of individuals to migrate. The effect is sizeable (column 4): A one-standard-deviation increase in Pull (Push) increases migration probability by 3 (0.2) percentage points, which equals an increase of 57% (4.5%) of the sample mean.

Table C13: Intention-to-treat Estimates

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Migration Decision				
	Logit coef.	Logit dy/dx	Cloglog dy/dx	LPM coef.	LPM coef.
Detainment of network members	0.127	0.006	0.007	0.022*	0.018*
	(0.164)	(0.008)	(0.007)	(0.012)	(0.010)
Network pull	0.327***	0.015***	0.014***	0.016***	0.012***
	(0.038)	(0.002)	(0.002)	(0.002)	(0.001)
Network push	0.973	0.045	0.033	0.115*	0.165***
	(0.729)	(0.033)	(0.033)	(0.059)	(0.052)
Year FE	✓	✓	✓	✓	
CoR x Year FE					✓
Observations	1,025,497	1,025,497	1,025,497	1,025,497	1,024,976
R-squared				0.06	0.10
Pseudo R-squared	0.12	0.12			
Mean of dependent var.	0.052	0.052	0.052	0.052	0.052
SD of dependent var.	0.222	0.222	0.222	0.222	0.222

NOTE: The dependent variable is an indicator for migration. The unit of observation is an individual in year t. The sample consists of individuals who were adults in 1933. Network pull measures the cumulative average Inclusive Utility across cities of residence of distant peers. Network push measures the cumulative average detainment rates in the residence cities of distant peers. 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, a dummy for whether the individual was born outside Germany, and the partial Inclusive Utility of individual i. Marginal effects reported in columns 2 and 3. Standard errors clustered at the city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.10 First Stage Results

Table C14 presents the results of the first stage. Overall, we observe that both instruments are relevant predictors of the endogenous variables network migration $\texttt{Mignet}_{m(i)t}$ and the Inclusive Utility I_{it}. Considering specifications with only year fixed effects in columns 1 and 2, we find that both instruments predict network migration, but only $\text{Pull}_{m(i),t}$ predicts the inclusive utility, as one could expect. When estimating models with city-of-residence $\times$ year fixed effects (columns 3 and 4), we find that both instruments impact the endogenous variables positively and significantly.

Table C14: First Stage

	(1)	(2)	(3)	(4)
Dependent variable:	Migration of network members	Inclusive Utility	Migration of network members	Inclusive Utility
Detainment of network members	-0.124***	-0.043	-0.123***	-0.029**
	(0.016)	(0.031)	(0.016)	(0.012)
Network pull	0.017***	0.318***	0.026***	0.104***
	(0.003)	(0.012)	(0.003)	(0.003)
Network push	0.712***	-0.516	0.868***	0.251***
	(0.114)	(0.356)	(0.106)	(0.074)
Year FE	✓	✓		
CoR x Year FE			✓	✓
Observations	1,025,497	1,025,497	1,024,976	1,024,976
R-squared	0.60	0.96	0.63	0.99
Mean of dependent var.	0.14	-0.45	0.14	-0.45
SD of dependent var.	0.154	1.099	0.154	1.099

NOTE: The unit of observation is an individual in year t. All regressions control for age, age squared, gender, personal detainment, and a dummy for whether the individual was born outside Germany. Standard errors clustered at the city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.11 Robustness Checks

Next, we explore how the fact that migration and deportation status is missing for some distant peers affect the estimated effects of network pull and push factors. In Table C15, we test the sensitivity of our findings to excluding from the sample decision makers with a larger share of distant peers with missing information. Column 1 replicates the baseline results that correspond to column 1 of Table 7, Panel A. Column 2 presents the results excluding the top 10 percent of the decision makers in terms of the share of peers with missing information; in columns 3 through 6, we exclude in each column an additional top 10 percent of the decision makers. Our migration spillover estimates are robust to these changes in the sample and are unlikely to be driven by sample selection among distant peers. The estimated effect of detainment of network members is less robust. As we exclude more decision makers from the sample, the point estimates get smaller and less precise.

Table C15: Robustness - Sample Selection Among Distant Peers

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Migration Decision					
Sample:	Baseline	Excluding top 10%	Excluding top 20%	Excluding top 30%	Excluding top 40%	Below Median
Detainment of network members	0.427***	0.437***	0.349**	0.269*	0.191	0.180
	(0.145)	(0.149)	(0.153)	(0.147)	(0.150)	(0.153)
Migration of network members	0.209***	0.260***	0.300***	0.330***	0.358***	0.365***
	(0.078)	(0.071)	(0.071)	(0.069)	(0.072)	(0.075)
Inclusive Utility	0.681***	0.634***	0.658***	0.699***	0.702***	0.684***
	(0.042)	(0.039)	(0.036)	(0.031)	(0.031)	(0.030)
Year FE	✓	✓	✓	✓	✓	✓
Observations	1,025,497	929,771	831,556	728,188	632,145	526,586
Pseudo R-squared	0.12	0.12	0.12	0.12	0.12	0.13
Mean of dependent var.	0.052	0.050	0.049	0.048	0.048	0.046
SD of dependent var.	0.222	0.219	0.217	0.215	0.213	0.210

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

In Table C16, we check the sensitivity of our estimates to dropping small and large cities (e.g. Berlin) from the sample. The measurement errors of the network are likely to be greater in big cities: The likelihood that all people from the same generation (born ±5 years apart) know each other gets smaller as the size of the Jewish community in a city gets larger. On the other hand, the likelihood of having outliers in terms of detainment and migration rate of network members is higher in cities with small Jewish communities. Our results are robust to excluding from the list of birth cities Berlin (column 2), other big cities with a Jewish community of more than 5,000 individuals (column 3) or 2,500 individuals (column 4), and small towns with a Jewish community of less than 25 individuals (column 5) or 50 individuals (column 6). In Table C17, we impose the same set of restrictions but for the sample of residence cities. Overall, the estimated migration spillover effects are robust to the changes in the sample. Excluding Berlin and other large towns from the sample of residence cities reduces the point estimate and precision of the effect of the detainment rate of network members.

Table C16: Robustness – Size of Birth Cities

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Migration Decision					
Sample:	All	Berlin excluded	CoB $<= 5000$	CoB $<= 2500$	CoB > 25	CoB > 50
Detainment of network members	0.427***	0.413***	0.327**	0.249*	0.481***	0.732***
	(0.145)	(0.142)	(0.139)	(0.131)	(0.152)	(0.183)
Migration of network members	0.209***	0.281***	0.320***	0.315***	0.189**	0.195**
	(0.078)	(0.065)	(0.064)	(0.056)	(0.081)	(0.092)
Inclusive Utility	0.681***	0.680***	0.651***	0.665***	0.680***	0.677***
	(0.042)	(0.038)	(0.038)	(0.036)	(0.042)	(0.046)
Year FE	✓	✓	✓	✓	✓	✓
Observations	1,025,497	870,790	780,398	729,629	1,006,977	926,504
Pseudo R-squared	0.12	0.13	0.13	0.13	0.12	0.12
Mean of dependent var.	0.052	0.052	0.052	0.051	0.052	0.053
SD of dependent var.	0.222	0.221	0.223	0.219	0.223	0.224

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

Table C17: Robustness – Size of Cities of Residence

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Migration Decision					
Sample:	All	Berlin excluded	CoR $<= 5000$	CoR $<= 2500$	CoR > 25	CoR > 50
Detainment of network members	0.427***	0.268*	0.267	0.256	0.415***	0.439***
	(0.145)	(0.163)	(0.169)	(0.182)	(0.151)	(0.161)
Migration of network members	0.209***	0.178**	0.250***	0.263***	0.173**	0.159*
	(0.078)	(0.075)	(0.070)	(0.073)	(0.080)	(0.084)
Inclusive Utility	0.681***	0.817***	0.697***	0.691***	0.690***	0.703***
	(0.042)	(0.040)	(0.040)	(0.043)	(0.043)	(0.048)
Year FE	✓	✓	✓	✓	✓	✓
Observations	1,025,497	650,688	408,171	343,136	994,801	958,930
Pseudo R-squared	0.12	0.13	0.12	0.13	0.12	0.12
Mean of dependent var.	0.052	0.053	0.058	0.056	0.053	0.053
SD of dependent var.	0.222	0.225	0.234	0.230	0.223	0.224

NOTE: Logit regressions with the dependent variable being an indicator for migration. The unit of observation is an individual in year t. All regressions control for age, age squared and gender. Standard errors clustered at the city-of-birth $\times$ year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Last, we tackle potential spatial spillovers of migration and persecution across cities, which could lead to correlated effects even though peers live in a different town than the decision maker. We focus in Table C18 on a sample of decision makers who themselves are internal movers (i.e. not living in their city of birth themselves) and impose a level of distance between the birth city and city of residence of decision makers. Column 1 presents the results for a sample including all internal movers; in columns 2, 3, and 4, we impose minimal distances of 100 kms, 200 kms, and 300 kms between the city of residence of decision makers and those of their peers. This addresses potential information spillovers from peers' residence cities to decision maker's residence city. The effects of migration rate of network members and inclusive utility are robust to these changes in the sample. However, we observe a substantial heterogeneity with respect to the effect of detainment rate of network members. We do not find a statistically significant effect on the likelihood of emigrating among decision makers who are internal movers in most specifications, even though its sign is positive as in the full sample.

Table C18: Robustness – Distance between decision maker's CoR and Distant Peers' CoR, Sample of internal movers as decision makers

	(1)	(2)	(3)	(4)
Dependent variable:	Migration Decision			
Sample:	Movers	$> 100km$	$> 200km$	$> 300km$
Detainment of network members	0.180	0.198	0.501***	0.155
	(0.135)	(0.138)	(0.164)	(0.179)
Migration of network members	0.342***	0.269***	0.297***	0.195***
	(0.057)	(0.059)	(0.064)	(0.063)
Inclusive Utility	0.745***	0.809***	0.842***	0.799***
	(0.034)	(0.038)	(0.044)	(0.053)
Year FE	✓	✓	✓	✓
Observations	655,924	593,563	509,841	401,019
Pseudo R-squared	0.12	0.12	0.13	0.12
Mean of dependent var.	0.049	0.049	0.050	0.053
SD of dependent var.	0.215	0.216	0.218	0.224

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

In Table C19, we present results with individual fixed effects that capture all unobserved heterogeneity at the individual level, for example individual differences in wealth, education, occupations, identity, and membership in political parties. We estimate positive and significant effects of network migration, as well as detainment in the network, on the likelihood of emigration. In terms of magnitude, the effect of network migration is about four times as large as in OLS specifications without individual fixed effects, and the effect of network detainment is roughly twice as large. Regarding the effect of the inclusive utility, the results are more mixed as the individual fixed effects together with the city × year fixed effects absorb more than 98.5% of the variation in the inclusive utility. Overall, the findings are in line with the results that we obtained without considering unobserved heterogeneity at the individual level.

Table C19: Individual Fixed Effects

	(1)	(2)
Dependent variable:	Migration Decision	
Migration of network members	0.226***	0.216***
	(0.016)	(0.018)
Inclusive Utility	0.028***	0.010
	(0.005)	(0.008)
Detainment of network members	0.075***	0.067***
	(0.015)	(0.013)
Individual FE	✓	✓
Year FE	✓	
CoR × Year FE		✓
Observations	1,022,727	1,022,204
R-squared	0.25	0.28
Mean of dependent var.	0.047	0.047
SD of dependent var.	0.211	0.211

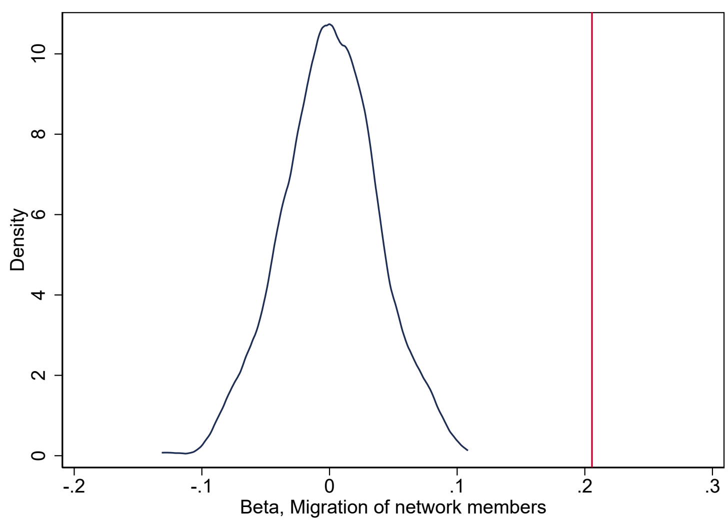
NOTE: LPM estimation with the dependent variable being an indicator for migration. The unit of observation is an individual in year t. The sample consist of individuals who were adults in 1933. All regressions control for age, age squared, gender, personal detainment, a dummy for whether the individual was born outside Germany, the partial Inclusive Utility of individual i, as well as detainment of network members. Standard errors clustered at the city-of-birth × year in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.12 Randomizing Peer Networks

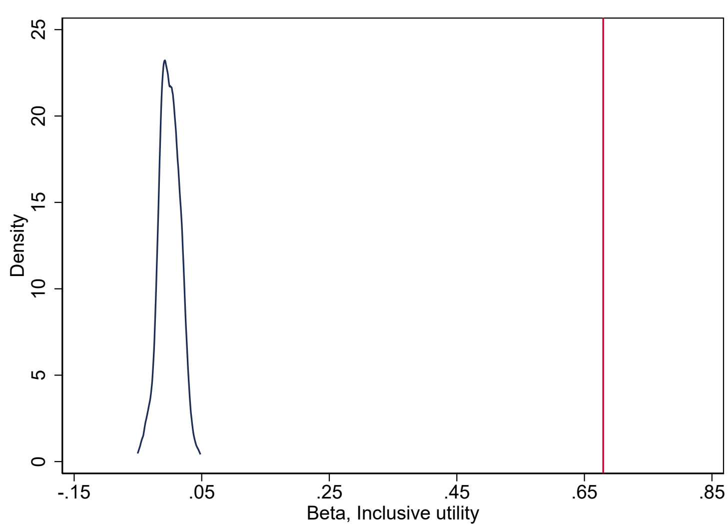
This sub-section presents the results of randomizing peer networks. We randomly assign to decision makers the peer networks of another individual and re-estimate column 1 of Table 7, Panel A. Figure C9 displays the coefficients on migration of network members and inclusive utility obtained from 1,000 such random assignments. The blue lines present the distribution of coefficients obtained from this randomization, whereas the red vertical lines represent the coefficients reported in column 1 of Table 7, Panel A. We observe that compared to the distribution of coefficients obtained from randomizing peer networks, the true coefficients are significantly larger in magnitude. Importantly, the likelihood of obtaining our estimates reported in Table 7 by chance is less than one in a thousand.

Figure C9: Placebo Test: Randomizing Peer Networks Across Individuals

(a) Coefficients on Migration of Network Members



(b) Coefficients on Inclusive Utility



D Counterfactual Simulations: Additional Material

D.1 Exact Hat Algebra

We use a simplified version of our theoretical setup, where migration predictors are defined at the city of residence (r) $\times$ city of birth (b) in the lower model (as in the estimation section) but also in the upper (outmigration) model. This means that `Detainment` and `Mignet`, used in the individual-level outmigration equation (12), are now given by:

$$\texttt{Detainment}_{rbt} \equiv \frac{\text{detained}_{r,t-1}}{\text{pop}_{r,t_0}} + \frac{\text{detained}_{b,t-1}}{\text{pop}_{b,t_0}}, \tag{3}$$

$$\texttt{Mignet}_{rbt} \equiv \frac{\text{stock}_{r,t-1}}{\text{pop}_{r,t_0}} + \frac{\text{stock}_{b,t-1}}{\text{pop}_{b,t_0}}. \tag{4}$$

where detained$_{r,t-1}$ (resp. stock$_{r,t-1}$) is the cumulative stock of Jews from city of residence r who has been detained (resp. has outmigrated) until year $t-1$. Note that we retain an additive functional form to follow the logic of the construction of the diasporas variables in the baseline model (see equation 14).

Indirect utilities can be rewritten in CES form as:[7]

$$\exp(V_{rbdt}) = \left(a_{dt} \times \texttt{diasp}_{rdt}^{\alpha_1} \times \texttt{diasp}_{bdt}^{\alpha_2}\right) / \left(\tau_{rdt} \times \tau_{bdt} \times \eta_{rdt} \times \eta_{bdt}\right), \tag{5}$$

where η_{rdt} and η_{bdt} are unobservable components of the migration costs, which will be assumed constant in the counterfactual scenarios. Regarding the choice of migrating abroad (under normalization assumption $W_{rd\neq \text{DEU}t} = 0$), we have

$$\exp(W_{rb\texttt{DEU}t}) = \texttt{mignet}_{rbt}^{-\beta} \times \texttt{detainment}_{rbt}^{-\gamma} \times \prod_k x_{krbt}^{\delta_k}. \tag{6}$$

Aggregating individual decision making at the $r \times b$ level as in equation (9), we obtain a CES equation for the observed bilateral share of migrants from $r \times b$ that choose d in year t (conditional on migrating):

$$s_{rbdt} \equiv \frac{\text{mig}_{rbdt}}{\text{mig}_{rbt}} = \frac{\left(a_{dt}\texttt{diasp}_{rdt}^{\alpha_1}\texttt{diasp}_{bdt}^{\alpha_2} / \tau_{rdt}\tau_{bdt}\eta_{rdt}\eta_{bdt}\right)^{1/\lambda_2}}{\sum_{d'\neq \texttt{DEU}} \left(a_{d't}\texttt{diasp}_{rd't}^{\alpha_1}\texttt{diasp}_{bd't}^{\alpha_2} / \tau_{rd't}\tau_{bd't}\eta_{rd't}\eta_{bd't}\right)^{1/\lambda_2}}. \tag{7}$$

At this stage, we can turn to the percentage change of that share: $\hat{s}_{rbdt} \equiv s'_{rbdt}/s_{rbdt}$, where the $'$ notation designates the level of the variable under the counterfactual scenario. The exogenous shocks relevant for the lower level that we consider are (i) percentage changes in national attractiveness ($\widehat{a}_{dt}$) and (ii) percentage changes in migration frictions as a function of changes in distance $\widehat{\tau}_{rdt} = \widehat{\texttt{dist}_{rd}}^{\tilde{\rho}_1}$ and $\widehat{\tau}_{bdt} = \widehat{\texttt{dist}_{bd}}^{\tilde{\rho}_2}$ (see equation 11). The resulting percentage changes in bilateral

[7] CES form is obtained after all migration predictors being exponentiated: $a_{dt} \equiv \exp(A_{dt}) = \frac{M_{dt}w_{dt}}{P_{dt}^{\theta}}$; $\texttt{diasp} \equiv \exp(\texttt{Diasp})$; $\texttt{mignet} \equiv \exp(\texttt{Mignet})$; $\texttt{detainment} \equiv \exp(\texttt{Detainment})$; $x_{krbt} \equiv \exp(X_{krbt})$ for all column vector X_k in the matrix of control variables $\mathbf{X}$; and $\texttt{i}_{rbt} \equiv \exp(I_{rbt})$ for the inclusive utility.

migrations can be expressed as:

$$\hat{s}_{rbdt} = \frac{\widehat{\text{mig}}_{rbdt}}{\widehat{\text{mig}}_{rbt}} = \frac{\left(\widehat{a}_{dt}\widehat{\texttt{diasp}}_{rdt}^{\alpha_1}\widehat{\texttt{diasp}}_{bdt}^{\alpha_2}/\widehat{\texttt{dist}}_{rd}^{\tilde{\rho}_1}\widehat{\texttt{dist}}_{bd}^{\tilde{\rho}_2}\right)^{1/\lambda_2}}{\widehat{\texttt{i}}_{rbt}}, \tag{8}$$

and associated percentage changes in (exponentiated) inclusive utility as:

$$\widehat{\texttt{i}}_{rbt} = \sum_{d' \neq \texttt{DEU}} s_{rbd't}\left(\widehat{a}_{d't}\widehat{\texttt{diasp}}_{rd't}^{\alpha_1}\widehat{\texttt{diasp}}_{bd't}^{\alpha_2}/\widehat{\texttt{dist}}_{rd'}^{\tilde{\rho}_1}\widehat{\texttt{dist}}_{bd'}^{\tilde{\rho}_2}\right)^{1/\lambda_2}. \tag{9}$$

Note that computing these changes does not require knowledge of the unobserved migration frictions (since η_{rdt} and η_{bdt} are assumed to be left unchanged).

We can now turn to the upper-level outmigration decision. Aggregated at the $r \times b$ level, the empirical probability of moving out of Germany (equation (7)) is simply:

$$E\left[\frac{\text{mig}_{rbt}}{\text{popatrisk}_{rbt}}\right] = \frac{\exp(\frac{\lambda_2}{\lambda_1}\ln \texttt{i}_{rbt})}{\exp(\frac{\lambda_2}{\lambda_1}\ln \texttt{i}_{rbt}) + \exp(\frac{1}{\lambda_1}W_{rb\text{DEU}t})}. \tag{10}$$

Following the same approach, we can compute the percentage change in outmigration shares:

$$\hat{S}_{rbt} = \frac{\widehat{\text{mig}}_{rbt}}{\widehat{\text{popatrisk}}_{rbt}} = \frac{(\widehat{\texttt{i}}_{rbt})^{\lambda_2/\lambda_1}}{S_{rbt}(\widehat{\texttt{i}}_{rbt})^{\lambda_2/\lambda_1} + (1-S_{rbt})\left(\widehat{\texttt{mignet}}_{rbt}^{-\beta}\widehat{\texttt{detainment}}_{rbt}^{-\gamma}\right)^{1/\lambda_1}}, \tag{11}$$

where the potential exogenous shock relates to local persecution ($\widehat{\texttt{detainment}}_{rbt}$).

The dynamics of our counterfactual analysis are driven by variables ($\texttt{diasp}$, $\texttt{mignet}$, popatrisk), which all need to be updated over time. Equations (14) and (4) link those variables to the current stocks at each destination from each city of residence/birth k. Stocks adjust in the following way: $\texttt{stock}_{kd(t+1)} = \texttt{stock}_{kdt} + \text{mig}_{kdt}$. Straightforward (but tedious) computations lead to the following percentage changes:

$$\begin{cases} \widehat{\texttt{diasp}}_{kd(t+1)} = \widehat{\texttt{diasp}}_{kdt} \times \left(\frac{\texttt{diasp}_{kd(t+1)}}{\texttt{diasp}_{kdt}}\right)^{(\widehat{\text{mig}}_{kdt}-1)} \\ \widehat{\texttt{mignet}}_{rb(t+1)} = \widehat{\texttt{mignet}}_{rbt} \times \left(\frac{\text{mig}_{rt}}{\text{popatrisk}_{rt_0}}\right)^{(\widehat{\text{mig}}_{rt}-1)} \times \left(\frac{\text{mig}_{bt}}{\text{popatrisk}_{bt_0}}\right)^{(\widehat{\text{mig}}_{bt}-1)} \\ \widehat{\text{popatrisk}}_{rb(t+1)} = \frac{\text{popatrisk}_{rbt}}{\text{popatrisk}_{rb(t+1)}} \times \left[\widehat{\text{popatrisk}}_{rbt} - \widehat{\text{mig}}_{rbt}S_{rbt}\right], \end{cases} \tag{12}$$

where percentage changes in migration flows in each city k are obtained by averaging across cities of birth:

$$\begin{cases} \widehat{\text{mig}}_{kdt} = \sum_b \frac{\text{mig}_{kbdt}}{\text{mig}_{kdt}} \times \widehat{\text{mig}}_{kbdt} \\ \\ \widehat{\text{mig}}_{kt} = \sum_b \frac{\text{mig}_{kbt}}{\text{mig}_{kt}} \times \widehat{\text{mig}}_{kbt} \end{cases} \tag{13}$$

Algorithm – As highlighted above, the procedure is quite parsimonious. The required data relates to city-level migration, i.e., the total population that is considering migration in r and b, and the bilateral flows to each country d: $\{\text{popatrisk}_{rbt}, \text{mig}_{rbdt}\}$ for years $1933 \leq t < 1942$. From those stocks and flows we can compute the observed shares (s_{rbdt}, S_{rbt}).

To fix ideas, let us consider a change in year T in the attractiveness of a given country c such that $\widehat{a}_{ct} \neq 1$ for $t \geq T$ and $\widehat{a}_{dt} = 1$ for $d \neq c$. The procedure can be decomposed into the following steps:

1. For a given period $t > T$, we set the policy changes at $\hat{a}_{dt}$ and we retrieve from the previous period the changes $\widehat{\text{popatrisk}}_{rbt}$ and $(\widehat{\texttt{diasp}}_{kdt}, \widehat{\texttt{mignet}}_{rbt})$ for each city k and destination d. Note that for the first year we have $\widehat{\text{popatrisk}}_{rbT} = \widehat{\texttt{diasp}}_{kdT} = \widehat{\texttt{mignet}}_{kT} = 1$.
2. Using (9) we compute $\widehat{\texttt{i}}_{rbt}$.
3. Using (11) we compute $\hat{S}_{rbt}$
4. Using (8) we compute $\hat{s}_{rbdt}$
5. We obtain $\widehat{\text{mig}}_{rbt} = \hat{S}_{rbt} \times \widehat{\text{popatrisk}}_{rbt}$.
6. We obtain $\widehat{\text{mig}}_{rbdt} = \hat{s}_{rbdt} \times \widehat{\text{mig}}_{rbt}$.
7. Using (13) we compute $\widehat{\text{mig}}_{kdt}$ and $\widehat{\text{mig}}_{kt}$ for each city k and destination d.
8. Using (12) we compute $\widehat{\texttt{diasp}}_{kd(t+1)}$ and $\widehat{\texttt{mignet}}_{rb(t+1)}$ and $\widehat{\text{popatrisk}}_{rb(t+1)}$ for the next period.
9. We set period at $t + 1$ and repeat steps 1 to 8 until we arrive at the final year of the scenario.

D.2 Policy Simulations

This sub-section presents the estimation results used for the counterfactual policy simulations of labor restrictions and travel subsidies.

Labor restrictions. For the simulation of counterfactual labor restrictions, we first estimate the effect of labor restrictions on the attractiveness of countries. Table C20 presents the results. 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". We regress the country attractiveness as measured by the predicted FE_{dt} obtained from the gravity specification on measures of labor market restrictions across countries. In column 1 we observe that countries with tighter labor market restrictions are significantly less attractive. We find similar results when conditioning on year fixed effects (column 2), when categorizing labor restrictions into four categories (column 3), and when adding country fixed effects and income per capita (column 4).

In the counterfactual simulation we ask, what if countries had removed their employment restrictions after the Nuremberg Laws (i.e. from 1936 onwards), instead of making access to the labor market difficult? Using the coefficient of column (1) of Table C20, we compute the counterfactual attractiveness of destination countries that results from removing labor market restrictions (i.e. moving all labor market policies to 1).

Table C20: Labor Restrictions

	(1)	(2)	(3)	(4)
Dependent variable:	Country Attractiveness (FE_{dt})			
Labor Restrictions	-0.749***	-0.739**		-0.507*
	(0.265)	(0.285)		(0.251)
Labor Restrictions = 2			-1.598*	
			(0.843)	
Labor Restrictions = 3			-1.700**	
			(0.723)	
Labor Restrictions = 4			-1.747***	
			(0.570)	
(log) GDP per capita				1.687
				(2.556)
Year FE		✓	✓	✓
Country FE				✓
Observations	173	173	173	155
R-squared	0.14	0.19	0.22	0.72
Mean of dependent var.	-2.01	-2.01	-2.01	-2.14
SD of dependent var.	2.06	2.06	2.06	2.03

NOTE: The unit of observation is a destination country. The dependent variable are the estimated fixed effects FE_{dt} based on the estimation of Table 5. Standard errors clustered at the country-level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Travel subsidies. In Table C21 we estimate the relationship between distance and the price of boat trips using historical ticket prices for 42 boat trips departing from Europe to overseas destinations. The information on boat trips (price and length of the trip in days) is taken from Löwenthal and Oppenheimer (1938).

As expected, we find a positive effect of distance on the price of boat trips measured either in Reichsmark (columns 1 and 2), or in the log of Reichsmark (columns 3 and). The elasticity reported in column 3 implies that a 1 percent increase in distance increases the price of boat trips by about 0.9 percent. Using the estimated elasticity, we calculate the travel cost (by boat) in Reichsmark from each city in Germany to each overseas destination. Based on this, we then simulate how a reduction in travel costs in the counterfactual analysis would affect migration flows.

Table C21: Travel Costs

	(1)	(2)	(3)	(4)
Dependent variable:	Price (RM)	Price (RM)	(log) Price (RM)	(log) Price (RM)
(log) Distance	454.002***	602.843***	0.932***	1.298***
	(54.258)	(118.139)	(0.071)	(0.120)
Days		-6.204		-0.015***
		(3.771)		(0.005)
Observations	42	42	42	42
R-squared	0.62	0.65	0.79	0.85
Mean of dependent var.	599.32	599.32	6.27	6.27
SD of dependent var.	291.69	291.69	0.53	0.53

NOTE: The unit of observation is a boat trip from a European port to overseas. The dependent variable is the price of the trip in Reichsmark. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

Adena, M., R. Enikolopov, M. Petrova, V. Santarosa, and E. Zhuravskaya (2015). Radio and the Rise of the Nazis in Prewar Germany. *The Quarterly Journal of Economics 130*(4), 1885–1939.

Bartrop, P. R. and Bartrop (2018). *The Evian Conference of 1938 and the Jewish Refugee Crisis*. Springer.

Cymet, D. (2010). *History vs. apologetics: the Holocaust, the Third Reich, and the Catholic Church*. Lexington Books.

Fouka, V. (2019). Backlash: The Unintended Effects of Language Prohibition in U.S. Schools after World War I. *The Review of Economic Studies 87*(1), 204–239.

Fryer Jr, R. G. and S. D. Levitt (2004). The Causes and Consequences of Distinctively Black Names. *The Quarterly Journal of Economics 119*(3), 767–805.

Goldberg, A. (2019). Canada and the Holocaust. The Canadian Encyclopedia.

Groot, M. D. (2010). Building the New Order: 1938-1945. The Spatial History Project. https://web.stanford.edu/group/spatialhistory/cgi-bin/site/pub.php?id=51.

Gruner, W. and C. Pearce (2019). *German Reich 1933–1937*. De Gruyter.

Hänisch, D. (1989). Inhalt und Struktur der Datenbank »Wahl-und Sozialdaten der Kreise und Gemeinden des Deutschen Reiches von 1920 bis 1933«. *Historical Social Research/Historische Sozialforschung*, 39–67.

Inklaar, R., H. de Jong, J. Bolt, and J. van Zanden (2018). Rebasing 'Maddison': new income comparisons and the shape of long-run economic development. *Maddison Project Working paper 10*.

Jackisch, C. (1994). Einwanderungspolitik und öffentliche Meinung in Argentinien 1933-1945. In *Alternative Lateinamerika. Das deutsche Exil in der Zeit des Nationalsozialismus: Internationales Symposium Exil in Spanien, Portugal und Lateinamerika; 1933-1945 vom 30. Sept. bis 2. Okt. 1991 am Ibero-Amerikanische Institut in Berlin*, pp. 43–53.

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

Klemperer, V. (2013). *I Shall Bear Witness: The Diaries Of Victor Klemperer 1933-41*. Orion.

Lohfeld, W. and S. Hochstadt (2006). Die Emigration jüdischer Deutscher und Österreicher nach Shanghai als Verfolgte im Nationalsozialismus. http://www.exil-archiv.de/html/themen/exilstationen/exilstationen.htm.

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

Maurer, T., M. A. Kaplan, S. M. Lowenstein, and R. Liberles (2005). *Jewish Daily Life in Germany, 1618-1945*. Oxford University Press on Demand.

Mayer, T. and S. Zignago (2011). Notes on CEPII's distances measures: The GeoDist database. Working Papers, CEPII.

Meyer, K. (1998). *Keiner will sie haben: Die Exilpolitik in England, Frankreich und den USA zwischen 1933 und 1945*. Peter Lang.

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

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

Roberg, K. W. (2009). *A Visa or Your Life!* AuthorHouse.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2019). Ipums usa: Version 9.0 [dataset]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V9.0.

Schreiber, B. (2005). *Versteckt: jüdische Kinder im nationalsozialistischen Deutschland und ihr Leben danach : Interpretationen biographischer Interviews*. Biographie- und Lebensweltforschung. Campus.

Stiftung Jüdisches Museum Berlin (Ed.) (2006). *Heimat und Exil: Emigration der Deutschen Juden nach 1933*. Jüdischer Verlag im Suhrkamp Verlag.

Strauss, H. A. (1971). The Immigration and Acculturation of the German Jew in the United States of America. *The Leo Baeck Institute Yearbook 16*(1), 63–94.

Totten, S. and S. Feinberg (2016). *Essentials of Holocaust education: Fundamental issues and approaches*. Routledge.

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

Wegner, S. (2013). *Zuflucht in einem fremden Land: Exil in Uruguay 1933-1945*. Assoziation A Verlag, Berlin - Hamburg.

Windolf, P. (2011). The German-Jewish Economic Elite (1900 to 1930). *Zeitschrift für Unternehmensgeschichte 56*(2), 135–162.

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

Zimmermann, N. M. (2013b). The supplementary cards stating information about origin and educational background of the population census of 17 May 1939. Federal Archives Germany.