Forced Migration and Human Capital Accumulation: Evidence from Post-WWII Population Transfers*

Sascha O. Becker
University of Warwick
CAGE, CEPR and CESifo

Irena Grosfeld
Paris School of Economics
CNRS

Nico Voigtländer
UCLA
NBER, CEPR and CAGE

Ekaterina Zhuravskaya
Paris School of Economics
CEPR and EHESS

Preliminary. This version: 24 February 2018

Abstract

We exploit a unique historical setting to study the long-run effects of forced migration on investment in education. As a result of World War II, the Polish borders were redrawn, resulting in large-scale migration. Poles were forced to move from the Eastern Borderlands (taken over by the USSR) and resettled to the newly acquired Western Territories, from which Germans were expelled. The largely emptied Western Territories were also the destination of voluntary migrants from Central Poland. We can thus compare *forced* migrants with *voluntary* migrants from the same ethnicity, allowing us to bypass typical confounding factors such as different cultural or linguistic background of migrants, and competition with natives in the destination labor market. We combine historical censuses with newly collected survey data to show that descendants of forced migrants are more educated than the descendants of voluntary migrants. This difference is not driven by selection of either group of migrants, by pre-war differences, or by local labor market conditions. Instead, survey evidence suggests that forced migration led to a shift in preferences, away from material possessions and towards investment in mobile assets such as human capital. The effects persist over three generations.

JEL: N33, N34, D74, I25

Keywords: Poland, Forced Migration, Uprootedness, Human Capital

*We received excellent comments at the Chicago Booth Miniconference on Economic History, at the 'Workshop in Political Economy and Economic Policy' at Queen Mary University London, at the Oxford-Warwick-LSE (OWL) Workshop in Economic History at Oxford, at DIW Berlin, and at the Universities of Bristol, Frankfurt and Warwick. We thank Luis Candelaria for very insightful discussions. Sascha O. Becker acknowledges financial support by the ESRC Centre for Competitive Advantage in the Global Economy (CAGE). Ekaterina Zhuravskaya thanks the European Research Council (ERC) for funding from the European Union's Horizon 2020 Research and Innovation program (grant agreement No. 646662).

1 Introduction

A large literature studies the economic effects of migration. This research typically focuses on two broad topics: the effect of migrants on short-run and long-run economic outcomes at their destinations,[1] and socio-economic effects on migrants themselves and on their descendants.[2] Both literatures often use refugee supply shocks due to forced migration to identify causal effects.[3] The literature often compares migrants to locals in the receiving areas, or examines changes in outcomes within either group due to migration. This implicitly defines treatment as a combination of the migration per se, plus the (often traumatic) experience that led to displacement, such as expulsions of minorities (Hornung, 2014) or large-scale population displacements after wars (Peters, 2017). Thus, interpreting these results as specific to migration is potentially problematic. It implicitly assumes that the forceful events only affected subsequent outcomes by triggering migration, but that they did not alter migrants' behavior or preferences otherwise.

In this paper, we explore a unique historical setting that allows us to compare *forced* migrants with *voluntary* migrants over the long run – both from the same ethnic group. We study population transfers of millions of Poles in the aftermath of WWII. Figure 1 illustrates the re-drawing of Poland's borders. The former eastern Polish territories (Kresy) became part of the Soviet Union, while the former German areas in the West (Western Territories) became Polish. The latter had been home to more than 8 million Germans before WWII, who had to resettle, leaving largely empty land and capital stock, with only about one million native Poles remaining there. Soviet officials forced Poles from the Kresy territories to leave – the vast majority was resettled to the largely emptied Western Territories. At the same time, the Polish administration encouraged people from Central Poland (which had been and remained Polish) to voluntarily migrate to the Western Ter-

[1] See the comprehensive discussions in Borjas (2014), Card and Peri (2016) and Dustmann, Schönberg, and Stuhler (2016).

[2] C.f. Dustmann, Frattini, and Lanzara (2012) for an overview of the literature on second-generation immigrants.

[3] For example, Card (1990) Borjas and Monras (2017), and Braun and Mahmoud (2014) use forced migration to identify the effect of migration on economic outcomes at the destination. Bauer, Braun, and Kvasnicka (2013), Nakamura, Sigurdsson, and Steinsson (2017), and Sarvimäki, Uusitalo, and Jäntti (2016) study the effect of migration on migrants themselves.

ritories. This historical experiment allows us to compare forced migrants with voluntary migrants within their destination locations in the Western Territories.

We use novel survey data with detailed information on each respondent's family history of migration. We find that descendants of forced migrants have significantly higher education than those of voluntary migrants, who in turn have higher schooling levels than the descendants of native Poles in Western Territories. Importantly, education levels of forced migrants were not higher before WWII. We also show that selection of voluntary migrants from Central Poland is unlikely to drive the observed differences. Instead, a likely explanation for our findings is that the loss of physical possessions due to *forced* migration led to a shift in preferences towards investing in mobile assets – and in particular, in human capital. This mechanism has been pointed out prominently by both historians and economists (c.f. Kuznets, 1960). However, in common settings it is notoriously difficult to single out this mechanism because forced migrants typically differ along other socio-economic characteristics such as ethnicity, language, or religion; and because they often face labor market competition with locals, affecting their educational choices. Our empirical setting bypasses these issues by comparing Poles with Poles, arriving simultaneously in a largely empty territory.

Our results suggest that the long-run impact of forced migration differs from the effect of voluntary migration, by increasing the inclination to invest in human capital. This is further corroborated by survey evidence showing that descendant of forced migrants value material goods less, while having a stronger aspiration for education of their children. Importantly, this interpretation holds only if voluntary migrants are not negatively selected with respect to their (preferences for) education. To address this issue, we provide several tests. First, we show that literacy in 1921 was if anything higher in the origin counties of voluntary migrants, as compared to forced migrants. Second, descendants of voluntary migrants in Western Territories today are on average more educated than their cousins in the Central Polish counties of origin. Third, descendants of native Poles in Western Territories (autochthons) have even lower education than the offspring of voluntary

migrants (but levels similar to natives in the rest of Poland). These results are easier to reconcile with *positive* selection of voluntary migrants from Central Poland. This would imply that if anything, our results underestimate the education advantage of forced migrants.[4] We discuss further evidence along these lines, and show that our results hold equally for urban and rural population in Western Territories. In addition, the strength of results does not vary with local features such as industrial composition, soil quality, or historical literacy in either origin or destination counties.

Our paper is related to a large literature on the effects of forced migration (see Ruiz and Vargas-Silva (2013) for a survey).[5] Key drivers of forced migration are natural disasters, international wars, and civil wars – events that lead to economic and psychological suffering. This was a concern in the immediate aftermath of WWII, with millions of Europeans displaced from their homes (see Shils, 1946) and has been shown to affect long-term health outcomes (Haukka, Suvisaari, Sarvimäki, and Martikainen, 2017).

Many papers look at relatively short-run effects of natural disasters[6] Several papers look at exogenous variation in mobility caused by public relocation programs. Examples are public housing demolitions in Chicago which forced households to relocate to private market housing using vouchers (Chyn, 2017), or large-scale resettlement programs in Indonesia (Bazzi, Gaduh, Rothenberg, and Wong, 2016) during peace times, arguably quite a different experience from forced migration as a result of wars.

Our focus is on the effects of forced migration after WWII over the very long term, in the

[4]The raw data is displayed in Figure 2. In 1921, literacy rates of Roman Catholics in Kresy (Eastern Borderlands) were lowe than in Central Poland (CP). Today, the secondary schooling rate for those residing in Western Territories (WT) and CP is nearly identical. However, in both WT and CP, those with Kresy ancestors are more educated. We will discuss potential selection issues in more detail later.

[5]One way to group migration episodes is by size of migration flow and voluntary or forced nature: migration in 'steady state' can be thought of as voluntary migration by relatively few people. Forced migration of relatively small groups describes the expulsion of minorities (e.g. Jews, Huguenots). Voluntary migration of large numbers of migrants has been studied, for instance, in the context of the Age of Mass Migration (1850–1913) to the US (Abramitzky, Boustan, and Eriksson, 2014), or the flow of >1 million Eastern Europeans to the UK after EU enlargement in 2004 (Becker and Fetzer, 2016). Our case is that of large numbers of forced (and voluntary) migrants.

[6]For instance, (see Sacerdote, 2012) looks at effects on test scores of students of students displaced from New Orleans after Hurricane Katrina. Nakamura et al. (2017) study the labour market outcomes of families displaced by the eruption of a volcano on an Island off the coast of Iceland in 1973.

generations of children, grandchildren and great-grandchildren of adult expellees. The closest literature to our context is the one of other groups of WWII expellees. Bauer et al. (2013) study the economic integration of Germans expelled from Poland's Western Territories into the West German economy. One of their findings is that migrant children tend to acquire more education than their native peers. They explain their finding by the loss of family wealth, businesses, and farms, which forced children of migrants to prepare for occupations in the wider labor market. One key element of this mechanism is congestion: farmer families had to look for work outside agriculture because agricultural land in West Germany was already held by native Germans.

Closely related is the work by Sarvimäki et al. (2016) who study the forced migration of 11% of the Finnish population after WWII. Most were farmers who were given land and assistance to continue farming in areas resembling their origin regions. Nevertheless, a quarter of a century later, they were 10–16 percentage points more likely to hold a non-agricultural job and earned 11–28% more than plausible control groups. The authors explain their findings by habit formation whereby some of those forced to leave their given environment moved into different occupations despite the support of the government to help them continue in farming. In the Finnish context, despite the government's effort to give land to farmers who were resettled, there are already autochton farmers who are well established in the market when the newly resettled expellees arrive. So while, different from the German context, as farmers do get access to farmland, they still enter a market that is already well-functioning. In contrast, in the Polish context most areas of Western Territories see a complete turnover from near 100% German population to 100% Polish population. (Our empirical analysis takes care of areas with a substantial number of autochthons.)

Finally, our results lend support to a literature that hypothesised that minorities who repeatedly faced physical expulsions tend to invest more in mobile assets compared to immobile assets. Brenner and Kiefer (1981) write that *"...a discriminated-against group which has had physical capital confiscated in the past might tend to take the probability of confiscation of an asset into consideration when making an investment. Further, a group which had been compelled to emigrate from*

a country might take the portability of an asset into consideration when making an investment in a new country, especially if it continues to face discriminations". In support of this argument, Arbatli and Gokmen (2016) observe that Armenian and Greek minorities were more educated in the Ottoman Empire. The argument has also often been made in connection with Jews (starting with Kuznets, 1960), who have been subject to repeated persecution for centuries. However, Botticini and Eckstein (2005, 2007, 2012) have challenged this view, documenting that the preference among Jews for education is religiously motivated, as a means to read the Thorah. This illustrates that in typical settings, the forced migration – education channel is hard to pin down, as minorities vary along many other dimensions.

Relative to the existing literature, we make several contributions. First, we analyze the hitherto unexplored mass population movements in post-WWII Poland, where Poles expelled from Eastern Territories (Kresy) are resettled to ex-German Western Territories. Second, we compare forced migrants from Kresy to voluntary migrants from Central Poland in order to isolate the effect of uprootedness separately from the migration experience per se. Third, we explore a unique setting, where forced and voluntary migrants move into largely empty housing stock left behind be expelled Germans, making congestion at destination less of a concern. Our results suggest that forced migration led to shifts in education preferences for families affected by it and to higher educational attainment of their descendants. This is arguably the result of 'uprootedness' caused by the expulsion experience.

The remainder of the paper is organized as follows. Section 2 provides historical background, Section 3 describes the data. Section 4 details the empirical approach, section 5 presents results. Section 6 concludes.

2 Historical Background

Before World War I, modern-day Poland was split between the German Empire, the Russian Empire, and the Austro-Hungarian (Habsburg) Empire. As a result of World War I, the Polish state

was recreated in 1918, commonly known as the Second Polish Republic. According to the 1921 census, the number of inhabitants was 27.2 million. Figure 1 shows the area of the Second Polish Republic shaded in red. The Western part the Second Polish Republic was comprised of former Prussian parts around Poznan plus the Polish Corridor separating the two parts of Prussia in the North and West of it that after WWII became called Western (and Northern) Territories (WT for short). The vast majority of the Second Polish Republic were the formerly Russian areas in the central and eastern parts. Finally, a relatively small area in the South was formerly Austro-Hungarian territory.

2.1 Forced Migration from Kresy Territories

After WWII, in accordance with the decisions taken during the Tehran, Yalta and Potsdam conferences, Poland 'moved to the West' (c.f. Davies, 1981). Poland gained the former German territories of Silesia, Pomerania and East Prussia, called by the communist propaganda "Recovered territories" and later "Western Territories" (WT). At the same time, Poland lost the *Kresy* territories (Eastern Borderlands), east of the so-called Curzon line. This post-WWII borders are shown in dark red in Figure 1. The part of Poland that belonged to the Second Polish Republic before WWII and continued to be Polish after WWII is commonly called 'Central Poland' as it is positioned between the newly acquired Western Territories and the formerly Polish Eastern Borderlands (Kresy).

As a result of the changing borders after WWII, mass population movements occurred. Before WWII, 8.5m Germans, 1m Poles, 100,000 Jews lived in what became Poland's Western Territories. At the end of WWII, an estimated 2.5-3.4m Germans[7] and 1m Poles were still in WT. Germans were expelled from WT and had to resettle to the German territories to the west of the Oder-Neisse line, leaving behind empty houses, land and factories. At the same time Poles from Kresy were forced to leave the area east of the Curzon line. This was part of a larger resettlement, where Ukrainians, Russians and Byelorussians left Poland, while Poles from the Ukrainian, Byelorussian and Lithuanian Soviet Republics were also forced to leave and resettle in the new borders of

[7] Dziewanowski (1977).

Poland.[8] The Polish State Repatriation Bureau tried to ensure an orderly movement of Poles from Kresy mainly into WT, but de facto a close to a quarter of Kresy people settled in Central Poland, especially those with existing family ties.

2.2 Voluntary Migration from Central Poland

At the same time, voluntary migrants from Central Poland moved to WT to benefit from housing, land, and infrastructure left by the Germans. Voluntary migrants from Central Poland started coming to Western Territories in spring 1945. These early settlers were mainly coming from districts across the border, sometimes on foot, or by horse cars, trucks. They were mostly looters trying to pick up the best pieces and leaving to their original places. Later on people from Central Poland were coming by trains from the regions further away from the border, in order to benefit from housing, land, and infrastructure left by the Germans (Halicka, 2013; Zaremba, 2012).

2.3 Aggregate Statistics on Population Transfers

Table 1 shows the distribution of the three origin groups, plus Poles living abroad before WWII, and their location in post-WWII Poland in either WT or CP. In the first post-WWII Polish Census in 1950, 96% of WT's total 5.6m population in 1950 were Poles, i.e., Roman Catholics and Polish native speakers. The rest were Jews (an estimated 250,000 to 300,000 Polish Jews survived WWII), who subsequently left, and Ukrainians, who were forced migrants from CP ("Operation Vistula"). Post-WWII Poland is thus largely ethnically and religiously homogeneous, composed of ethnic Poles of Roman Catholic faith that differ only in their pre-WWII origin: autochthons Poles in WT, Poles in Central Poland, and Poles from the Kresy lands. Table 1 uses data from the 1950 Census to show the distribution of the three origin groups, plus Poles living abroad before WWII, and their location in post-WWII Poland in either WT or CP. The 1950 Polish Census provides information

[8]The main reason for people from Kresy to leave Kresy was fear and coercion: from direct terror, arrests, to forced ukrainization or lithuanization, and professional or educational discrimination of Poles. The decision to leave was also facilitated by the understanding that it is better to risk uncertain fate in the West than go for certain fate in the East (in parallel with the registration for migration there was forced resettlement to the deep USSR , e.g. to Donbas mines, going on). See Kochanowski (2015).

on the place of residence on 1 September 1939. Data are available at the regional level, which can be mapped into the three major areas (Kresy, CP, and WT). Of the total population (24.6 million people), 22.8% (5.6 million people) lived in Western Territories. Within Western Territories, 2.785 million (49.7%) were from Central Poland, while 1.554 million (27.7%) were of Kresy origin, and 1.112 million (19.9%) were authochtons. The remaining 152,000 (2.7%) came from abroad. Within Central Poland, 18.355 million (96.5%) of the population were of CP origin; only 583,000 (3.1%) were from Kresy. A mere 19,000 (0.1%) in CP had lived in Western Territories before WWII, and 53,000 (0.3%) came from abroad. The numbers also show that nearly three quarters of Kresy people did move to Western Territories as intended, and about one quarter moved to CP. At the same time, people of CP origin make up close to 50% of the WT population. Autochthons are a minority with just 20% of the population. The vast majority of WT inhabitants in 1950 were migrants from either CP or Kresy (and a small number from elsewhere).

3 Data

We use numerous data sets for modern-day and historic Poland in our analysis. Two large-scale surveys capture modern-day outcomes, in particular education: the Diagnoza (Social Diagnosis) Survey, a large-scale household survey, as well as a our own Ancestry Survey. We complement and link these data with historic censuses: the 1950 Census captures the situation at the end of the mass population transfers. We also use censuses from inter-war Poland (1921 and 1931) and from before WW I: the Russian Empire Census from 1897. We describe each of these data sources in turn.

3.1 Diagnoza Survey data

The Diagnoza Survey is a large-scale household survey comparable to similar surveys in the US (PSID) or the UK ('Understanding Society'). It is a representative sample of the Polish population with 8 waves between 2000 and 2015.[9] We use the 2015 wave, for which we lobbied the statistical

[9]For further detail of the survey see http://diagnoza.com/index-en.html.

agency to include a question on whether any of the ancestors of the respondent were from Kresy. This cross-section has ca. 30,000 observations and allows us to compare education outcomes of respondents with *any* ancestors from Kresy to those without.[10] Education of respondents is given by education degrees in 4 categories (primary or less; vocational and less than secondary; secondary; more than secondary). We use the education data in various ways: (1) take the categorical variable as an outcome; (2) use a dummy for (at least) completed secondary education (compared to less than completed secondary education); (3) use a dummy for (at least) completed tertiary education; (4) use years of education (generated from the degree dummies).

3.2 Our own Ancestry Survey

In 2016, we conducted our own survey to reconstruct the geographic origin of ancestors of WT residents. We asked a professional survey company to draw a representative sample of the WT population (3,169 respondents) as well as an additional representative sample of people in WT with Kresy origin (900 respondents).[11] Specifically, we asked detailed questions about the place of living of respondents' ancestors for each ancestor in the generation of the youngest adults in 1939. For instance if the youngest person who was at least 18 years old in 1939 was the respondent's mother, we would ask information about mother and father of the respondent. If the youngest adult in 1939 was the respondent's grand-mother, we would ask information about all four grandparents. If the youngest adult in 1939 was the respondent's great-grand-father, we would solicit information about all eight great-grandparents.[12] The 4,069 respondents gave information about 13,223 ancestors in total. The vast majority of respondents was able to give a place of living for all of their ancestors, even when they were great-grandparents. In some cases, respondents might not recall the exact place of residence of some ancestors before WWII and would merely indicate a

[10]The question is "Is there anybody in your household who himself or his parents or grand parents were living before WWII in the Eastern Borderlands (Kresy)?"

[11]This oversample of 900 additional respondents with Kresy origin was done via 'random route' sampling, i.e. after interviewers had interviewed one of the randomly drawn 3.169 respondents, they would go from door to door in the neighborhood until they found a respondent with Kresy origins.

[12]We drop a handful of respondents who were themselves 18 in 1939. They are too few to constitute a meaningful group of their own.

broad area. Overall, we were able to identify and geo-code the *exact* place of residence for 11,928 out of 13,223 ancestors.

Typically, when one parent is from Kresy, the other is as well, unless the Kresy parent was single when expelled from Kresy. As a result, in the parent generation, the share of those with Kresy origin is typically either zero or one, but 0.5 only when the Kresy parent married a non-Kresy person after being expelled. Even if a young adult expelled from Kresy marries after expulsion, assortative mating between those of similar origin might generate Kresy couples. Similarly, if one particular grandparent is from Kresy, their spouse usually is from Kresy as well, unless their marriage took place after the Kresy ancestor was expelled. For descendants reporting about the grandparent generation, the share of Kresy ancestors is therefore typically 0, 0.5, or 1, but rarely 0.25 or 0.75. Finally, the same argument translates to the great-grandparent generation, where couples of great-grandparents are typically from the same region of origin.

3.3 Historical censuses

Post-WWII Polish Census 1950

The first post-War Polish Census in 1950 has important information on population movements. It asked which region (wojewodztwa) in Poland or which country people lived in before the war. In Western Territorties, this information is available by county (powiat) of residence. In the rest of Poland, the information is available by wojewodztwa.

Inter-War Polish Censuses: 1921 and 1931

We use two Censuses from the Second Polish Republic. The census closest to WWII is the 1931 one, which has information on literacy rates by region, but without further splits of the data. The 1921 Census, in contrast, has literacy rates by religious denomination, allowing us to separately measure the literacy rates of Roman Catholics and other religious denominations. Since post-WWII Poland is homogeneously Roman Catholic, the pre-WWII literacy rates of Roman Catholics are of particular interest to us – they are a very close proxy for the literacy of ethnic Poles.

Pre-WWI: Russian Empire Census 1897

The 1897 Census of the Russian Empire (Troynitsky, 1899) provides information on literacy rates in the Russian language, and also separately in the native language for other ethnicities. For our purposes, we extract the literacy of Poles in their native language.

3.4 Benchmarking Survey Data against Historic Census Data

One potential worry in our survey is recall bias. For example, more educated respondents may have more information on the location of origin of their ancestors. This is an issue with the Diagnoza Survey, which only asks about Kresy origin. This leads to a potential bias in estimating the effect of forced migration from Kresy on today's education. In our own Ancestor Survey, we ask about *all* origins for ancestors, which makes this source of bias less likely. In fact, the majority of respondents knew the name of the locality of origin of their ancestors and not only the broad region of origin.[13] The fact that 11,928 out of 13,223 ancestors gave an exact location of origin is a first indication that respondents have a good memory regarding the origin of their ancestors.

While we have no way to confirm the accuracy of information of individual respondents, we can benchmark whether, on average, the share of respondents in a location who say that their ancestors are from Kresy is in line with the share of the population in 1950, in the first post-WWII Census that comes from Kresy. Of course, this exercise is complicated by mobility of ancestors and respondents between 1950 and the survey years 2015 (Diagnoza) and 2016 (Ancestry Survey). Still, we find that the ancestry information from the two surveys lines up quite well with 1950 Census data.

Figure 3 displays the share of migrants from Kresy in the 1950 Census, plotted against the share of respondents with ancestors from Kresy in the 2015 Diagnoza Survey. Note that this exercise is done across wojewodztwa (regions) in all of Poland in 1950 because both Diagnoza and the Census cover all Polish regions. The shares of population/respondents with Kresy origin

[13] In our survey, we were impressed by how survey respondents engaged with the questionnaire. Many seemed to be so fascinated by being asked about their ancestry that they made every effort to respond accurately. Some would even check documents to make sure they are giving the most precise answer possible.

in 1950 and 2015 line up very well for most regions.[14] Information about Kresy origin is available at the more disaggregated level of powiaty (counties) in Western Territories. We repeat the above consistency check in Figure 4. This county-level exercise is more challenging for two reasons. First, mobility across county boundaries is more likely than across region boundaries. Second, in the Diagnoza survey, the number of respondents in some counties is quite small, so that measuring the share of respondents with Kresy origin becomes noisier. Despite these caveats, Figure 4 shows a tight relationship even at the county level.

Figure 5 repeats the above exercise using our 2016 Ancestry Survey in combination with the 1950 census. Since our Ancestry Survey also covers all origin locations, we can map these data to 16 regions that are also covered in the census.[15] The results Figure 5 show that there's a strong positive relationship between the two data sources, supporting the reliability of our Ancestor Survey results. In sum, the benchmarking exercises make us confident that respondents in Diagnoza and in our Ancestry Survey gave reasonable answers to the questions about their ancestral places of origin.

4 Empirical Results

Our analysis relates modern-day education outcomes to the place of origin of historical ancestors. Using our two micro data sets – the Diagnoza Survey and our own Ancestry Survey – we estimate regressions at the respondent level i:

$$Edu_i = \beta\, Kresy_i + \phi'\, \mathbf{X}_i + \eta_i\, D_{Locality}(i) + \varepsilon_i \tag{1}$$

[14] There are a few exceptions: for instance, Warsaw is below the regression line, suggesting that while in 1950 few people of Kresy origin lived there (because the majority moved straight to the Western Territories), in 2015 the share of Warsaw survey respondents with Kresy ancestors is considerably larger, likely because the capital city attracts people from all parts of the country and hence also many descendants of Kresy people. On the other side of the line are regions like Poznan with a large share with Kresy origin in 1950, which is lower among the survey respondents in 2015.

[15] The 16 origin territories include Kresy, Western Territories, and 14 regions in Central Poland.

where Edu_i denotes different measures for individual i's education. In the Diagnoza survey (which only provides information on whether respondents have any ancestor from Kresy), $Kresy_i$ is a dummy variable that takes on value one if *any* ancestor is from Kresy. When using our own Ancestry Survey (which provides information on all ancestors), we use the *share* of ancestors from Kresy. $\mathbf{X}_i$ is a vector of the respondent's demographics: age, age squared, gender, and a rural/urban dummy. Finally, $D_{Locality}(i)$ represents county (powiat) or municipality (gmina) fixed effects that control for the local socio-economic environment, such as labor market conditions. In our survey, we additionally control for the share of ancestors from rural areas, as well as for the share of ancestors from Western Territories or abroad (with the share from Central Poland as the reference group).

4.1 Diagnoza Survey Results

In this section we present our main results from the Diagnoza survey, whose 2015 wave provides information on Kresy ancestry. The advantage of Diagnoza is that it covers all of today's Polish territories with a large number of respondents. The downside is that it only includes (dichotomous) information for whether respondents had *any* ancestors from Kresy territories, but not for ancestors from other regions of origin. We first show results for education as the outcome variable and then turn to values.

Education

Table 2 presents our main results from the Diagnoza survey. It shows that – across all four measures – individuals whose ancestors were expelled from Kresy territories have significantly higher levels of education today. The relationship is not only statistically but also economically significant. Panels A and B show that individuals with ancestors from Kresy score about 0.25 points higher on the 4-category education variable (which has a mean of 2.47), and they get almost one extra year of education (mean: 11.7). In panels C and D, family roots in Kresy are associated with an 11.4 p.p. higher probability of finishing secondary education (relative to an average of 0.48),

and with 8.9 p.p. higher odds of finishing college (with mean of 0.19). Note that all regressions control for county fixed effects, which makes it unlikely that our results are confounded by regional differences in the destination of migrants (and thus the place of residence of their descendants). The coefficient sizes are very similar for respondents in rural and urban areas (columns 2 and 3). This makes it unlikely that our results are confounded by the availability of agricultural land to forced migrants. In addition, results are very similar for respondents in Central Poland and Western Territories (columns 4 and 5). In other words, it does not matter if the descendants of forced migrants from Kresy live in Central Poland or Western Territories – they enjoy an educational advantage everywhere. As we will discuss in more detail below, this makes it unlikely that our main results are driven by selection of the *control* group into Western Territories, i.e., that voluntary migrants who came from Central Poland to Western Territories were less educated, leading to a less educated comparison group for Kresy migrants in Western Territories.

Table 3 shows that the above results hold across cohorts. The coefficient on Kresy is relatively stable, but somewhat larger for older cohorts. This, together with the fact that the mean of education is higher for younger cohorts suggests that the *relative* effect of Kresy origin is stronger for older cohorts. This makes sense, given that the intergenerational transmission of preferences is not one-to-one, even when taking into account local peer effects and assortative mating of parents (c.f. Dohmen, Falk, Huffman, and Sunde, 2012).

Values

Table 4 examines questions on values from the Diagnoza survey. In the first two columns, the outcome variable is a categorical variable that reflects respondents' aspiration for the education of their children.[16] People with Kresy ancestors score 0.15 higher, relative to a mean of 4.31 and a standard deviation of 1.09. Remarkably, this result holds even after we control for the respondent's own education (column 2). In the remaining columns we examine answers to the

[16]The survey question was: 'What level of education would you like your children to attain?' The five categories 1-5 correspond to 1:vocational school ; 2: secondary ; 3: technical secondary : 4: higher BA level; 5: higher Msc level

question "What is the main condition for success in life?". Columns 3 and 4 use a dummy equal to 1 if the answer to the statement "the measure of a successful life is the possession of various material goods" is definitely yes, yes or rather yes.[17] In columns 5-6 and 7-8 the same type of questions was used: "what do you consider as the main condition of a successful life? ". If money or freedom were indicated, it is equal to 1. Columns 3 and 4 show that, among people with Kresy ancestors, a significantly lower proportion answers that "material goods" are a good measure of a successful life. The same result holds – albeit being less robust – for answers to the question whether money is a main condition for a successful life (columns 5 and 6). Finally, columns 7 and 8 show that descendants of Kresy migrants value freedom more than the rest of the Polish population. These results lend support to the interpretation that forced migration shifted the affected people's preferences towards investment in education, and away from material possessions, which can be confiscated.

4.2 Ancestry Survey Results – Respondent Level

We now turn to our Ancestor Survey, which has information on the origin of *all* ancestors in a respondent's family tree, for the generation with the youngest adults at the beginning of WWII. We can thus compute the *share* of ancestors from Kresy (average 36.7%), as well as from Central Poland (avg. 48.9%), autochthons from Western Territories (avg. 13.1%), and from abroad (avg. 2.2%). We use the share of migrants from Central Poland as the reference category. Column 1 in Table 5 shows that the share of ancestors from Kresy is associated with significantly higher levels of education. The coefficient is practically unchanged when we control for the share of ancestors from other regions and, separately, from rural areas (column 2). The results show that descendants of Poles from Western Territories (autochthons) have lower education levels (as compared to the control group – descendants from Central Poland's migrants). Overall, the ranking in terms of education by ancestors from highest to lowest is thus Kresy, Central Poland, Western Territories – or forced migration, voluntary migration, no migration.

[17] Other answers could be neither yes nor no, rather not, no, definitely no.

Individuals with ancestors from rural areas also have lower education levels. This raises the concern that our results may be confounded by rural migrants ending up in rural areas, where education is lower. In particular, if migrants from Central Poland were more rural than their Kresy counterparts, and if they were also more likely to move to rural areas, then the coefficient on Kresy would be biased upward in regressions that do not control for urban origin or destination. We explore this dimension further in columns 3 and 4, where we split the sample into respondents from rural and urban areas (i.e., by *destination* of migrants), and in addition control for the *origin* of ancestors. The results are remarkably similar for both subsamples. Finally, columns 5 and 6 show that the share of Kresy ancestors is also significantly related to the probability of finishing secondary and higher education.

While Panel A controlled for respondents' county of residence fixed effects, Panel B is even more restrictive, including municipality (gminy) fixed effects. These are smaller than local labor markets. Coefficients in both panels are very similar, suggesting that local socio-economic characteristics are unlikely to confound our results.

The main coefficient of interest in Table 5 reflects the magnitude of changes in education when moving from zero to one in the share of Kresy ancestors. The magnitude is very similar to the results in Table 2, which used a dummy for *any* ancestor from Kresy. This similarity makes sense, given that the majority of respondents with any Kresy ancestor in our dataset had *all* ancestors from Kresy (see Section 3.2).

4.3 Ancestry Survey Results – Ancestor Level

In the ancestry data, we can gain complementary insights by using the data at the ancestor (a) level. For each respondent i we then run the following regression:

$$\begin{aligned} Edu_i = \beta_1 Kresy_{ia}^{P} + \beta_2 Kresy_{ia}^{GP} + \beta_3 Kresy_{ia}^{GGP} + \gamma_1 P_{ia} + \gamma_2 GP_{ia} + \gamma_3 GGP_{ia} + ... \\ ... + WT_{ia} + Rural_{ia} + \phi' \mathbf{X}_{ia} + \eta\, D_{Locality}(i) + \varepsilon_{ia} \end{aligned} \tag{2}$$

where $Kresy_{ia}^{P}$ indicates whether ancestor a (in this case, for the parent generation P) of respondent i were from Kresy. Similarly, $Kresy_{ia}^{GP}$ and $Kresy_{ia}^{GGP}$ are indicators for Kresy origin of grandparents and great-grandparents, respectively.[18] The corresponding coefficients shed light on whether the expulsion experience fades out over time or is stable. We also control for the generation of ancestors, as reflected by the indicators P_{ia}, GP_{ia}, and GGP_{ia}, and we include indicators for whether ancestors were from Western Territories (WT_{ia}) and/or rural areas ($Rural_{ia}$). The ancestor level regressions also allow us to control for a number of ancestor level characteristics, as captured by $\mathbf{X}_{ia}$. As before, $D_{Locality}(i)$ reflect county or municipality fixed effects for the location of respondent i. We cluster error terms by respondents to account for the fact that all ancestry information of a given respondent come from the same source and education of the respondent does not vary across ancestors.

We present our ancestor level regressions in Table 6. Odd columns include county fixed effects, while even columns control for municipality fixed effects. For our degree indicator (columns 1 and 2) we find a strong positive relationship with Kresy ancestors from the parent and grandparent generations. The coefficients for the latter are less than one-half in size, but remain highly significance. The coefficients for Kresy ancestors from the great-grandparent generation are even smaller and statistically insignificant (although statistically indistinguishable from those for the grandparent generation). A very similar pattern (with declining coefficients for older ancestor generations) holds also for secondary education (columns 3-4). This suggests that the effect of a history of forced migration fades over the generations. However, there is an exception to this pattern: higher education. As columns 5 and 6 show, the coefficient here is largest for Kresy ancestors from the great-grandparent generation. Much of this is due to the fact that the great-grandparent ancestor generation is associated with younger respondents, who are more likely to obtain higher education. The proportion of respondents with higher education is 16% for those with parent migrants, 25%

[18]For example, a respondent whose mother migrated from Kresy while the father was from Western Territories would have two observations at the ancestor level: $Kresy_{ia}^{P} = 1$ and $P_{ia} = 1$ for the mother, as well as $WT_{ia}^{P} = 1$ and $P_{ia} = 1$ for the father.

for grandparent migrants, and 37% for great-grandparent migrants. When adjusting the coefficients in columns 5 and 6 accordingly, their *relative* magnitude becomes similar across generations, but they do not decline. Thus, the results suggest a stronger persistence of the expulsion-education relationship for higher education.

5 Interpretation of Results

5.1 Pre-WWII Differences in Education

An obvious concern is that Poles who were expelled from Kresy may already have been more educated before WWII. As a first glance at this issue, we use the literacy rates of Roman Catholics in the 1921 Polish Census. Their Roman Catholic religion distinguished Poles from Russians, Jews, and other ethnicities living in interwar Poland.[19] In Kresy territories, Roman Catholics had a literacy rate of 58.9%, as compared to 65.4% in Central Poland. Thus, if anything, Poles from Kresy were less educated before they were forced to migrate. This pattern holds also when we differentiate between rural and urban areas: In Kresy, urban (rural) literacy was 73.6% (55.4%), as compared to 74.1% (63.2%) in Central Poland.

5.2 Possible Selection of Migrants

Of course, the above-mentioned education statistics for 1921 only reflect averages, and selection of more or less educated migrants remains a concern. Selection is unlikely among Kresy migrants, especially for those from urban areas where expulsion was complete (see Section 2.1). However, the possibility remains that migrants from Central Poland were selected. In particular, negative selection of Central Polish migrants would mean that the control group in our study has too low education, biasing the coefficient on Kresy upward. In the following, we implement a three-step argument to address this concern.

It is important to note that selection affects our results if it affects education today, i.e., our main outcome variable. We will thus examine various potential sources of selection through this lens.

[19] The borders of interwar Poland coincide roughly with Kresy plus Central Poland (see Figure 1).

We will also differentiate between selection at the origin and selection into destinations. In terms of the latter, note that (possible) selection of educated migrants into attractive destinations within Western Territories would not affect our results, as it is captured by county or municipality fixed effects in our analysis. The two channels for selection into destinations that would bias our results are i) that educated migrants from Kresy ended up in Western Territories and/or ii) that voluntary migrants from Central Poland to Western Territories were less educated. We show below that both are unlikely to affect our results.

Selection of Migrants from Kresy?

Selection *at the origin* is unlikely among Kresy migrants, due to the large-scale efforts to expel Poles. However, some historical sources suggest that forced migration out of Kresy was not fully homogenous. In particular, the pressure on Poles to leave may have been lower in rural areas in the Belarussian and Lithuanian parts of Kresy. In Ukraine, in contrast, there was substantial animosity between Poles and Ukrainians, leading to a complete exodus of Poles (see Section 2.1). In the following, we explore this variation by restricting the sample to urban areas, and to the Ukrainian parts of Kresy. Table 7 presents our results. The dependent variables are our education categories in Panel A and an indicator for secondary education in Panel B. Regressions are run at the ancestor level, as given by (2).

Column 1 shows our main result: descendants of Kresy migrants have significantly higher education today. In columns 2 and 3 we present results for ancestors from rural and urban areas, respectively. For rural areas, we obtain coefficients that are similar in size to the main result. Crucially, the coefficient is – if anything – even larger when we restrict the sample to ancestors from urban areas, i.e., to those ancestors for whom complete expulsion was likely the case. In columns 3-6 we repeat the previous analysis, but we further restrict the sample to ancestors from the Ukrainian part of Kresy. The coefficient in column 4 (for both urban and rural locations) is very similar to the one when using all Kresy regions (column 1). This suggests that our main results indeed reflect the situation of complete expulsion of migrants. In addition, columns 5 and 6 also

show a similar pattern as columns 2 and 3: coefficients are highly significant for both rural and urban ancestors, and they are somewhat larger for the latter. These results make it unlikely that our results are driven by selection of migrants at the origin.

Education of Kresy Descendants in Central Poland and Western Territories

Nevertheless, there may have been selection *into destinations*. As Table 1 has shown, while the majority of Kresy migrants settled in Western Territories, about one quarter moved to Central Poland. If the most capable Kresy migrants moved to Western Territories, than our main results (within WT) would biased. We show that this is unlikely, elaborating on the results from Table 2, columns 4 and 5, which showed that the coefficient on Kresy ancestors is essentially the same in Central Poland and in Western Territories. Table 8 uses Diagnoza data (which are available for both CP and WT). Columns 1-3 restrict that sample to CP, while columns 4-6 present results for WT. The dependent variable is the four-category education measure in Panel A, and a dummy for secondary education in Panel B. We find that the coefficient on Kresy ancestors is very similar within both regions. In addition, within WT the coefficients are also almost identical in rural and urban areas. The only difference in coefficients is observed in CP, where the coefficient is slightly larger in urban areas. Thus, the only extent to which selective migration may affect (some of) our results is by particularly educated people with Kresy origin moving to urban areas in Central Poland. None of our main results within WT would be affected by this.

Selection of Migrants from Central Poland?

Was there selection of migrants in our control group, i.e., of voluntary movers from Central Poland to Western Territories? In particular, our results on the 'effect' of Kresy origin would be biased upward if the control group was negatively selected, that is, if people with lower education were more likely to move from Central Poland to Western Territories. We shed light on this issue in two steps. First, we look at possible selection at the regional level. We show that – if anything – migrants from Central Poland came from counties with higher literacy than Kresy migrants. Second, we examine the possibility of negative selection from within counties, at the individual

level. While we do not have historical data, we can use contemporaneous education to show that this is unlikely to affect our results: Respondents in Western Territories with ancestors from Central Poland are actually *more* educated than a reasonable comparison group – people in their Central Polish counties of origin.

1. Historical Literacy in Migrant Ancestors' Counties of Origin. In the following, we compare the historical literacy rates in the county of origin of migrants from Kresy vs. Central Poland. We combine the 1921 literacy rates with information on the place of origin of each respondent's ancestors from our Ancestor Survey. For each respondent in our survey, we thus know the literacy in the county of origin for each of their ancestors. Regressions are run at the ancestor level, and the results are presented in Table 9. For comparison, we first report our main result within the subsample for which the historical literacy data at the ancestors' origin are available. We use secondary education as the outcome variable because it is the closest to 'literacy' among our various measures for education.[20] As shown in column 1, today's secondary education is about 7 p.p. higher for respondents with ancestors from Kresy.

Next, we use literacy in 1921 as the dependent variable. The coefficient on $Kresy_{ia}$ in this regression shows the difference in literacy in 1921 in the average Kresy county of origin, as compared to the average Central Poland county of origin in our Ancestor Survey.[21] According to the coefficient in column 2, the average county of origin of Kresy ancestors actually had a 3 p.p. lower literacy rate (relative to a mean of 62%). Columns 3-6 show that a similar pattern holds when we restrict the sample to ancestors from rural origin locations or to those with urban origins. Finally, our results hold also when we use literacy rates from the 1897 Russian Census, when both (most of) Central Poland and Kresy were parts of Russia. The number of observations here is lower because the southern-most part of CP and Kresy was part of the Austro-Hungarian Empire. Nevertheless,

[20] Primary education of respondents in our Ancestor Survey is 99.7%, without any meaningful variation. Also, secondary education has a mean of 0.57, which is comparable to historical literacy rates (mean 0.62).

[21] Note that the 1921 Polish Census did not cover Western Territories (which still belonged to Germany). The historical literacy of ancestors from abroad is also not available. Thus, the exercise in Table 9 can only be performed for Kresy vs. Central Poland as ancestors' origins – the main groups of interest for our study.

even in this smaller sample we find that descendants of Kresy migrants have significantly higher rates of secondary education today (odd columns), while their ancestors came from counties with lower literacy (even columns).

2. Modern-day Education of Migrants from Central Poland

The argument above shows that – if anything – regional selection of migrants would work against our main findings. However, the possibility remains that voluntary migrants were selected from *within* Central Polish counties. In particular, if the least educated from within counties left Central Poland, then this would still introduce an upward bias for our main results. Ultimately, to tackle this problem, we would need historical individual-level data on the education of migrants. Since these are not available, we implement a check directly where the selection concern matters: if one is worried about negative selection of migrants from Central Poland, then this would be in the context of persistent lower education today, so that our control group would have unrepresentative, low education. In the following we thus check whether descendants of migrants from Central Poland have lower education than a valid comparison group today – individuals in the counties of their ancestors' origin.

To implement this check, we focus on respondents whose ancestors moved from Central Poland to WT. From our Ancestry Survey, we have information on their county of origin in Central Poland. We also know the education level today in these origin counties from respondents in the Diagnoza Survey.[22] Using the combined information, we construct, for each respondent the variable

$$DiffEdu_{ia} = Edu_i - \overline{Edu}\left(county_{ia}^{CP}\right) \tag{3}$$

where the last term reflects the average education in the county of origin of respondent i's ancestor a. Since we only look at descendants of migrants from CP, all such counties are in Central Poland. Table 10 presents the results of t-tests for the null that $DiffEdu_{ia} = 0$ (clustering standard errors at the respondent i level). The significantly positive differences indicate that descendants of CP

[22] We use only Diagnoza respondents in Central Poland without any ancestors from Kresy.

migrants who now live in WT have on average *higher* education than their 'cousins' in the origin counties in CP. This pattern holds also when we restrict attention to the much smaller subsamples where both origin and destination locations were urban (columns 3 and 4), and when both locations were rural (columns 5 and 6). This makes it unlikely that the significant differences emerge mechanically due to predominant migration from rural CP areas to urban WT locations.

The results from Table 10 complement Table 9 in suggesting that – if anything – migrants from Central Poland were positively selected. Ultimately, we cannot show whether there was actually selection in historical migrants, or whether other mechanisms drive the observed education gap between migrants from CP and people in their origin locations.[23] Nevertheless, the results are relevant for interpreting the coefficient on Kresy origin in our main regressions. They suggest that our control group – descendants of migrants from CP who now live in WT are *better* educated than their closest comparison groups.[24] Consequently, we *underestimate* the effect for Kresy migrants in WT.

5.3 Heterogeneity

If the effect of Kresy origin has to do with the move itself and not with the origin locations, origin locations should not interact with Kresy origin. To analyze this aspect, we enrich the specifications in regression (2) by adding interaction terms between the dummy of Kresy origin with characteristics of the location of origin, or destination or the difference between destination and origin characteristics.

Specifically, use the Polish Censuses in 1931 interact Kresy origin with the share of Roman Catholics, share of native Polish speakers, share of native Russian speakers, literacy rate, urbanization. Similarly, using the Polish Census in 1921, we use interaction terms with the literacy rate,

[23]For example, an alternative story is that migrants, even when not forced, revise upward the importance of human capital. This would be similar to the mechanism for forced migrants, but not as strong – thus placing voluntary migrants between stayers and forced migrants in terms of their education. Another possible explanation are labor market spillovers in Western Territories, from educated Kresy migrants onto CP migrants. This would be consistent with spillovers found in Semrad (2015).

[24]Note also that on average, education in CP and WT today is very similar (see Figure 2). Consequently, it is unlikely that CP migrants merely benefitted from a generally better education system in WT.

share of Roman Catholics, share of native Polish speakers, and the share of literate people among Roman Catholics.

Going beyond characteristics of the population, we also look at climate variables at the place of origin. This exercise is based on the idea that a large share of the population was working in agriculture, so land suitability, temperature, precipitation-evatranspiration ratio, and ruggedness are key features of the environment in which farmers work. The rationale of this exercise is that forced migrants may benefit or suffer from moving depending on the whether they benefited (or suffered) from conditions in their place of origin .

Tables 13 and 14 show that none of the interaction terms is statistically significant. We interpret this as evidence that the effect of uprootedness we uncover is universal and not driven by specific circumstances at place of origin.

6 Conclusion

Forced migration is an issue in both historic and modern times. The UNHCR estimates that more than 65 million people are currently displaced from their home regions as a result of interstate wars, civil wars, and disasters. While the immediate experience of expulsion is certainly dramatic, the long-run effects on the displaced and their descendants are less clear. In times of peace, natural disasters may only create temporary scars to survivors. Times of war may be more dramatic experiences and long-run effects depend on the circumstances that await those displaced at their destination location.

Effects of forced migration per se are notoriously difficult to distinguish from confounding factors. We study the long-run education effects of post-WWII population movements of Poles expelled from the Eastern Borderlands of Poland ('Kresy') that were taken over by the Soviet Union. In exchange, Poland gained formerly German areas that become Poland's new Western Territories. Our strategy is to compare forced migrants from Kresy (who were resettled to the newly acquired Western Territories) with voluntary migrants from Central Poland. Importantly,

both forced and voluntary migrants in Western Territories were moving into largely empty territory left behind by millions of expelled Germans. This suggests, and our empirical results confirm, that congestion is not an issue. We interpret our finding of an education effect for forced migrants as the result of migration per se. These results offer a glimmer of hope for descendants of those who experienced expulsion.

In view of large refugee flows in many parts of the world, a policy recommendation that emerges from our study is that governments in countries receiving forced migrants would be well advised to give full access to education to the forced migrants and their children. There are potentially high returns from access to education for people who have suffered from forced migration.

References

Abramitzky, R., L. Boustan, and K. Eriksson (2014). A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration. *Journal of Political Economy 122*(3), 467–506.

Arbatli, C. E. and G. Gokmen (2016). Minorities, Human Capital and Long-Run Development: Persistence of Armenian and Greek Influence in Turkey. CESifo Working Paper Series 6268, CESifo Group Munich.

Bauer, T. K., S. Braun, and M. Kvasnicka (2013). The Economic Integration of Forced Migrants: Evidence for Post-War Germany. *The Economic Journal 123*, 998–1024.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2016, September). Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia. *American Economic Review 106*(9), 2658–2698.

Becker, S. O. and T. Fetzer (2016). Does Migration Cause Extreme Voting. *CAGE Working Paper 306*.

Borjas, G. J. (2014). *Immigration Economics*. Cambridge, MA, USA: Harvard University Press.

Borjas, G. J. and J. Monras (2017). The Labour Market Consequences of Refugee Supply Shocks. *Economic Policy 32*(91), 361–413.

Botticini, M. and Z. Eckstein (2005, December). Jewish Occupational Selection: Education, Restrictions, or Minorities? *The Journal of Economic History 65*(04), 922–948.

Botticini, M. and Z. Eckstein (2007, 09). From Farmers to Merchants, Conversions and Diaspora: Human Capital and Jewish History. *Journal of the European Economic Association 5*(5), 885–926.

Botticini, M. and Z. Eckstein (2012). *The Chosen Few: How Education Shaped Jewish History, 70-1492*. Princeton University Press.

Braun, S. and T. Mahmoud (2014). The Employment Effects of Immigration: Evidence from the Mass Arrival of German Expellees in Postwar Germany. *Journal of Economic History 74*(1), 69–108.

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *ILR Review 43*(2), 245–257.

Card, D. and G. Peri (2016). Immigration Economics by George J. Borjas: A Review Essay. Volume 54, pp. 1333–1349.

Chyn, E. (2017). Moved to opportunity: The long-run effect of public housing demolition on labor market outcomes of children. *American Economic Review 2nd round RnR*.

Davies, N. (1981). *God's Playground. A History of Poland*. Offord: Oxford University Press.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2012). The Intergenerational Transmission of Risk and Trust Attitudes. *Review of Economic Studies 79*, 645–677.

Dustmann, C., T. Frattini, and G. Lanzara (2012). Educational Achievement of SecondâĂŘGeneration Immigrants: an International Comparison. *Economic Policy 27*(69), 143–185.

Dustmann, C., U. Schönberg, and J. Stuhler (2016). The Impact of Immigration: Why Do Studies Reach Such Different Results? *Journal of Economic Perspectives 30*(4), 31–56.

Dziewanowski, M. K. (1977). *Poland in the Twentieth Century*. New York: Columbia University Press.

Halicka, B. (2013). *Polens Wilder Westen. Erzwungene Migration und die kulturelle Aneignung des Oderraumes 1945âĂŞ1948*. Ferdinand SchÃűningh Verlag, Paderborn.

Haukka, J., J. Suvisaari, M. Sarvimäki, and P. Martikainen (2017). The Impact of Forced Migration on Mortality: A Cohort Study of 242,075 Finns from 1939–2010. *Epidemiology 28*(4), 587–593.

Hornung, E. (2014). Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia. *American Economic Review 104*(1), 84–122.

Kochanowski, J. (2015). Wysiedlenie zwane repatriacja. *Kresy Rzeczpospolitej. Wielki mit Polakow 2*.

Kuznets, S. (1960). Economic Structure and Life of the Jews. In L. Finkelstein (Ed.), *The Jews: Their History, Culture, and Religion*, Chapter 39, pp. 1597–1666. Jewish Publication Society of America.

Nakamura, E., J. Sigurdsson, and J. Steinsson (2017). The Gift of Moving: Intergenerational Consequences of a Mobility Shock.

Peters, M. (2017). Refugees and Local Agglomeration – Evidence from Germany's Post-War Population Expulsions.

Ruiz, I. and C. Vargas-Silva (2013). The economics of forced migration. *The Journal of Development Studies 49*(6), 772–784.

Sacerdote, B. (2012). When the Saints Go Marching Out: Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita. *American Economic Journal: Applied Economics 4*(1), 109–135.

Sarvimäki, M., R. Uusitalo, and M. Jäntti (2016). Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations.

Shils, E. A. (1946). Social and Psychological Aspects of Displacement and Repatriation. *Journal of Social Issues 2*(3), 3–18.

Troynitsky, N. (1899). *The First Total Census of Russian Empire in 89 Volumes*. The Central Statistical Bureau of the Ministry of Internal Affairs of the Russian Empire, St. Petersburg,1899-1905.

Zaremba, M. (2012). *Wielka Trwoga. Polska 1944-1947*. Kraków: Wydawnictwo Znak.

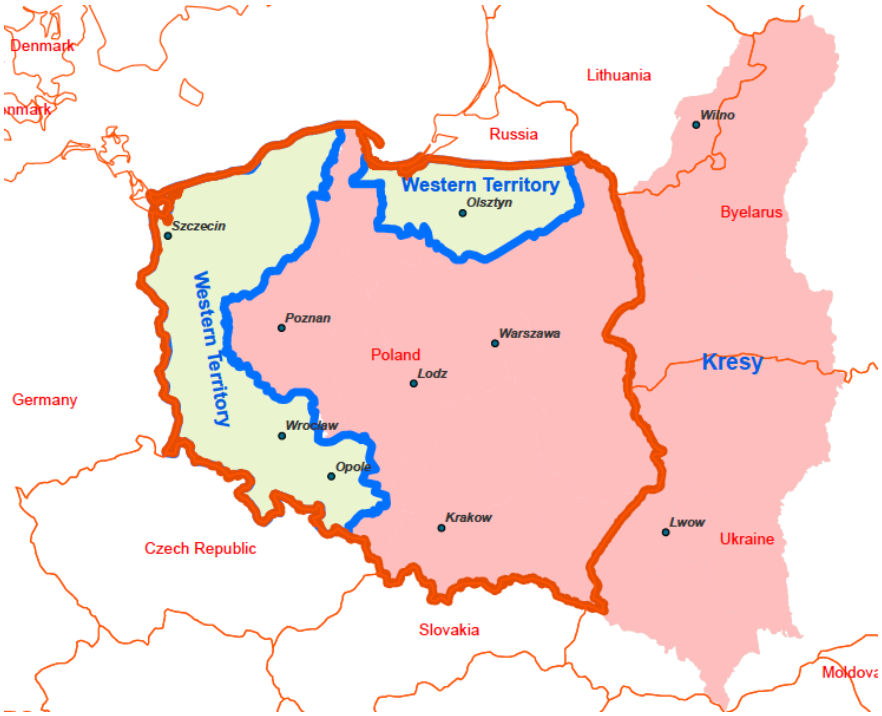


Figure 1: Map of Poland's Territorial Change after WWII

Note: The figure illustrates the re-drawing of Poland's borders after WWII. The former eastern Polish territories (Kresy) became part of the Soviet Union, while the former German areas in the West (Western Territories) became Polish. Soviet officials forced all Poles from the Kresy territories to leave – the vast majority was resettled to the emptied Western Territories. At the same time, the Polish administration encouraged people from Central Poland (which had been and remained Polish) to voluntarily migrate to the Western Territories.

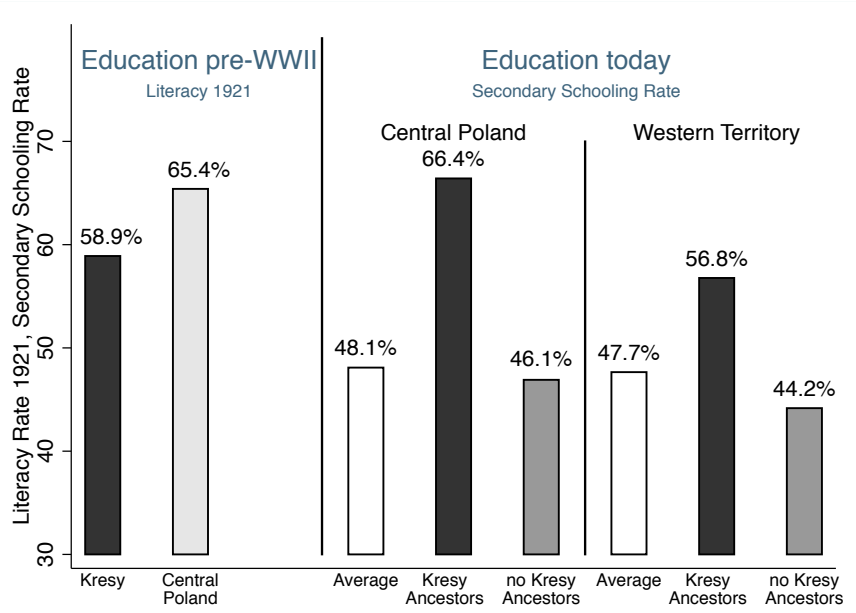


Figure 2: Overview of Historical and Contemporaneous Patterns in Education

Note: The figure shows that a) the origin location of forced migrants (Kresy) had lower education before WWII, and b) descendants of forced Kresy migrants today have higher educational attainment. For 1921, the figure displays literacy rates of Roman Catholics in Kresy (Eastern Borderlands) and Central Poland (CP). Literacy rates were lower in Kresy than in CP. For today, the figure shows the secondary schooling rate for those residing in Western Territories (WT) and CP, respectively. In both regions, the figure shows the average independent of ancestry, which is nearly identical (47.7% in WT and 48.1% in CP), as well as for those with and without Kresy ancestors. In both WT and CP, those with Kresy ancestors are more educated. Those with Kresy ancestors in WT, are – if anything – less educated than those with Kresy origin in CP (this difference disappears after accounting for urban vs. rural education differences).

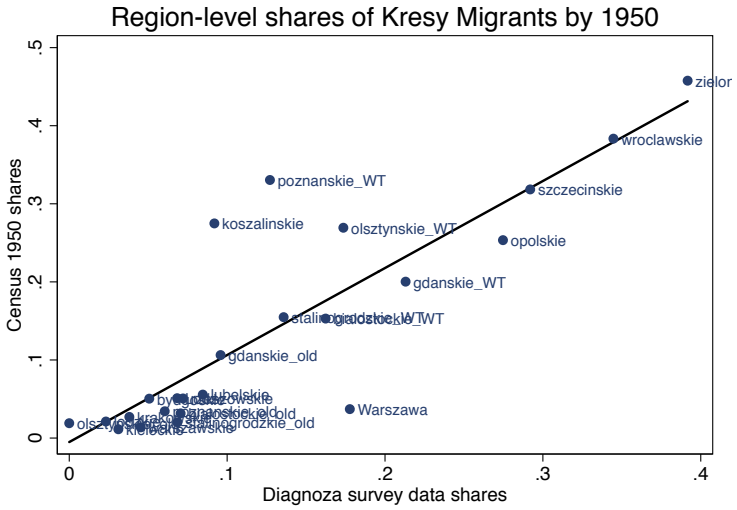


Figure 3: Region-level: Migrants from Kresy in the Diagnoza survey (2015) vs. the 1950 Census

Note: The figure plots the regional share of migrants from Kresy territories in the 1950 census (y-axis) against the Kresy migrant share from the 2015 Diagnoza data. The variation is at the region (voivodeship). Data are available for 24 regions, covering all of Poland. The regression coefficient is 1.00 with a standard error of 0.057 and R^2 of 0.74.

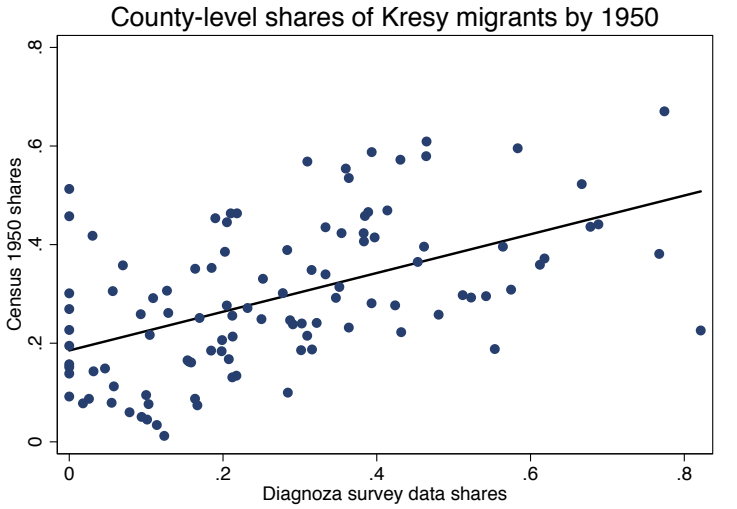


Figure 4: County-level migrants from Kresy in the Diagnoza survey (2015) vs. the 1950 Census

Note: The figure plots the county (powiaty) level share of migrants from Kresy territories in the 1950 census (y-axis) against the Kresy migrant share from the 2015 Diagnoza data. Data are available for 107 counties in the Western Territories of Poland. The regression coefficient is 0.39 with a standard error of 0.071 and R^2 of 0.27.

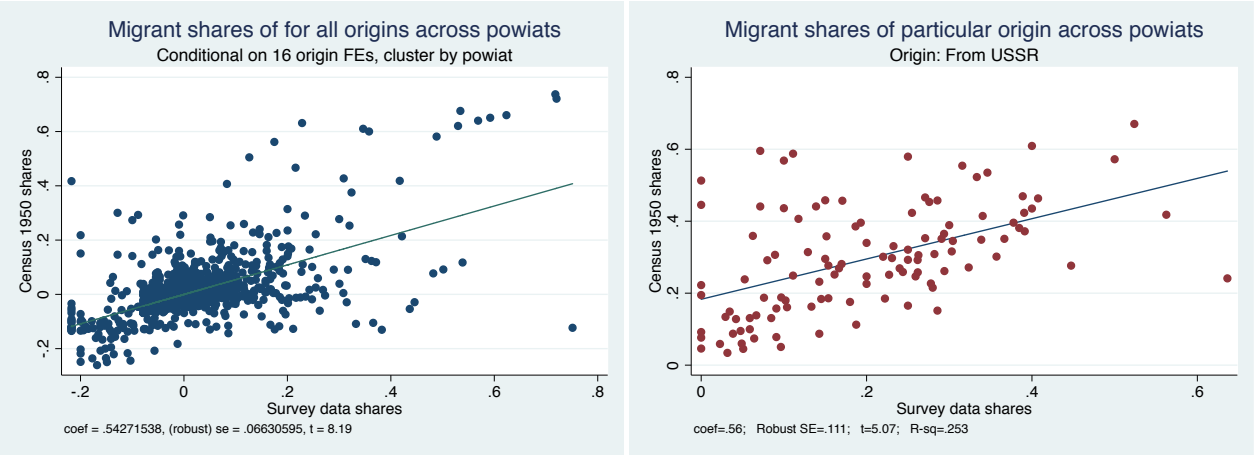


Figure 5: Data quality check of our Ancestry Survey

Note :The left panel plots the county (powiaty) level share of migrants from 16 origin territories in the 1950 census (y-axis) against the migrant share from the 2016 Ancestry Survey. The 16 origin territories include Kresy, Western Territories, and 14 regions of pre-WWII Poland. The regression coefficient is 0.54 with a standard error of 0.066 and R^2 of 0.62. The right panel repeats this exercise, but using only migrants from Kresy. The regression coefficient is 0.56 with a standard error of 0.11 and R^2 of 0.25.

TABLES

Table 1: Overview: Population Census in 1950 (in thousands)

	Western Territories (WT)	Central Poland (CP)	Share in Western Territories
Total population, 1950	5,602	19,012	22.8%
By Region of Origin:			
Lived in Central Poland in 1939	2,785	18,355	13.2%
	(49.7%)	(96.5%)	
Lived in USSR (Kresy) in 1939	1,554	583	72.7%
	(27.7%)	(3.1%)	
Lived in Western Territories in 1939	1,112	19	98.3%
	(19.9%)	(0.1%)	
Lived abroad (not USSR) in 1939	152	53	74.0%
	(2.7%)	(0.3%)	

Notes: The table shows the population of Poland in 1950 in by area of residence, as well as origin. The three major areas are Kresy (which became part of the Soviet Union after WWII), Central Poland (which had been and remained Polish), and Western Territories (which had been German and became Polish).

Table 2: Average education of individuals in the 2015 Diagnoza

Dependent variable: Individual-Level Education, as indicated in table

	(1)	(2)	(3)	(4)	(5)
Sample:	All	Rural	Urban	Central Poland	Western Territories
Panel A. Dep. Var.: Level of degree (4 categories)					
Ancestor from Kresy	0.257***	0.227***	0.272***	0.263***	0.246***
	(0.023)	(0.040)	(0.027)	(0.033)	(0.031)
Mean Dep. Var.	2.47	2.21	2.74	2.48	2.46
Observations	29,104	14,656	14,448	21,807	7,297
Panel B. Dep. Var.: Years of education					
Ancestor from Kresy	0.849***	0.719***	0.917***	0.937***	0.752***
	(0.073)	(0.118)	(0.093)	(0.111)	(0.098)
Mean Dep. Var.	11.73	10.95	12.51	11.75	11.65
Observations	29,161	14,677	14,484	21,854	7,307
Panel C. Dep. Var.: Secondary education dummy					
Ancestor from Kresy	0.114***	0.111***	0.119***	0.114***	0.111***
	(0.011)	(0.019)	(0.013)	(0.016)	(0.015)
Mean Dep. Var.	0.48	0.36	0.61	0.48	0.48
Observations	29,104	14,656	14,448	21,807	7,297
Panel D. Dep. Var.: Higher education dummy					
Ancestor from Kresy	0.089***	0.062***	0.102***	0.116***	0.064***
	(0.010)	(0.015)	(0.013)	(0.015)	(0.013)
Mean Dep. Var.	0.19	0.12	0.28	0.20	0.19
Observations	29,104	14,656	14,448	21,807	7,297
Respondent county FE	✓	✓	✓	✓	✓
Controls‡	✓	✓	✓	✓	✓

Notes: The table shows that individuals whose ancestors were expelled from the USSR Kresy territories have significantly higher levels of education today. Regressions are run at the individual level; standard errors are clustered at the household level. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties.

Table 3: Education and Kresy ancestors – Across Cohorts

Dependent variable: Individual-Level Education, as indicated in table. Full sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age:	21-30	31-40	41-50	51-60	61-70	71-80	81+
Age in 1945:	-	-	-	-	-	0-10	11+
Panel A: Dep. Var.: Level of degree (4 categories)							
Ancestor from Kresy	0.141***	0.241***	0.330***	0.355***	0.283***	0.303***	0.318***
	(0.044)	(0.050)	(0.053)	(0.048)	(0.051)	(0.077)	(0.116)
Mean Dep. Var.	3.05	3.00	2.61	2.49	2.33	2.04	1.72
R-squared	0.28	0.29	0.28	0.25	0.26	0.36	0.50
Observations	4,165	4,070	4,129	5,111	4,738	2,558	1,338
Panel B: Dep. Var.: Secondary education dummy							
Ancestor from Kresy	0.084***	0.076***	0.144***	0.174***	0.103***	0.139***	0.142***
	(0.024)	(0.024)	(0.026)	(0.024)	(0.024)	(0.034)	(0.050)
Mean Dep. Var.	0.77	0.68	0.50	0.45	0.43	0.36	0.27
R-squared	0.20	0.25	0.25	0.23	0.24	0.33	0.46
Observations	4,165	4,070	4,129	5,111	4,738	2,558	1,338
Panel C: Dep. Var.: Higher education dummy							
Ancestor from Kresy	0.041	0.125***	0.144***	0.126***	0.093***	0.068**	0.054
	(0.027)	(0.029)	(0.026)	(0.023)	(0.022)	(0.031)	(0.041)
Mean Dep. Var.	0.30	0.38	0.21	0.16	0.15	0.14	0.09
R-squared	0.30	0.27	0.23	0.19	0.16	0.22	0.33
Observations	4,165	4,070	4,129	5,111	4,738	2,558	1,338
Respondent county FE	✓	✓	✓	✓	✓	✓	✓
Controls‡	✓	✓	✓	✓	✓	✓	✓

Notes: The table shows that the results from Table 2 hold across different age cohorts. Regressions are run at the individual level; standard errors are clustered at the household level. * p<0.1, ** p<0.05, *** p<0.01.
‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties.

Table 4: Values on Education and Material Possession

Dependent variable: Individual-Level Attitudes, as indicated in table. Full sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Aspiration for education of own children (1 low – 5 high)		What is the main condition for success in life?					
			Material goods		Money		Freedom	
Ancestor from Kresy	0.151***	0.136*	-0.073***	-0.051***	-0.028**	-0.011	0.020***	0.020***
	(0.051)	(0.072)	(0.013)	(0.013)	(0.011)	(0.011)	(0.006)	(0.006)
Degree: Vocational/incompl. secondary		0.365***		-0.009		-0.019*		-0.008*
		(0.066)		(0.011)		(0.010)		(0.004)
Degree: Secondary edu.		1.268***		-0.057***		-0.067***		-0.014***
		(0.095)		(0.012)		(0.010)		(0.004)
Degree: Higher edu.		1.766***		-0.165***		-0.117***		-0.008
		(0.120)		(0.013)		(0.012)		(0.005)
log HH income		0.120***		0.005		-0.020***		-0.001
		(0.024)		(0.004)		(0.004)		(0.002)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	4.31	4.18	0.56	0.56	0.28	0.28	0.05	0.05
R-squared	0.21	0.33	0.12	0.13	0.11	0.12	0.08	0.08
Observations	8,470	4,573	22,087	21,440	21,732	21,092	21,623	20,983

Notes: The table shows that descendants of Kresy migrants have stronger preferences for the education of their children, value material goods less, and value freedom more. Regressions are run at the individual level; standard errors are clustered at the household level. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties.

Table 5: Results from the Ancestor Survey: Education in Western Territories

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Education Degree (4 categories)				Secondary	Higher
Notes on sample:			rural	urban		
Panel A: Respondent County Fixed Effects						
Share of Ancestors, Kresy	0.298***	0.265***	0.208***	0.260***	0.118***	0.077***
	(0.038)	(0.043)	(0.076)	(0.052)	(0.021)	(0.018)
Share of Ancestors, WT		-0.300***	-0.163	-0.398***	-0.150***	-0.115***
		(0.061)	(0.101)	(0.079)	(0.031)	(0.024)
Share of Ancestors, abroad		-0.272	-1.071**	0.005	-0.005	0.061
		(0.259)	(0.437)	(0.286)	(0.115)	(0.100)
Share of Ancestors, rural		-0.147***	-0.141	-0.164***	-0.060**	-0.035*
		(0.050)	(0.115)	(0.056)	(0.024)	(0.020)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	2.73	2.64	2.31	2.80	0.47	0.22
R^2	0.24	0.27	0.31	0.25	0.19	0.18
Observations	3,716	3,668	1,110	2,558	3,668	3,668
Panel B: Respondent Municipality Fixed Effects						
Share of Ancestors, Kresy	0.297***	0.240***	0.151*	0.270***	0.101***	0.075***
	(0.039)	(0.045)	(0.079)	(0.055)	(0.022)	(0.019)
Share of Ancestors, WT		-0.302***	-0.314***	-0.336***	-0.154***	-0.110***
		(0.062)	(0.108)	(0.080)	(0.034)	(0.026)
Share of Ancestors, abroad		-0.107	-0.831**	0.147	0.063	0.089
		(0.203)	(0.382)	(0.237)	(0.101)	(0.100)
Share of Ancestors, rural		-0.171***	-0.092	-0.174***	-0.076***	-0.043**
		(0.049)	(0.103)	(0.057)	(0.024)	(0.020)
Controls‡	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	2.73	2.64	2.31	2.80	0.47	0.22
R^2	0.33	0.37	0.46	0.32	0.29	0.26
Observations	3,716	3,668	1,110	2,558	3,668	3,668

Notes: The table uses data from our 2016 Ancestor Survey in Western Territories, showing that the share of ancestors from Kresy in a respondent's family tree is associated with higher levels of education. Regressions are run at the individual level; robust standard errors indicated in parenthesis. * p<0.1, ** p<0.05, *** p<0.01. Excluded category is ancestors from Central Poland. Average origin of ancestors: 48.9% from Central Poland, 36.7% from Kresy, 13.1% from Western Territories (autochthons), 2.2% from abroad.

‡ Controls include respondents' gender, age, age^2 and an indicator for rural location.

Table 6: Ancestor-level regressions for education outcomes

Dependent variable: Education at the Respondent Level in 2016

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Degree (4 categories)		Secondary education		Higher education	
Ancestor from Kresy (Parent)	0.290***	0.277***	0.135***	0.124***	0.039*	0.045**
	(0.054)	(0.054)	(0.025)	(0.026)	(0.020)	(0.022)
Ancestor from Kresy (Grandparent)	0.122***	0.108***	0.056***	0.048**	0.037**	0.033*
	(0.039)	(0.038)	(0.021)	(0.020)	(0.019)	(0.019)
Ancestor from Kresy (Great-grandparent	0.079	0.064	0.030	0.017	0.110***	0.092**
	(0.060)	(0.057)	(0.034)	(0.031)	(0.038)	(0.036)
Grandparent	0.114**	0.106*	0.040	0.042	0.000	-0.008
	(0.054)	(0.054)	(0.026)	(0.027)	(0.021)	(0.023)
Great-grandparent	0.401***	0.345***	0.213***	0.189***	0.090**	0.070*
	(0.075)	(0.076)	(0.040)	(0.040)	(0.037)	(0.039)
Ancestor from WT	-0.278***	-0.268***	-0.150***	-0.142***	-0.106***	-0.100***
	(0.044)	(0.043)	(0.025)	(0.025)	(0.020)	(0.021)
Ancestor from rural area	-0.161***	-0.164***	-0.072***	-0.074***	-0.042***	-0.041***
	(0.031)	(0.029)	(0.016)	(0.015)	(0.015)	(0.015)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓		✓		✓	
Municipality FE		✓		✓		✓
Mean Dep. Var.	2.83	2.83	0.55	0.55	0.26	0.26
R^2	0.27	0.36	0.21	0.30	0.19	0.26
Observations	11,548	11,548	11,548	11,548	11,548	11,548

Notes: The table shows that the relationship between forced migration (ancestors from Kresy) and education fades over the generations – except for higher education, where it is relatively stable (accounting for the fact that the proportion of respondents with higher education is larger among younger respondents, with great-grandparent ancestors from WWII). Regressions are run at the ancestor level; standard errors clustered by individual respondents. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, and an indicator that takes on value one if the respondent lives in a rural area.

Table 7: Results from Ancestor Survey, by origin of ancestors

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
"Ancestors from Kresy" includes:	All Kresy Ancestors			Only Kresy Ancestors from Ukraine		
Notes on sample:	all	rural	urban	all	rural	urban
Panel A: Dep. Var.: Education Degree (4 categories) in 2016						
Ancestor from Kresy	0.163***	0.143***	0.200***	0.147***	0.120***	0.175***
	(0.029)	(0.034)	(0.050)	(0.035)	(0.042)	(0.057)
Ancestor from rural area	-0.160***			-0.152***		
	(0.031)			(0.033)		
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	2.83	2.77	3.02	2.81	2.75	2.98
R^2	0.27	0.30	0.29	0.27	0.31	0.30
Observations	11,548	8,598	2,950	10,237	7,669	2,568
Panel B: Dep. Var.: Secondary Education in 2016						
Ancestor from Kresy	0.075***	0.061***	0.110***	0.067***	0.047**	0.102***
	(0.015)	(0.018)	(0.025)	(0.018)	(0.022)	(0.029)
Ancestor from rural area	-0.072***			-0.071***		
	(0.016)			(0.017)		
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.55	0.52	0.65	0.54	0.51	0.63
R^2	0.21	0.22	0.26	0.20	0.23	0.27
Observations	11,548	8,598	2,950	10,237	7,669	2,568

Notes: The table uses data from our 2016 Ancestor Survey in Western Territories, showing that the coefficient on Kresy ancestors is, if anything, larger for ancestors from urban areas (where expulsion from Kresy was complete), and that the coefficient is robust to using only the Ukrainian part of Kresy, where expulsions were also nearly complete, leaving essentially no scope for selection at the origin locations. Regressions are run at the individual level; robust standard errors indicated in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, an indicator for respondents living in rural locations, and an indicator for ancestors from Western Territories. Excluded category is ancestors from Central Poland.

Table 8: Kresy origin and education: Central Poland vs. Western Territories

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Only Central Poland			Only Western Territories		
Locations included:	all	rural	urban	all	rural	urban
Panel A: Dep. Var.: Education Degree (4 categories) in 2016						
Ancestor from Kresy	0.263***	0.196***	0.308***	0.246***	0.246***	0.240***
	(0.033)	(0.060)	(0.038)	(0.031)	(0.053)	(0.038)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	2.48	2.23	2.77	2.46	2.13	2.68
R-squared	0.31	0.31	0.27	0.28	0.25	0.26
Observations	21,807	11,736	10,071	7,297	2,920	4,377
Panel B: Dep. Var.: Secondary Education in 2016						
Ancestor from Kresy	0.114***	0.102***	0.126***	0.111***	0.115***	0.112***
	(0.016)	(0.029)	(0.018)	(0.015)	(0.026)	(0.019)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.48	0.37	0.61	0.48	0.31	0.58
R-squared	0.23	0.21	0.20	0.22	0.18	0.19
Observations	21,807	11,736	10,071	7,297	2,920	4,377

Notes: The table uses data from our 2016 Ancestor Survey in Western Territories, showing that the coefficient on ancestors from Kresy is very similar in Central Poland and in Western Territories, in both rural and urban respondent locations. Regressions are run at the individual level; robust standard errors indicated in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, an indicator for respondents living in rural locations, and an indicator for ancestors from Western Territories. Excluded category is ancestors from Central Poland.

Table 9: Education Today and in Counties of Origin of Ancestors

Dependent variable: as indicated in table header

Dependent Variable:	(1) Secondary Edu in 2016	(2) Historical Literacy	(3) Secondary Edu in 2016	(4) Historical Literacy	(5) Secondary Edu in 2016	(6) Historical Literacy
Sample. Ancestor from:	Rural & Urban origin		Rural origin		Urban origin	
Panel A: Literacy from the 1921 Polish Census						
Ancestor from Kresy	0.075***	-0.030***	0.061***	-0.041***	0.112***	-0.002
	(0.015)	(0.004)	(0.018)	(0.005)	(0.025)	(0.005)
Ancestor from rural area	-0.063***	-0.169***				
	(0.017)	(0.003)				
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.57	0.62	0.54	0.58	0.66	0.75
R^2	0.20	0.39	0.22	0.11	0.24	0.25
Observations	9,744	9,744	7,199	7,199	2,545	2,545
Panel B: Literacy from the 1897 Russian Census						
Ancestor from Kresy	0.147***	-0.032***	0.139***	-0.031***	0.148**	-0.031***
	(0.030)	(0.003)	(0.034)	(0.004)	(0.072)	(0.008)
Ancestor from rural area	-0.046	0.004				
	(0.034)	(0.004)				
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.58	0.16	0.57	0.16	0.63	0.15
R^2	0.31	0.28	0.31	0.30	0.55	0.52
Observations	2,177	2,177	1,744	1,744	433	433

Notes: The table shows that descendants of Kresy migrants have significantly higher rates of secondary education today (odd columns), while their ancestors came from counties with *lower* literacy (even columns). Regressions are run at the ancestor level; standard errors clustered by individual respondents. * p<0.1, ** p<0.05, *** p<0.01.
‡ Controls include respondents' gender, age, age^2, and an indicator that takes on value one if the respondent lives in a rural area.

Table 10: Education difference between destination and origin of migrants from CP to WT

Dep. Var.: Difference in education between destination and origin

Dep. Var.:	(1) Degree (4 cat.)	(2) Secondary Education	(3) Degree (4 cat.)	(4) Secondary Education	(5) Degree (4 cat.)	(6) Secondary Education
Sample:	All ancestors from CP		Urban origin & destination		Rural origin & destination	
Diff. in Mean	0.453***	0.112***	0.388***	0.076***	0.246***	0.032
	(0.025)	(0.012)	(0.047)	(0.024)	(0.051)	(0.025)
Observations	5,395	5,395	1,079	1,079	1,196	1,196

Notes: The table shows that respondents in WT, who are descendants of migrants from Central Poland, are better educated than people in places of their origin (in Central Poland). Regressions are run at the ancestor level; standard errors clustered by respondent. * p<0.1, ** p<0.05, *** p<0.01.

Table 11: Congestion? The Role of Autochthons in Western Territories

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)
	Respondents in WT			
	Degree (4 cat.)	Secondary education	Degree (4 cat.)	Secondary education
Ancestor from Kresy	0.217***	0.104***	0.221***	0.107***
	(0.032)	(0.017)	(0.031)	(0.016)
Ancestor from Kresy × Share of autochthons, 1950	0.193	0.075		
	(0.142)	(0.074)		
Share of autochthons, 1950	0.236**	0.051		
	(0.113)	(0.059)		
Ancestor from Kresy × Share of Polish speakers, 1900			0.159	0.055
			(0.104)	(0.063)
Share of Polish speakers, 1900			0.316***	0.123**
			(0.114)	(0.056)
Controls‡	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
Mean Dep. Var.	2.46	0.48	2.46	0.48
R^2	0.20	0.15	0.20	0.15
Observations	7,297	7,297	7,297	7,297

Notes: The table shows that the coefficient on Kresy ancestry does not vary significantly with the share of autochthons in the 1950 Census. This suggests that congestion is *not* an important part of the story. Regressions are run at the level of respondents in DIAGNOZA; standard errors clustered by powiat. * p<0.1, ** p<0.05, *** p<0.01.
‡ Controls include respondents' gender, age, age^2, as well as ancestor level dummies.

Table 12: Congestion Regressions at the Ancestor Level

Dependent variable: as indicated in table header

	(1) Degree (4 cat.)	(2) Degree (4 cat.)	(3) Incomplete (Secondary)	(4) Complete (Secondary)
Ancestor from Kresy	0.169***	0.070***	0.163***	0.072***
	(0.036)	(0.019)	(0.034)	(0.018)
Ancestor from WT	-0.311***	-0.163***	-0.288***	-0.146***
	(0.087)	(0.039)	(0.082)	(0.036)
Ancestor from Kresy × Share of autochtons, 1950	-0.042	0.053		
	(0.216)	(0.123)		
Ancestor from WT × Share of autochtons, 1950	0.236	0.102		
	(0.228)	(0.127)		
Share of autochtons, 1950	0.169	0.053		
	(0.202)	(0.114)		
Ancestor from Kresy × Share of Polish speakers, 1900			0.039	0.041
			(0.178)	(0.095)
Ancestor from WT × Share of Polish speakers, 1900			0.168	0.041
			(0.175)	(0.089)
Share of Polish speakers, 1900			0.219	0.096
			(0.172)	(0.082)
Controls‡	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
Mean Dep. Var.	2.83	0.55	2.83	0.55
R^2	0.20	0.14	0.20	0.14
Observations	11,548	11,548	11,548	11,548

Notes: The table shows that the coefficient on Kresy ancestry does not vary significantly with the share of authochthons in the 1950 Census. This suggests that congestion is *not* an important part of the story. Regressions are run at the ancestor level; standard errors clustered by powiat. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, as well as ancestor level dummies.

Table 13: No heterogeneous effects w.r.t. origin diversity:

Dependent variable: as indicated in table header

	Degree (4 categories)							
Ancestor from Kresy	0.170***	0.178***	0.145***	0.164***	0.185***	0.162***	0.162***	0.147***
	(0.040)	(0.040)	(0.030)	(0.033)	(0.031)	(0.033)	(0.029)	(0.054)
1931 share Rom. Catholics	-0.016							
	(0.084)							
1931 share Rom. Catholics × Ancestor from Kresy	0.058							
	(0.130)							
1931 share Polish speakers		0.027						
		(0.096)						
1931 share Polish speakers × Ancestor from Kresy		0.024						
		(0.141)						
1931 share Russian speakers			3.489					
			(2.133)					
1931 share Russian speakers × Ancestor from Kresy			-2.767					
			(2.171)					
1931 literacy rate				-0.029				
				(0.147)				
1931 literacy rate × Ancestor from Kresy				0.057				
				(0.176)				
1931 urbanization rate					0.093			
					(0.080)			
1931 urbanization rate × Ancestor from Kresy					-0.136			
					(0.106)			
1921 literacy rate						0.047		
						(0.142)		
1921 literacy rate × Ancestor from Kresy						-0.063		
						(0.177)		
1921 literacy rate among Rom. Catholics							0.057	
							(0.152)	
1921 literacy rate among Rom. Catholics × Ancestor from Kresy							-0.044	
							(0.204)	
1921 share Rom. Catholics								0.098
								(0.109)
1921 share Rom. Catholics × Ancestor from Kresy								-0.202
								(0.170)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	2.83	2.83	2.83	2.86	2.87	2.86	2.86	2.86
R^2	0.27	0.27	0.27	0.27	0.28	0.27	0.27	0.27
Observations	11,534	11,547	11,547	9,766	8,614	9,744	9,744	9,744

Notes: The table shows that the coefficient on Kresy ancestry does not vary significantly with average characteristics of the population at the place of origin. This suggests that uprootedness effect is pretty universal and not dependent on pre-war socio-economic environment. Regressions are run at the ancestor level; standard errors clustered by individual respondents. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2.

Table 14: No heterogeneous effects w.r.t. soil quality at origin

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Degree (4 cat.)							
Ancestor from Kresy	0.186***	0.186***	0.161***	0.154***	0.184***	0.189***	0.177***	0.177***
	(0.032)	(0.033)	(0.043)	(0.042)	(0.036)	(0.035)	(0.032)	(0.032)
Ancestor from rural area	-0.178***	-0.177***	-0.183***	-0.180***	-0.181***	-0.183***	-0.181***	-0.187***
	(0.040)	(0.040)	(0.039)	(0.040)	(0.038)	(0.039)	(0.039)	(0.039)
Land suitability for wheat at origin	-0.000	-0.001						
	(0.001)	(0.002)						
Land suitability for wheat at origin X Ancestor from Kresy		0.001						
		(0.002)						
Annual temperature at origin			-0.020	0.012				
			(0.026)	(0.042)				
Annual temperature at origin X Ancestor from Kresy				-0.070				
				(0.047)				
Precip.-evatranspiration ration at origin					-0.080	-0.026		
					(0.129)	(0.167)		
Precip.-evatranspiration ration at origin X Ancestor from Kresy						-0.206		
						(0.242)		
Ruggedness at origin							0.002	0.022
							(0.027)	(0.031)
Ruggedness at origin X Ancestor from Kresy								-0.076
								(0.058)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	2.87	2.87	2.87	2.87	2.87	2.87	2.87	2.87
R^2	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28
Observations	8,794	8,794	8,794	8,794	8,794	8,794	8,794	8,794

Notes: The table shows that the coefficient on Kresy ancestry does not vary significantly with aoil quality at the place of origin. This suggests that uprootedness effect is pretty universal and not dependent on agricultural features at the location of origin. Regressions are run at the ancestor level; standard errors clustered by powiat. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, as well as ancestor level dummies.