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

Sascha O. Becker
University of Warwick

Irena Grosfeld
Paris School of Economics

Pauline Grosjean
UNSW

Nico Voigtländer
UCLA

Ekaterina Zhuravskaya
Paris School of Economics

This draft: 18 April 2018. Latest version here.

Abstract

We exploit a unique historical setting to study the long-run effects of forced migration on investment in education. After World War II, the Polish borders were redrawn, resulting in large-scale migration. Poles were forced to move from the *Kresy* territories in the East (taken over by the USSR) and were resettled mostly to the newly acquired *Western Territories*, from which Germans were expelled. We combine historical censuses with newly collected survey data to show that, while there were no pre-WWII differences in education, Poles with a family history of forced migration are significantly more educated today. Descendants of forced migrants have on average one extra year of schooling, driven by a higher propensity to finish secondary or higher education. This result holds when we restrict ancestral locations to a subsample around the former Kresy border, and when including fixed effects for the destination of migrants. Because Kresy migrants were of the same ethnicity and religion as other Poles, we can bypass these typical confounding factors in other cases of forced migration. Labor market competition with natives or selection of migrants are also unlikely to affect our results. Instead, survey evidence complements our results on education, suggesting that forced migration led to a shift in preferences, away from material possessions and towards investment in a mobile asset – 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 Harvard University, Simon Fraser University, 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 (grant no. ES/L011719/1). 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

Forced migration is a life-changing experience leaving deep scars in the memory of those who experience it. Does the experience affect also subsequent generations? In his autobiographical novel, Nobel Laureate Amoz Oz writes "*It was always like that with Jewish families: they believed that education was an investment for the future, the only thing that no one can [...] take away from your children, even if, God forbid, there's another war, ... another migration*." The idea that forced migration may affect preferences for education has not only been made for Jews, but more generally, in the academic literature, e.g. by Brenner and Kiefer (1981): *"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."* However, this "uprootedness" hypothesis has proved controversial. Even for the most prominent case – that of the Jews – Botticini and Eckstein (2012) have convincingly challenged the idea that expulsion and discrimination are behind their educational lead.[1] It is notoriously difficult to convincingly identify the link between forced migration and investment in education. Forced migrants typically differ from locals along other socio-economic and cultural characteristics such as ethnicity, language, or religion. In addition, they often face labor market competition with locals, affecting their educational choices.

In this paper, we explore a unique historical setting that allows us to study the effect of forced migration on human capital investment, absent the typical confounding factors. 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 about 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. Poles were forced to leave from the Kresy territories and the vast majority of them resettled in the largely-emptied Western Territories. We can thus shed light on the long-run effects of uprootedness, by comparing the descendants of Poles that were forced to migrate with all other Poles – of the same ethnicity, language, and religion.

To study the long-run effects of forced migration, we cooperated with the large-scale Polish social survey 'Diagnoza' to include questions about respondents' ancestors from Kresy in their 2015 wave. Among the almost 30,000 respondents, more than 11% had ancestors from Kresy. We find that descendants of forced migrants have significantly higher education today, as compared to all other Poles. The education advantage of descendants of forced migrants is quantitatively

[1] Instead, Jewish preferences for education are explained by religious motives – Jewish boys were expected to read the Torah. These preferences trace back to the time of the fall of the second temple in Jerusalem (in 70 CE), before Jews started to be repeatedly expelled.

important: they have on average one extra year of schooling, driven by a higher propensity to finish secondary or higher education. Importantly, education levels of forced migrants were *not* higher before WWII. Figure 2 illustrates our main results: Before WWII, when Poland consisted of the Kresy territories and Central Poland (CP), Poles in the former had if anything lower literacy rates. Today, Poland consists of CP and Western Territories (WT). The figure shows that in both regions, average education is similar.[2] However, people with ancestors from Kresy have significantly higher education in both areas. Since Kresy migrants were forced to leave their homeland, selection is unlikely to confound our results. This is further supported when we present the educational advantage of Kresy descendants by birth cohort. Figure 3 shows that those forced migrants who had likely finished school by the time they were expelled from Kresy (i.e., the cohort born before 1930) do not differ from other Poles in terms of their education. For younger cohorts, we find a significant education advantage for Kresy descendants, even for those born two generations after their ancestors had been expelled.[3]

We complement our main results by an additional Ancestry Survey that we conducted in 2016 in the Western Territories, where the majority of Kresy migrants were transported after WWII. We collected a representative sample of about 4,000 respondents, who recalled the detailed location of almost 12,000 ancestors from all over Poland, as well as from Kresy. This dataset allows us to compare descendants of forced migrants from Kresy with those whose ancestors were voluntary migrants from Central Poland after WWII, and with Poles who had already lived in WT before the war (autochthons).[4] This historical experiment allows us to compare forced migrants with voluntary migrants within their destination locations in the Western Territories. We find that descendants of migrants from Kresy are the most educated, followed by descendants of voluntary migrants. Descendants of autochthons are the least educated group in Poland's Western Territories today. These results suggest that forced migration has an even stronger long-term impact on investment in human capital than voluntary migration.

The detailed ancestor locations reported in our Ancestry Survey also allow us to confirm our main results in a particularly restrictive border sample analysis. We restrict the sample to people whose ancestors in 1939 lived within less than 150km of the Kresy border. At the same time, we include municipality fixed effects for the location of today's survey respondents. We find that

[2]We use the share of people with a secondary degree in this analysis because it is comparable to literacy rates in 1921 in terms of its nationwide average. In our empirical analysis, we use years of schooling as the main outcome variable.

[3]The results shown in Figure 3 use county fixed effects, which absorb potential local differences in the education system or in the labor market environment.

[4]As discussed in detail in Section 2.2.2, the Polish administration encouraged people from Central Poland to voluntarily migrate to the Western Territories. This raises the possibility of self-selection. We perform a host of empirical tests and conclude that self-selection among voluntary migrants is unlikely to affect the results of our Ancestry Survey.

among respondents who live in the same town or village today, those with ancestors from a few km to the East of the Kresy border have significantly higher education today than those whose ancestors lived in Central Poland, a few km to the West of the Kresy border. We also show that our results from the Ancestry Survey are unlikely to be driven by selection of voluntary migrants from Central Poland to Western Territories after WWII. In particular, if voluntary migrants from CP were negatively selected, this would yield a less educated control group in WT today, biasing the coefficient on Kresy origin upward. To address this possibility, we provide several tests. First, as shown in Figure 2, literacy in 1921 was if anything higher in the origin counties of voluntary migrants (CP), as compared to forced migrants (Kresy). Second, descendants of voluntary migrants in Western Territories today are on average more educated than their cousins whose ancestors remained 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 Central Poland). These results are easier to reconcile with *positive* selection of voluntary migrants from Central Poland, which implies that, if anything, our results underestimate the education advantage of forced migrants. We also show that our results hold equally for urban and rural respondent locations 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.

In order to interpret our results, we examine several possible explanations. These include differential returns to human capital for Kresy migrants, congested labor markets, and differential out-migration. None of these account for the observed education premium of Kresy descendants. Instead, the most 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 human capital. This is further corroborated by survey evidence showing that descendants of forced migrants value material goods less, while having a stronger aspiration for education of their children. They also possess fewer assets, relative to the number of assets they could afford.

Our work is related to a large literature that 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,[5] and socio-economic effects on migrants themselves and on their descendants.[6] Migration can be voluntary or forced.[7] Key drivers of forced migration are

[5]See the comprehensive discussions in Borjas (2014), Card and Peri (2016) and Dustmann, Schönberg, and Stuhler (2016). For long-run effects of migrants at their destinations see Hornung (2014), Peters (2017), and Murard and Sakalli (2018).

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

[7]The literature on the effects of forced migration is surveyed in Ruiz and Vargas-Silva (2013). For example, Card (1990), Borjas and Monras (2017), Bharadwaj, Khwaja, and Mian (2015), and Braun and Mahmoud (2014) use forced

natural disasters, international wars, and civil wars.[8] Finally, a large body of work has examined the effects of *voluntary* migration, for instance in the context of the Age of Mass Migration to the US (Abramitzky, Boustan, and Eriksson, 2014; Sequeira, Nunn, and Qian, 2017).[9]

Our focus is on the effects of forced migration after WWII over the very long term, in the generations of children, grandchildren, and great-grandchildren of adult expellees. In the context of forced migration due to WWII, two related papers are Bauer, Braun, and Kvasnicka (2013) and Sarvimäki, Uusitalo, and Jäntti (2016). Bauer et al. (2013) study the economic integration of Germans expelled from Poland's Western Territories into the West German economy. They find that migrant children tend to acquire more education than their native peers. One key element of this mechanism is congestion: former farming families had to look for work outside agriculture because agricultural land in West Germany was already held by native Germans. We show below that this mechanism is unlikely to be at play in the largely emptied Western Territories. Sarvimäki et al. (2016) study the forced migration of 11% of the Finnish population after the Soviet invasion in 1939. Their focus is on income of those adults who were expelled themselves, but not on outcomes for later generations.

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) were resettled into the largely empty ex-German Western Territories. This unique setting allows us to bypass common confounding factors associated with forced migration, such as different ethnicity, language, or religion. Second, we break new ground be studying the long-run effects of forced migration on the descendants of migrants over several generations. Third, our Ancestor Survey allows us to compare descendants of forced migrants from Kresy to those of voluntary migrants from Central Poland. Finally, our results lend strong support to the prominent, yet controversial, hypothesis that uprootedness leads to human capital investment.

The remainder of the paper is organized as follows. Section 2 provides historical background, Section 3 describes the data. Section 4 shows the main results using survey data for all of Poland. Section 5 shows results for Western Territories using data from our own Ancestry Survey. Section

migration to identify the effect of migration on economic outcomes at the destination.

[8]Many papers examine relatively short-run effects of natural disasters. For instance, Sacerdote (2012) looks at the effects on test scores of students displaced from New Orleans after Hurricane Katrina. Nakamura, Sigurdsson, and Steinsson (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. Jacob (2004) and Chyn (2017) exploit exogenous variation in mobility caused by public housing demolitions in Chicago. While both papers do not find effects on educational attainment, displaced children have better labour market outcomes as adults.

[9]Bazzi, Gaduh, Rothenberg, and Wong (2016) study lottery-driven variation in voluntary migration during peace times in Indonesia to show that farmers are more productive in destination locations with agroclimatic endowments similar to where they come from.

6 looks at mechanisms. Section 7 concludes.

2 Historical Background

2.1 The change of Poland's borders

At the end of World War I, in 1918, the independent Polish state was recreated as the Second Polish Republic (SPR) after having been split for over a century among the three European super powers: the Prussian (German) Empire, the Russian Empire, and the Austro-Hungarian (Habsburg) Empire. The Second Polish Republic ceased to exist as an independent state with the Nazi and Soviet invasions of Poland in September 1939 following the Molotov–Ribbentrop Pact, which divided the SPR between the Soviet Union and Nazi Germany. At the end of Word War II, an independent Poland reemerged within new borders that 'moved Poland to the West' (c.f. Davies, 1981). These new borders were established in accordance with the decisions taken during the Tehran, Yalta and Potsdam Conferences. Poland gained the former German territories of Silesia, Pomerania and East Prussia, called by the communist propaganda "Recovered territories" and later "Western Territories" (WT) but lost the Eastern Borderlands, known as *Kresy*, located to the east of the Curzon line. Figure 1 illustrates the change in Polish borders: the shaded pink area depicts the Second Polish Republic (i.e., Poland between WWI and WWII), whereas the post-WWII borders are shown with thick red line.[10]

Henceforth, we refer to the part of Poland that belonged to the Second Polish Republic before WWII and continued to be Polish state after WWII as 'Central Poland.' Thus, the territory of Poland before WWII was comprised of Central Poland and Kresy (Eastern Borderlands lost to the USSR), whereas Polish territory after WWII is comprised of Central Poland (CP) and Western Territories (newly acquired from Germany). The Kresy territory was divided among the Soviet Republics of Lithuania, Belorussia, and Ukraine.

According to the 1931 Polish Census – the last census of the SPR – the total population of Poland was 31.7 million people one third of whom lived in Kresy. Between the two world wars, the Central Poland part of the SPR was relatively homogeneous: Poles constituted 85%, Jews 10%, Germans 3%, and Ukrainians 2% of total population. In contrast, Kresy had a heterogenous population: only 35% of the Kresy population were Poles; the rest were Ukrainians (39%), Jews (14%), Belorussians (8%), Germans (2%), Russians (1%), and Lithuanians (1%).

Before WWII, according to the 1939 German census, 8.8 million people lived in areas that

[10]The Eastern border of the Second Polish Republic was established at the signature of the peace treaty in Riga which marked the end of the Soviet-Polish war of 1919-1921. The borders of the Second Polish Republic around Silesia and East Prussia were adjusted as a result of several referenda in 1920-1922. Throughout the analysis and on the map, we consider the final SPR border as of 1922.

after WWII became the Polish Western Territories. Nearly all of them declared to be 'Germans'.[11] The population shares in 1939 are estimated to be: 88% Germans, 10% Poles, and about 1 % Jews (Dziewanowski, 1977).

2.2 Post-WWII Mass Population Movements

As a result of the change in borders after WWII, mass population movements occurred. At the end of WWII, an estimated 2.5-3.4 million Germans and 1 million Poles were located in Western Territories (Dziewanowski, 1977). Germans were expelled from Western Territories and many fled as the Red Army was advancing westward even before the expulsion took place. Germans left behind empty houses, land and factories and had to resettle in Germany to the west of the Oder-Neisse line. Importantly, Polish and Soviet authorities had agreed on a mass population exchange program following the change in the borders, according to which Poles from Kresy, which fell under the control of Soviet Republics of Ukraine, Belorussia and Lithuania were forced to resettle within the new borders of Poland, while Ukrainians, Belorussians, and Lithuanians had to leave Poland and resettle in the USSR. These mass movements of people started in 1944. The original plan was to complete them within two years. In reality, they were largely completed by 1948.

The population exchange agreements were signed between the so-called Polish Committee of National Liberation, a puppet provisional government of Poland controlled by the Soviet Union, and the governments of the three Soviet Republics (Ciesielski, 1999). The official language of these agreements did not explicitly specify that the ethnic groups in question were expelled from the two respective sides of the Curzon line, however, historians agree (see, e.g., Davies, 1981; Kersten, 1986) that the members of these groups had no viable alternative but to move, especially, if they did not inter-marry with the title ethnicities of each respective state.[12] The main reason for this was the fear of ethnic persecution: ranging from direct terror (e.g., for Kresy Poles, the most relevant being the memory of Stalin's war-time deportation of Poles to Siberia and Central Asia as well as massacres of Poles in Volhynia and Eastern Galicia perpetrated by Ukrainian insurgents[13]) to professional and educational discrimination (as each of the Soviet republics as well as Poland took serious steps to become mono-ethnic states) (e.g., Kochanowski, 2015).

2.2.1 Forced Migration from Kresy Territories

By 1950, 2.1 million Poles were forced to move from the territories that Poland lost to locations within the new borders of the Polish State. The Polish State Repatriation Bureau tried to ensure

[11] German sources (Wirtschaft und Statistik, 1956) use the somewhat odd wording 'Deutschtum' = 'Germanness'. After 1911, mother tongue was no longer surveyed in schools and population censuses (Cinnirella and Schueler (2018)).

[12] Note that inter-marriage rates were particularly low between Poles and Ukrainians.

[13] See Magocsi (1996); Snyder (2003) on the Volhynian massacres.

an orderly movement of Poles from Kresy directly to Western Territories, which was hard to implement in reality because of the war-related devastation, in part related to the war destruction of infrastructure and the lack of adequate transport. De facto close to a quarter of Kresy migrants settled in Central Poland many of whom had family ties there before the war. The aim of the Polish authorities was to resettle Kresy deportees in those places in Western Territories that had soil and climatic conditions most closely resembling the conditions at the origin locations, which in practice meant that trains brought people to the Western Territories from Kresy along the same latitude. Kresy Poles had to leave most but not all of their possessions behind. Repatriation Bureau statistics reported that resettlers from Kresy brought with them to WT 71 213 horses, 139 272 cattle, 34 991 pigs, 50 586 sheep and 168 960 tools for agriculture.

Even though historians agree that Poles were forced to move from Kresy, not everybody left during the post-war population exchange. Importantly, most of these people left to Poland during the so-called second repatriation of Poles from the USSR in 1955-1959. In 1945-1946, authorities in the Lithuanian and Belorussian SSR were concerned that agricultural production could be halted by a drop in agricultural labor and tried to prevent Poles in rural areas from leaving. In contrast, Ukrainian authorities did not attempt to prevent rural Poles from leaving due to the very high levels of animosity between Poles and Ukrainians (e.g., Ciesielski, 1999). In all three Soviet republics, pressure to leave on urban Polish population was very high.

2.2.2 Voluntary Migration from Central Poland

Land, housing, infrastructure, and capital stock were abundant after the expulsion of German population from Western Territories, as before the war these territories were developed and densely populated. This made Western Territories an attractive destination for voluntary migrants from Central Poland, who sought a better fortune than in their homeland. The flow of migrants from Central Poland started as early as the spring 1945. Some of this voluntary migration was spontaneous (mostly from the neighbouring Polish areas, sometimes on foot, or by horse cars and trucks), some was triggered by an advertising campaign organised by the Polish authorities that promoted a move to Western Territories in order to populate the newly acquired land as quickly as possible. The campaign advertised Western Territories as the land of abundant resources. In the online appendix Figure A.1 we present an example of a poster dating back to this campaign. As a result of this campaign, voluntary migrants came to Western Territories from all over Central Poland, many of whom travelled long distance by train (Halicka, 2013; Zaremba, 2012).

2.2.3 The timing of arrival of Kresy and CP migrants

Upon arrival to Western Territories, Poles (irrespective of whether they came from Kresy or Central Poland) were allocated capital that expelled Germans had to leave behind; in rural areas, this

primarily meant houses, land plots, and agricultural machines, in urban areas – apartments, townhouses, shops, office buildings.[14] The large capital stock, both in urban and rural areas, such as German town factories and large farms, was nationalised/collectivised by the Polish authorities and incoming settlers got jobs in these productive establishments.

Importantly, forced Kresy migrants and voluntary migrants from Central Poland arrived to Western Territories at the same time. The Ministry of Recovered Territories collected statistics on the rates of arrival of migrants by month during 1946 and 1947. We present these data in Figure A.7. Panel A portrays the stock of migrants from both Central Poland and Kresy who arrived to WT by a particular month and Panel B shows the share of forced Kresy migrants in this stock of total arrivals. By the end of 1947 the Kresy migration of the first repatriation wave came to an end. Moreover, some of the Kresy migrants who first arrived to WT subsequently left (in particular, this was the case for the Polish Jews who survived the war).

2.3 The Origin of Poles in WT and CP after the Mass Population Movements

The first reliable post-WWII population census in Poland took place in 1950.[15] In addition to statistics normally collected during censuses, it provided the information about the mass movements of Polish population by asking about the place of residence on 1 September 1939. In particular, respondents indicated if they lived within the post-war Polish borders, and if so, in which region (voivodeship). If in 1939, respondents lived outside post-war Poland, they had to indicate the modern country within borders of which they resided in 1939. Thus, forced Kresy migrants indicated that they lived in the USSR before the war. Table 1 reports aggregate statistics about the origin of Polish population separately in Western Territories and in Central Poland from 1950 census. Of the total 24.6 million of Polish population, 23% (5.6m) lived in Western Territories. Within Western Territories, about 50% (2.8m) came from Central Poland, 28%(1.6m) came from Kresy, and 20% (1.1m) were autochtons, i.e., Poles who lived in Western Territories when these territories belonged to Germany before the war. The remaining 2.7% came from other countries, mostly from France. Within Central Poland, 96.5% (18.4m) of the population had CP origin and only 3% (583 thousand) came from Kresy. Very few of CP's population came from Western Territories or from abroad, other than USSR (0.1% and 0.3%, respectively). In Western Territories, the vast majority of autochtons were concentrated in a few counties, so that the vast majority of WT's inhabitants in most localities were either voluntary migrants from CP or forced migrants from Kresy.

Importantly, the post-WWII Poland was largely an ethnically and religiously homogeneous

[14] In the early months of the population exchange, occasionally Poles had arrived before Germans left; in those cases, the two groups co-existed for a short period of time.

[15] There was another census in 1946. However, this census is widely considered unreliable because population movements were still ongoing.

country, composed of ethnic Poles of Roman Catholic faith that differed only in their pre-WWII origin. Western Territories, as the rest of Poland, were also ethnically and religiously homogenous: according to 1950 census, 96% of WT's total 5.6m population were Poles, i.e., Roman Catholics and Polish native speakers. The rest of WT's population in 1950 were Jews, most of whom subsequently left (fewer than 300,000 Polish Jews survived WWII and some came to WT right after the war), and Ukrainians, who were forced migrants from CP to WT (during the so-called "Operation Vistula").[16]

Overall, post-war Europe had witnessed a lot of forced movements of people. Yet, migration of Poles to Western Territories was a unique experiment allowing us to separate the effect of being forced to move from the effect of the migration itself or of the challenges of integration at the destination locations. The reason for this is that the natives (Germans) were expelled from these territories and both voluntary and forced migrants arrived at the same time, were treated similarly, and belonged to the same ethnic and religious group.

3 Data

We use numerous data sets for modern-day and historical Poland in our analysis. In particular, we combine individual-level survey data from two surveys with locality-level data from several historical censuses. The two surveys capture modern-day outcomes, such as educational attainment and attitudes, and ask questions about the history of migration of respondents' ancestors as a result of the change in Polish borders. We complement these data with data from historical censuses that describe the population that lived in Poland before and after WWII. We use two censuses conducted in inter-war Poland (SPR) in 1921 and 1931 and two censuses from before WWI, i.e., the German Empire Census of 1900 and the Russian Empire Census of 1897, to describe the situation before the move of Polish borders. We rely on the post-WWII 1950 Polish Census for information on the situation at the end of the mass population transfers and the aggregate statistics on the movements of population. We describe each of these data sources in turn.

3.1 Modern-day survey data

3.1.1 Diagnoza Survey

The Diagnoza (Social Diagnosis) 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.[17] We commissioned adding several questions to the 2015 wave which inquired whether any of the ancestors of the respondent came

[16] See http://en.wikipedia.org/wiki/Operation_Vistula.

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

from Kresy and if so where from exactly. The main 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)?" If the answer was yes, respondents were asked to indicate up to three localities where their relatives lived in Kresy in the summer of 1939. The 2015 wave has ca. 30,000 observations and allows us to compare education and other outcomes for respondents with *any* ancestors from Kresy to those without *any* ancestors from Kresy. Our primary interest is investment in human capital, thus, we focus on the education of respondents. In the survey, it is given by education degrees in 4 categories (primary or less; vocational and less than secondary; secondary; more than secondary). We use this information in the following ways: (1) take the categorical variable as an outcome; (2) use a dummy for (at least) secondary education (as compared to less than completed secondary education) and (3) use a dummy for (at least) completed tertiary education. In addition, the survey contains the question on the total number of years that respondents spent in any educational institution. We also use a number of questions about the attitudes of respondents toward the education of their children and toward the accumulation of material wealth.

3.1.2 Our own Ancestry Survey

In 2016, we conducted our own survey in Western Territories 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).[18] 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 adult generation was the respondent's parents, we asked whether the mother and the father of the respondent lived on September 1, 1939. If the generation of the youngest adults in the family in 1939 was the respondent's grand-parents, we asked about the exact place of residence of each of the four grandparents. Similarly, if the generation of the youngest adults in 1939 was the respondent's great-grandparents, we solicited information about place of living for all eight great-grandparents.[19] Overall, 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 the generation of youngest adults in the family was the great-grandparents. In some cases, respondents might not recall the exact place of residence of some ancestors before WWII and would merely indicate a broad area. Overall, we

[18] 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 neighbourhood until they found a respondent with Kresy origins. Our regressions use appropriate sampling weights.

[19] A handful of respondents were themselves over 18 year old in 1939. We drop them from the analysis, as they are too few to constitute a meaningful group of their own.

were able to identify and geo-code the *exact* place of residence for 11,928 out of 13,223 ancestors.[20] As the main variable of interest, we use the share of ancestors of respondents, who come from Kresy.

We report summary statistics for the variables that measure the origin of respondents in Table A.1 and for the variables that measure the respondents' education in Table A.2 in the online appendix. Each of these tables summarises data in both datasets.

3.2 Historical censuses

Post-WWII Polish Census 1950 – The Polish Census in 1950 contains information on population movements. It asked which region (voivodeship) in post-WWII Poland or which post-WWII country people lived in before the war. In Western Territories, this information is available by county (powiat) of residence. In the rest of Poland, the information is available by region (voivodeship).

Inter-War Polish Censuses: 1921 and 1931 – We use two censuses conducted in the Second Polish Republic. The census closest to WWII was conducted in 1931, it gives information on literacy rates, languages, and religion by locality, but without cross-tabulations of the data. The 1921 Census, in contrast, has literacy rates by religious denomination, allowing us to separately measure the literacy rates among Roman Catholics and among people with other religious denominations. As post-WWII Poland is homogeneously Roman Catholic and pre-WWII Poland was rather diverse, especially in the Kresy area, the pre-WWII literacy rates of Roman Catholics are of a particular interest to us – they are a closest proxy for the literacy of ethnic Poles before WWII.[21]

Pre-WWI: Russian Empire Census 1897 – The 1897 Census of the Russian Empire (Troynitsky, 1899) provides information on literacy rates in Russian language and in the native language for each native language in the empire. For our purposes, we extract the literacy of native Polish speakers in native language.

Pre-WWI: German Empire Census 1900 – We use the share of Polish speakers in 1900 across localities in Western territories to proxy for the autochthon WT population.

3.3 Benchmarking Survey Data against Historical Census Data

One potential worry in using survey data about ancestral origin is the recall bias. For example, more educated respondents may have more information on the location of origin of their ancestors. This is a particularly important issue in the Diagnoza Survey, which only asks about Kresy origin. This could lead to a potential bias in estimating the effect of forced migration from Kresy on today's education. In our own Ancestry Survey, we ask about origins of *all* ancestors, which

[20] Figure A.6 displays the origin of ancestors in our Ancestor Survey.

[21] In SPR, only Poles were Roman Catholics, other groups had other religious affiliations: mostly, Orthodox Christians, Greek Catholics, and Jews.

makes this source of bias less likely. De facto, the majority of respondents knew the exact name of the locality of origin of their ancestors, and not only the broad region of origin, which highlights the importance of the mass population movements in the family histories of Poles.[22] The fact that 11,928 out of 13,223 ancestors gave the exact location of origin is an important indication that respondents had no problem recalling the origin of their ancestors.

While we have no way to confirm the accuracy of information of individual respondents, we can benchmark the movements implied by the difference between the location of residence of respondents and their ancestors in our surveys with the data on mass population movements from 1950 Census. Of course, this exercise cannot take into account the mobility of people after 1950. Nonetheless, we find that the ancestry information from the two surveys lines up quite well with the 1950 Census data.

In Figure A.4, we compare the results of the Diagnoza survey with the 1950 census. The left panel 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 regions (voivodeships) in all of Poland in 1950 because both the Diagnoza survey and the Census cover all of Polish post-war territory and the data in the 1950 Census is available at the regional level for the whole of Poland. The shares of population with Kresy origin in 1950 and of Diagnoza respondents with Kresy origin in 2015 line up very well for most regions.[23] Information about the pre-WWII origin of population in Western Territories is available at a more disaggregated level of counties (powiaty) in 1950 Census. We use it to repeat the consistency check on Diagnoza data in the right panel of Figure A.4. The fit in this county-level exercise is bound to be less precise for two reasons. First, the post 1950 mobility across county boundaries is higher than across regional 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, the right panel of the Figure shows a tight relationship, also at the county level.

Figure A.5 repeats this exercise using our 2016 Ancestry Survey in combination with the 1950 census. Since our Ancestry Survey asks about origin locations of all ancestors including those ancestors who came to WT from Central Poland and not only from Kresy, we can map the origin

[22] In our survey, we were able to monitor the interview process and we were impressed by how survey respondents engaged with the questionnaire. Many respondents were fascinated by our questions about their ancestry to the extent that they made every effort to respond accurately. Many even checked family archives to make sure that they gave the most precise answer possible. Some even called back to tell us their family stories.

[23] There are a few exceptions. For instance, Warsaw is considerably below the regression line, suggesting that while in 1950 few people of Kresy origin lived there because the majority moved straight to Western Territories, in 2015 the share of Warsaw survey respondents with Kresy ancestors is considerably larger, likely because the capital city attracts people from different parts of the country, including many descendants of Kresy people.

location data from the Ancestry Survey to the 16 origin areas that are covered in the 1950 Census.[24] This comparison is displayed on the left panel of Figure A.5. The right panel presents the share of people of Kresy origin by country from our Ancestry Survey and from the 1950 Census. Both panels show a strong positive relationship between the data in the two data sources, supporting the reliability of our Ancestry Survey. In sum, the benchmarking exercises make us confident that respondents in the Diagnoza and the Ancestry Survey gave reasonable answers to the questions about their ancestral places of origin.

4 Forced Migration from Kresy and Education

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

$$Y_i = \beta \, Kresy_i + \phi' \, \mathbf{X}_i + \eta_{Locality_{(i)}} + \varepsilon_i, \tag{1}$$

where Y_i denotes different outcomes of respondent i, such as measures of i's education and attitudes. In the Diagnoza survey (which only provides information on whether respondents have any ancestor from Kresy), $Kresy_i$ is a dummy variable that takes the value of one if *any* ancestor was 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: gender, age, age^2, dummies for six age groups, as well as indicators for whether the respondent lives in a rural area, in an urban county, and in Western Territories.[25] Finally, $\eta_{Locality_{(i)}}$ represents fixed effects for the locality of respondents' residence. This controls for the local socio-economic environment, such as labor market conditions. In particular, we use county (powiat) or municipality (gmina) fixed effects. Because in Diagnoza several respondents may come from the same household, we cluster the error term at the household level.

4.1 Diagnoza Survey Results – Education

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 Poland and has a large number of respondents. The downside is that it only includes information on whether respondents had *any* ancestors from Kresy territories, but no indication how many of them

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

[25] Note that by using six age group dummies together with age and age^2, we allow for non-linear effects of age *within* age groups.

and no indication of origin of ancestors from regions other than Kresy.[26] We first show results for education as the outcome variable and then turn to values.

Table 2 presents the comparison of education levels of respondents with and without Kresy origin using the Diagnoza survey. It shows that individuals whose ancestors were expelled from Kresy territories have significantly higher levels of education today. Panel A presents our main results for the common outcome variable 'years of education.' The results for the full sample, with approximately 29,000 respondents (out of whom about 3,300 had Kresy ancestors) show that Kresy ancestry is associated with 0.85 additional years of schooling, relative to an average of 11.7 years. Next, columns 2 and 3 restrict the sample to respondents in rural and urban areas, respectively. The coefficient sizes are similar in both subsamples. This makes it unlikely that our results are confounded by the availability of agricultural land to forced migrants. In addition, the 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. Note that all regressions control for county fixed effects. Consequently, regional differences in the destinations of migrants from different origins (and thus the place of residence of their descendants) cannot drive our results.

In Panel B we use an alternative outcome variable for education – a categorical variable that indicates the level of degree (1 – primary or less; 2 – vocational or incomplete secondary; 3 – secondary; 4 – More than secondary).[27] This variable confirms the previous results: Individuals with ancestors from Kresy score about 0.25 points higher on the 4-category education variable (which has a mean of 2.5 and standard deviation of 1). The coefficients are also very similar across the subsamples in columns 2-5. Finally, in Panels C and D the dependent variable is an indicator for secondary and higher education, respectively. We find that descendants of Kresy migrants are 11 percentage points more likely to finish secondary education (relative to a mean of 0.5), and 9 percentage points more likely to graduate from college (relative to a mean of 0.2). Thus, in *relative* terms, the association between Kresy origin and education is strongest for higher education.

[26]For respondents who state that they had any ancestors from Kresy, the 2015 Diagnoza asked them to provide up to three locations of origin for their ancestors. We geocoded these places. What is missing in Diagnoza is the exact ancestor who lived in these locations (mother, father, grandmother, etc). We collect this information in our Ancestory Survey (see below).

[27]Using this alternative measure can address the following factors that may render years of schooling noisy: 1) year repetitions if a student performs badly; 2) changes in regulations for the minimum years of schooling and for the number of years needed for a degree; 3) the number of years studied to obtain a post-graduate degree may indicate neither the quality nor of the quantity of education.

Kresy Border Sample

Could the higher educational attainment by Kresy descendants today be driven by pre-WWII differences? For example, culture (and in particular, attitudes towards education) may have been different in the eastern territories of pre-WWII Poland, even if literacy rates were very similar (see Figure 2). To shed more light on this possibility, we restrict the sample to an area of less than 150 km on each side of the Kresy-Poland border. This arguably provides a culturally more homogenous area. At the same time, we face a challenge in constructing this sample. Because people from Kresy were expelled, there are no Diagnoza respondents living in the eastern part of the border sample today. We thus use information on the location of ancestors that is provided in Diagnoza to identify respondents with ancestors within less than 150 km to the east of the Kresy border.[28] As for the area within 150 km to the *west* of the Kresy border (i.e., in today's Poland), we assume that respondents without Kresy ancestors who live there today have also family roots in the area.[29]

We begin by checking whether there were pre-existing differences in education along the Kresy border. The left panel of Figure 4 shows that this is not the case: literacy among Poles (identified by their Roman Catholic religion in the 1921 census) was very similar to the east and west of the Kresy border. There is also no significant trend in distance on either side of the border. Next, the right panel of Figure 4 shows education levels in 2015 – west of the border for respondents (without Kresy ancestors) who live within 150 km of the Kresy border today, and east of the border, for respondents with ancestors from Kresy who lived within 150 km. The figure confirms that ancestors from Kresy have substantially higher education levels, even in a subset of individuals with ancestors from close to the Kresy border.

Column 6 in Table 2 presents our border sample regression results. Note first that in the Kresy border sample, the means of the dependent variables are very similar to the overall sample means in Diagnoza (column 1). This renders the results directly comparable. In addition to the standard controls, column 6 also includes a quadratic polynomial in latitude and longitude to capture unobservables that may vary around the Kresy border (Dell, 2010).[30] Turning to the results, we find strong and significant coefficients for Kresy ancestors for all education measures. Coefficients are even larger than those in the main sample. One reason for the difference could be that we now use only those Kresy-origin respondents who also remember the locations where their ancestors

[28] Whenever a respondent gave the location of more than one Kresy ancestor (see footnote 26), we use the minimum distance to the Kresy border among all ancestors.

[29] Note that most of the descendants of forced migrants live in the Western Territories, far away from the historical Kresy border. Our border analysis compares these to individuals who still live close to the Kresy border (to its west). That is, our border sample compares individuals who potentially live far apart today, although their ancestors were likely living close to each other.

[30] Following the argument in Gelman and Imbens (2014) that cubic and higher-order polynomials can yield misleading estimates, we use a second order polynomial.

lived in 1939. This may be a subsample with particularly vivid memories of the forced migration experience, augmenting the long-run effects on education. Before we move to the next analysis, we note that our border sample analysis has a number of shortcomings – most importantly that the individuals with Kresy roots now largely live in Western Territories, where the socioeconomic environment also differs from the area close to the Kresy border. We show in Sections 5.1 and 5.2 below that our results are unlikely to be driven by these differences, by focusing exclusively on descendants of migrants who live in Western Territories today.

Cohort Analysis

Table 3 presents our results for different birth cohorts. Column 1 includes only individuals born before 1930 – the oldest respondents in the Diagnoza survey. Among this group, respondents with Kresy ancestors are likely to be Kresy migrants themselves. The pre-1930 cohort was 16 or older in 1945 and thus would already have finished their secondary education (if they had any). Panel B shows a very small and insignificant coefficient on Kresy ancestry for secondary education. In other words, in the cohorts that was old enough to have finished secondary education, the proportion with a secondary degree is very similar for individuals expelled from Kresy and other Poles. This implies that our results are unlikely to be driven by pre-existing differences in migrants' education. Panel C reports the results for higher education. We find a negative (but insignificant) coefficient on Kresy ancestry. One interpretation of this finding is that young adults who were expelled from Kresy had to organize their lives in their new homes, leaving little scope to pursue further education.

Columns 2-8 in Table 2 focus on younger cohorts, i.e., those that had not finished schooling by 1945 or had not even been born. The coefficient on Kresy ancestry 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. We visualize this in Figure 3, which uses ln(years of education) as dependent variable and therefore reports the relative coefficient sizes that can be directly compared across cohorts. Among the 1930 birth cohort (i.e., school-age children in 1945), people with Kresy origin have about 12% higher years of schooling.[31] For later birth cohorts, the coefficient size declines continuously. 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,

[31] Historical accounts suggests that the supply of schools was well organized as early as 1946, even in the Western Territories. There was a great effort to ensure good educational opportunities (free and obligatory for the primary schools). The first schools in WT were established relying on the initiative of individual teachers. Very quickly, however, the communist authorities created special institutions to help develop a unified educational system in WT and in CP (Online PWN Encyclopedia, accessed 28 March 2018).

Huffman, and Sunde, 2012).[32]

Income and Labor Market Outcomes

In Table 4, we verify that higher education of descendants of Kresy migrants translates into better labor market outcomes. Column 1 shows that respondents with ancestors from Kresy have about 10% higher incomes. Column 2 suggests that the higher income is at least partially driven by their higher education – once we control for education, the coefficient on Kresy becomes smaller and statistically insignificant. Columns 3-6 show that people with Kresy ancestors are more likely to have white collar occupations and to become entrepreneurs. At the same time, they are less likely to be unemployed (cols 7-8). These results remain statistically significant even after we control for education, but the coefficients on Kresy origin become smaller in magnitude.

4.2 Diagnoza Survey Results – Values

What explains the higher education among people with Kresy ancestors? In Table 5, we take a first pass at the issue by examining values towards education and material possessions. In the first two columns, the outcome variable is a categorical variable that reflects respondents' aspiration for the education of their children.[33] People with Kresy ancestors score 0.15 higher, relative to a mean of 4.3 and a standard deviation of 1.1. Remarkably, this result holds even after we control for the respondent's own education (column 2). In columns 3-6 we examine answers to the question: "What is the main condition for success in life?" In particular, we consider as dependent variables dummies which indicate whether respondents agreed that the possession of material goods, money, or freedom constitute the main condition for a success in life.[34] Columns 3 and 4 show that respondents with Kresy ancestors are significantly less likely to believe that material goods determine a successful life. Columns 5 and 6 show that descendants of Kresy migrants value freedom more than the rest of the Polish population. In columns 7 and 8, we explore whether the lower value placed on material wealth among descendants of Kresy migrants translates into actual choices about accumulating assets. Diagnoza asks about the possession of 20 different assets (e.g., vacation house, ebook, home theatre, boat, etc.). For those assets *not* possessed, respondents were asked if this was for financial reasons. The dependent variable in columns 7-8 is the number assets not owned for non-financial reasons (i.e., assets that the household could afford, but does not own),

[32]Note that among the 1990s birth cohort, many respondents were still in schooling/university age by the time of the Diagnoza survey in 2015. Thus, the 1990s coefficient needs to be interpreted with caution.

[33]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.

[34]In particular, we deem as positive the following answers: "definitely yes," "yes," or "rather yes." The other potential answers were "neither yes nor no," "rather not," "no," and "definitely no."

divided by the number of all non-possessed assets.[35] Consistent with the results on stated preferences from cols 3 and 4, we find that Kresy migrants own fewer assets, even if they could afford them. 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.

5 Ancestry Survey

So far, we have shown that people with Kresy ancestors have higher education than other Poles who live in the same county. This holds both in Central Poland and in the Western Territories. The control groups in these two regions are potentially different – in Central Poland, 'other Poles' are largely those who had already lived there before WWII; in WT, many of the Poles without Kresy ties are descendants of voluntary migrants who moved to the largely empty WT from Central Poland after WWII. When interpreting our findings it is important to know the composition of the control group – especially in WT – and to address potential selection issues. The Diagnoza data offer only limited possibilities to perform such checks, because the survey only asked about ancestors with Kresy origin. We will thus largely rely on our own Ancestry Survey from 2016, which covers ancestors from all regions. Before moving to these checks and the interpretation of our results in Section 6, we show that our main results also hold in the Ancestry Survey.

5.1 Ancestry Survey Results – Respondent Level

Our Ancestry Survey 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 ancestors from Kresy as our main explanatory variable in equation 1 and add the shares of autochthons and foreigners as controls. Note that the ancestral share from Central Poland is thus the reference group. In addition, we also control for the share of each respondent's ancestors who lived in rural areas to capture possibly differences between migrants from rural and urban areas.

Column 1 in Table 6 shows that the share of ancestors from Kresy is associated with significantly higher levels of education among respondents. The coefficient changes very little when we control for the shares of ancestors from Western Territories and from abroad as well as the share of ancestors from rural areas (column 2). In columns 3 and 4, we explore whether the gaps in education levels by origin are different in rural and urban areas by splitting the sample into rural and urban *destinations* of migrants). Note that, in addition, in all regression, we control for share

[35] Note that, not surprisingly, Kresy migrants on average own a larger number of assets, as they earn higher incomes due to their higher levels of education. Controlling for the overall number of assets owned by each household does not change our results.

of ancestors who come from rural *origin* locations. We find that 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.

The negative coefficients on the share of ancestors from WT show that in Western Territories (where the Ancestry Survey was conducted) autochthons have lower education levels as compared to the control group – the descendants from Central Poland's migrants. Overall, the ranking of respondents in terms of education by the origin of ancestors from highest to lowest is: Kresy, Central Poland, Western Territories – or forced migration, voluntary migration, no migration.

While Panel A controls for respondents' county of residence fixed effects, Panel B imposes even more restrictions by including municipality (gminy) fixed effects. These are typically smaller than local labor markets. Coefficients in both panels are very similar, suggesting that local socio-economic characteristics do confound our results.

The main coefficient of interest in Table 6 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 60% of respondents with any Kresy ancestor in our dataset had *all* ancestors from Kresy.

5.2 Ancestry Survey Results – Ancestor Level

In the analysis above, we used our Ancestry Survey at the respondent level. In the following, we gain complementary insights by using the data at the ancestor (a) level, where each ancestor a of each respondent i is a separate observation. The ancestor-level data are particularly useful when controlling for characteristics of ancestors' origin locations. In particular, we estimate the following equation on the sample of ancestors:

$$Y_i = \gamma_1 Kresy_{a_{(i)}} + \psi' \mathbf{A}_{a(i)} + \varphi' \mathbf{O}_{a_{(i)}} + \phi' \mathbf{X}_i + \eta_{Locality_{(i)}} + \varepsilon_{a_{(i)}}, \tag{2}$$

where Y_i is respondent i's education, as above, and $Kresy_{a_{(i)}}$ indicates whether ancestor a of respondent i came from Kresy.[36] In addition to all standard controls for respondents' demographics ($\mathbf{X}_i$) and destination location fixed effects $\eta_{Locality_{(i)}}$ (as described above), we control for ancestor characteristics $\mathbf{A}_{a(i)}$: dummies for whether ancestor a is a parent, grandparent, or great-grandparent of respondent i.[37] $O_{a_{(i)}}$ denotes characteristics at the origin location of ancestor

[36] It is important to note that econometrically respondent-level and ancestor-level regressions are not equivalent. In Appendix B we explore (using Monte Carlo simulations) the conditions under which that coefficient estimates from the ancestor level regressions are comparable to the respondent level regressions from Section 5.1. In short, coefficients are comparable if a given respondent's ancestors (in the generation of youngest adults in 1939) tend to come from the same historical area (Kresy/CP/WT). This condition holds in our data.

[37] 21% of all ancestors are the parents; 54.5% are the grandparents; and 24% are the great-grandparents.

a, such as whether a came from a rural area, and the ethnic, religious, and linguistic composition at a's county of origin before WWII. We also include dummies indicating whether the ancestor was an autochthon or from abroad (leaving Central Poland origin as the comparison group). We cluster error terms by respondents to account for the fact that all ancestry information for a given respondent comes from the same source and education of the respondent does not vary across ancestors.[38]

Table 7 presents our results at the ancestor level. We report the coefficient on Kresy origin separately for the cases where the youngest adult ancestors in 1939 were the respondent's parents, grandparents, or great-grandparents.[39] Odd columns include county fixed effects, whereas even columns control for municipality fixed effects. For years of education (columns 1 and 2) we find a strong positive relationship with Kresy ancestors in the parent and the grandparent generations of ancestors. The coefficient size for the grandparent generation is less than one-half of that for the parent generation, but it remains highly statistically significant. The magnitude of the coefficient for the great-grandparent generation is even smaller and statistically insignificant (although it is not statistically different from the one for the grandparent generation). A very similar pattern (with declining coefficients for older ancestor generations) holds also for secondary education (columns 3-4). These results are consistent with the cohorts results in Table 3 which showed that the relationship between Kresy origin and education is smaller for younger respondents (who in turn are more likely to have great-grandparents as ancestors from 1939). However, the results for the higher education attainment (presented in columns 5 and 6) show a different pattern. Here, the coefficient on Kresy origin is largest for the great-grandparent generation of ancestors. 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 levels of education. The proportion of respondents with higher education is 16% for those with parent migrants, 25% for grandparent migrants, and 37% for great-grandparent migrants. For comparison, the rate of secondary education attainment across generations is: 43% for those with parent migrants, 53% for grandparent migrants, and 69% for great-grandparent migrants. When we compare the coefficients in columns 5 and 6 to the mean of higher educational attainment, their *relative* magnitude is similar across generations. Thus, the results are consistent with a strong persistence of the expulsion-education relationship, where the effect moves across generations to the level of education at which the potential margin of influence is highest.

[38] When using interaction terms with county-of-origin characteristics, we use two-way clustering both at the respondent i level and at the level of ancestors' county of origin.

[39] The oldest ancestor generation in our survey are great-grandparents, and the youngest are parents. We drop the few observations of very old respondents who were themselves expelled from Kresy. Thus, when reporting the coefficients on Kresy separately for the three generations, the Kresy dummy itself is not separately identified.

Border Sample based on Ancestry Survey

Finally, our Ancestry Survey also allows us to perform an alternative border sample analysis, complementing the results from Section 4.1. In particular, in the border sample in Figure 4 we had to assume that individuals still living close to the (west of the) Kresy border today also had ancestors there. In addition, we had to compare individuals who potentially live far away from each other today – those still living close to the Kresy border and those whose ancestors migrated mostly to Western Territories. Our Ancestry Survey allows us to address both points. First, it includes the origin location of all ancestors, which includes a sufficiently high number from Central Poland close to the Kresy border. Second, all respondents in the Ancestry Survey live in Western Territories today, so that we only compare education across people in WT, with ancestors from either side of the Kresy border. In fact, the Ancestry data allow us to go even further: We can use municipality fixed effects for respondents, thus comparing people who live in the same town or village in WT today, but with ancestors from the West or East of the Kresy border.[40]

In Figure 5, we restrict the sample to people with ancestor roots in the area of less than 150 km around the Kresy-Poland border. Underlying the figure is a regression of years of education on our standard controls and municipality fixed effects. Dots correspond to residuals from this regression aggregated into 10km bins (for visualization), while the lines are based on all underlying observations, with the shaded area representing 90% confidence intervals. The figure confirms and extends our previous border results: Comparing people who live in the same municipality in the Western Territories today, those whose ancestors were expelled from just a few km to the East of the Kresy border have significantly higher education than those whose ancestors lived a few km to the West. Table A.4 provides the corresponding regression results as well as additional robustness checks for the border sample analysis, including different distance cutoffs and controlling for spatial polynomials.

Overall, our Ancestry Survey – which includes respondents from Western Territories only – confirms both the magnitude and significance of the relationship between Kresy family origin and education. In the following we discuss possible interpretations of this finding, focusing also on potentially confounding effects of migrant selection.

[40]Including respondent county (let alone municipality) fixed effects was not possible in the border analysis in Section 4.1 because this would have absorbed all geographic variation (i.e., distance from the border) for people to the West of the Kresy border. The reason is that for these respondents, we did not have information on their ancestors and thus had to assume that their current location is identical to the place of origin of their ancestors.

6 What Explains the Higher Education of Descendants of Forced Migrants?

In this section, we consider several potential explanations for the educational advantage of descendants of forced Kresy migrants. First, we consider the possibility that the Polish Kresy population was more educated than other Poles before WWII, or that Kresy migrants were selected. Second, we examine whether selection of *voluntary* migrants from Central Poland to Western Territories may drive our ancestry survey results (the results for Poland overall would be unaffected by this type of selection). Third, we examine various other possible channels such as congested labor markets, differential returns to schooling, and out-migration. None of these potential mechanisms appear to confound our findings. We conclude that the most likely mechanism behind our results is that forced migration shifts preferences towards investment in education.

6.1 Pre-Existing Differences and Selection of Kresy Migrants?

We begin by examining possible pre-existing differences of Kresy migrants, as compared to other Poles. Such differences could result either from differences in education of Poles from Kresy before WWII, or from selection of Kresy migrants. We show that both are unlikely.

6.1.1 Were Poles in Kresy Territories already more Educated before WWII?

An obvious concern is that Poles who were expelled from Kresy may already have been more educated before WWII. We already provided a first glance at this issue in the border sample in Figure 4. Here, we provide a full analysis, covering all counties in the Kresy origin territories. We use the literacy rates of Roman Catholics in the 1921 Polish Census (when Kresy still belonged to Poland). Their Roman Catholic religion distinguished Poles from Ukrainians, Belorussians, Jews, and other ethnicities living in interwar Poland. As reported in Table A.3 in the online appendix, 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 on average before they were forced to migrate, compared to Poles in the rest of the Second Polish Republic. This pattern also holds 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.

6.1.2 Were Forced Migrants from Kresy Selected?

Selection *at the origin* is highly unlikely among Kresy migrants, contradicting the historical narrative of large-scale efforts to expel Poles from Kresy. However, some historical sources do suggest that forced migration out of Kresy was not fully homogenous (see Section 2.2.1). In particular, the pressure on Poles to leave was lower in rural areas in the Belarussian and Lithuanian parts of Kresy. In Ukraine, in contrast, a substantial animosity between Poles and Ukrainians led to an (almost) complete exodus of Poles even from rural areas. In what follows, we explore this vari-

ation by restricting the sample to urban areas and to the Ukrainian parts of Kresy. Our Ancestry Survey allows us to exploit this variation, because it provides the detailed location of origin for each ancestor (both in terms of rural vs. urban, and the region within Kresy).[41]

If selection of Poles from Kresy affected our results, we would expect the coefficient on Kresy origin to vary, depending on how much scope for selection a given ancestor's region of origin offered. Table 8 performs an analysis along these lines, using our main outcome variable – years of education. We create different subsamples, depending on ancestors' locations of origin. Correspondingly, regressions are run at the ancestor level, as given by (2). Column 1 replicates our main result, using all Kresy ancestors: 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 wholesale expulsion is historically documented. In columns 4-6 we repeat the analysis, restricting the sample to ancestors from the Ukrainian part of Kresy, where expulsions were also universal. 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 no selection at the source. 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 urban origin sub-sample. Overall, these results make it unlikely that our results are driven by selection of migrants at the origin. If anything, we find even larger coefficients where there was less scope for selection of migrants from Kresy.

6.1.3 Selection of Kresy Migrants into Destinations?

Even if selection from Kresy territories is unlikely, there may have been selection of Kresy migrants or their descendants *into destinations*. As Table 1 has shown, while the majority of Kresy migrants settled in Western Territories, about one quarter moved to Central Poland. For example, if the most capable Kresy migrants moved to Western Territories, than our results within WT would be biased. We begin by noting that the results from Table 2 (columns 4 and 5) show that the coefficient on Kresy ancestors is very similar in Central Poland and in Western Territories. Next, we restrict the Diagnoza sample to those respondents with Kresy origin. Within this subsample, we compare the level of education of those who live in CP (overall 1,314 respondents) with those who live in Western Territories (2,008 respondents). Years of education is 13.2 years for the for-

[41] Since the Ancestry Survey covers only Western Territories, its results may be affected by selection of the *control group* – migrants from Central Poland that moved to Western Territories. We show in Section 6.2 that this is unlikely.

mer, as compared to 12.2 years for the latter. In other words, those with Kresy ancestors who now live in Western Territories are actually somewhat less educated than Kresy descendants in CP.[42] This implies that if anything, our results for Western Territories underestimate the effect of Kresy origin.

6.2 Does Selection of Voluntary Migrants affect the Ancestry Survey Results?

In our results for Poland overall (i.e., using the Diagnoza survey), selection of the *control group* is not an issue – the control group comprises 'all other Poles.' However, our Ancestry Survey was conducted only in Western Territories, which was not only the destination of forced migrants from Kresy, but also of voluntary migrants from Central Poland. This raises the potential issue of selection of voluntary migrants. In particular, our Ancestry Survey coefficients on Kresy origin would be upward biased if the control group of less educated individuals were more likely to migrate from CP to WT after WWII. In the following, we perform several analyses that show that this is unlikely.

6.2.1 Regional Selection of Voluntary Migrants from CP to WT?

We first examine the possibility of regional selection – migrants from Central Poland coming from particularly poor areas. For each respondent in our Ancestry Survey, we know the place of origin of each of their ancestors; and from the historical censuses, we know the literacy rates at the county of origin. This allows us to compare the historical literacy rates in the county of origin of migrants from Kresy vs. Central Poland. Table 9 presents the results, running regressions at the ancestor level with secondary education as the contemporaneous measure for education in odd columns, and historical literacy in even columns.[43] Column 1 replicates our main results in the subsample for which the historical literacy data at the ancestors' origin are available: respondents with Kresy ancestors today have 7.5 percentage point higher secondary education rates than respondents with ancestors from Central Poland who live in the same county. Next, column 2 uses literacy rates among Poles in 1921 as the dependent variable. The coefficient on $Kresy$ in this regression shows the average difference in literacy in 1921 between counties in Kresy and in Central Poland from which respondents' ancestors originated. Again, because we use county fixed effects, we compare respondents who today live in the same counties in Western Territories. According to the results in

[42] The reason for this difference is probably migration of highly skilled individuals to urban centers such as Warsaw in Central Poland. In fact, the difference is particularly pronounced when comparing urban respondents, and smaller for rural respondents.

[43] We use secondary education as the contemporaneous outcome variable because it is the closest to 'literacy' among our various measures for education. Primary education of respondents in our Ancestry 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).

column 2, the average county of origin of Kresy ancestors actually had a 3 percentage point *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.

Panel B in Table 9 shows that our results also hold when we use literacy rates of Poles in Polish language from the 1897 Russian Census, focusing on the subsample of historical counties of the Russian Empire which after 1918 belonged to Central Poland and Kresy. The number of observations here is lower because the western part of Central Poland was part of the German Empire, and the southern-most part of Kresy and of Central Poland belonged to the Austro-Hungarian Empire. Nevertheless, 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 in 1897 (even columns).[44]

In sum, we find that forced migrants from Kresy came from locations where Poles were historically less educated as compared to locations of origin of voluntary migrants from Central Poland.

6.2.2 Individual Selection of Voluntary Migrants from CP to WT?

While we have shown that regional (county-level) selection is unlikely to affect our results, *individual* selection of voluntary migrants remains a possibility. In particular, it could potentially be the case that uneducated Poles from Central Poland decided to seek a better fortune in the Polish Western Territories, whereas the educated Poles *from the same origin counties* stayed in Central Poland. In the following, we show that this type of selection is unlikely to drive our results.

The negative selection of Central Polish migrants into the Polish Western Territories would imply that the control group in our study has too low education, biasing the coefficient on Kresy origin upward. To show directly that there was no negative individual selection, we would need historical individual-level data on the education of migrants and those who stayed in Central Poland. Since these are not available, we implement a check directly where the selection concern matters: if one were 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 lower education than one should expect. Building on this argument, we can use contemporaneous education to show that individual selection is unlikely to affect our results: In the following we show that respondents in Western Territories with ancestors from Central Poland (i.e., voluntary migrants) are actually *more* educated than a reasonable comparison group – people in counties in Central Poland, where their ancestors originated from.

[44]Note that the 1921 Polish Census and the 1897 Russian Empire Census did not cover Western Territories (which belonged to Germany). The historical literacy of origin locations 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.

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. We only use Diagnoza respondents in Central Poland without any ancestors from Kresy. Using the combined information, we construct, for each ancestor of each respondent, the variable:

$$\Delta Edu_{a_{(i)}} = Edu_i - \overline{Edu}\left(county^{CP}_{a_{(i)}}\right) \tag{3}$$

where Edu_i is today's education of ancestor a's descendant (i.e., respondent i) and the second term $\overline{Edu}\left(county^{CP}_{a_{(i)}}\right)$ denotes the average education in the county of origin of 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 $\Delta Edu_{a_{(i)}} = 0$ (clustering standard errors at the level of respondent i). We show results for both years of schooling and for secondary education. The significantly positive differences indicate that descendants of CP migrants who now live in WT have on average *higher* education than their counterparts in the origin counties in CP. This pattern also holds when we restrict attention to the much smaller subsamples in which *both* origin and destination locations were urban (columns 3 and 4), and in which both origin and destination were rural (columns 5 and 6). The results from the subsamples make it highly unlikely that the significant differences emerge mechanically due to predominant migration from rural CP areas to urban WT locations.

The results from Table 10 suggest 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.[45] Nevertheless, the results are relevant for interpreting the coefficient on Kresy origin in our Ancestry Survey 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.[46] Consequently, we *underestimate* the effect for Kresy origin in WT.

[45]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 is related to labor market spillovers in Western Territories from educated descendants of Kresy migrants onto descendants of CP migrants. This would be consistent with spillovers as documented by Semrad (2015).

[46]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.

6.3 Other Potential Channels: Congestion, Returns to Schooling, Out-migration

The previous literature (as discussed in the introduction) showed that migrants who lack access to local land resources (which are held be entrenched locals) often opt for education in order to get access to manufacturing jobs. This is unlikely to affect our results for several reasons. First, the Western Territories were largely empty, and the idea of the resettlement was to populate this 'empty space.' Second, as we described in the background section, migrants from Kresy and CP arrived to WT at the same time (see Figure A.7). Third, if local congestion drove up the incentives to invest in education, this would be captured by county or municipality fixed effects. Thus, a differential congestion effect for Kresy and CP migrants is *a priori* unlikely. Nonetheless, we test for this channel, using variation in the population of autochthons that remained in Western Territories after WWII. Autochthons were a minority in WT, but their share varied across localities. We measure the share of autochthons in two ways: we use the share of Polish native speakers in 1900 (from the German 1900 Census) and the share of autochthons in 1950 from the Polish Census. The share of autochthons in 1950 in the median county was 6.5% and the mean 15% (SD of 21). The corresponding figures for the share of Polish speakers in 1900 is as follows: median 1.4%, mean 13%, and SD 23%. Figure A.2 in the online appendix shows that the two measures are highly correlated.

While local fixed effects in our previous regressions would have captured any direct effect of congestion (i.e., the local share of autochthons) on education, it is still possible that congestion affected Kresy migrants differentially. To address this possibility, we add an interaction term between $Kresy$ and the autochthon share to regression 1. In addition, we include fixed effects at the regional (instead of county) level, so that the direct coefficient on the county-level share of autochthons is also identified. Columns 1 and 2 in Table 11 use our Diagnoza data for respondents in Western Territories, with years of education as dependent variable.[47] The coefficient on the share of autochthons is positive and significant, both when measured by the share of Polish speakers in 1900 (col 1) and by data from the 1950 Polish Census (col 2). While this result is compatible with congestion effects on education, it may also reflect other factors, such as systematic differences of former German counties with higher pre-WWII Polish population. We also find that the interaction effect between Kresy origin and the historical presence of autochthons is relatively small and insignificant. To facilitate the interpretation of coefficient sizes, we standardized the share of autochthons. Thus, the interaction coefficient of about 0.1 implies that a one standard deviation higher share of autochthons is only associated with 0.2 extra years of schooling among people with

[47] We can run this analysis either using Diagnoza or Ancestry Survey data. We use the former because the sample size in Diagnoza is bigger – even when the sample is restricted to respondents from WT (where autochthons were present). Results using the Ancestry Survey are very similar and available upon request.

Kresy ancestors – relative to a direct Kresy coefficient of almost 0.8. These results suggest that differential congestion effects for Kresy vs. CP migrants are unlikely to confound our findings.[48]

Could our results be driven by differential returns to schooling for Kresy migrants? We shed light on this question in columns 3 and 4 of Table 11. We use log household income as the dependent variable, and are interested in the interaction term between $Kresy$ and years of education. A significantly positive coefficient would imply higher returns to schooling for Kresy migrants. We find that the interaction term is small and insignificant in both the full Diagnoza sample and in the subsample of Western Territories.[49] This suggests that differential returns to schooling do not affect our results.

Finally, columns 5 and 6 of Table 11 examine whether differential migration from Poland to other countries may affect our results. For example, if uneducated people with Kresy origin (or educated people without Kresy origin) were more likely to leave Poland, then this could bias the coefficient on Kresy upwards. We use the fact that the Diagnoza survey asked respondents whether they "plan to go abroad within the next two years, in order to work?"[50] We find no relationship between Kresy ancestry and the intend to emigrate (col 5). In addition, the intereaction term between education and Kresy origin is also small and insignificant (col 6). This makes it unlikely that education and Kresy origin drove emigration in a fashion that would confound our results.

Another potential confounding factor is that Kresy migrants might be more likely to have moved in groups from the same location of origin. If moving in groups was beneficial to their descendant's education, this may have reinforce the education effect we observe. While we do not have census-type data on the number of migrants in a destination who are from the same origin, our Ancestry Survey allows us to generate a proxy measure. We can compute, for each municipality, the number of ancestors in the survey who are from the same county of origin. This is likely to be a noisy measure, as it is based on a count within our survey alone. We can, however, use this count variable as a left-hand side variable to check whether descendants of Kresy migrants have a larger number of people from the same place of origin living in same municipality. We find no statistically significant difference between Kresy and CP migrants in this regard.[51]

[48] Another type of congestion effect could theoretically have affected Kresy migrants if they had systematically arrived in Western Territories after migrants from Central Poland. However, this is not consistent with the historical evidence. As shown in appendix Figure , which uses month-by-month figures of arrival, both migrants from CP and from Kresy arrived in parallel throughout the years after WWII.

[49] The coefficient on Kresy itself is insignificant because we also control for years of education, confirming the result from Table 4.

[50] The Polish Census in 2011 included a similar question: "How many members of your HH have emigrated?" The response to this Census question is publicly available at the regional level. In Figure A.3 we show that there's a strong positive relationship between actual out-migration and the intent to migrate reported in Diagnoza. This validates our use of the latter as a proxy for emigration from Poland.

[51] In a regression of the number of migrants from the same county of origin on a dummy for Kresy origin and our

6.4 Heterogeneity of Origin Locations

If the effect of Kresy origin is driven by the move itself and not by characteristics of origin locations, then origin locations should not interact with Kresy origin. To analyze this aspect, we use our Ancestry Survey and enrich the specifications in regression (2) by adding interaction terms between $Kresy$ and characteristics of the location of origin (or destination, or the difference between destination and origin characteristics).

Specifically, using the Polish Census in 1931, we interact Kresy origin with the share of Roman Catholics, the share of native Polish speakers and native Russian speakers, the literacy rate, and the urbanization rate. We perform similar analysis using the Polish Census in 1921. Going beyond population characteristics, we also look at climate variables at the place of origin. A large share of the population was working in agriculture pre-1939. Thus, land suitability, temperature, the precipitation-evatranspiration ratio, and ruggedness are key features of the economic environment. Tables A.5 and A.6 in the online appendix show that none of the interaction terms with Kresy origin is statistically significant. We interpret this as evidence that the effect of uprootedness we uncover is universal and not driven by specific circumstances at the place of origin.

6.5 Interpretation of Results

Summing up, we have performed numerous checks whose results speak against selection as a driver of our results, and against alternative explanations such as differential returns of education or congestion of local labor markets. One explanation that is compatible with all our findings is the prominent – yet debated – argument that forced migration causes a shift in preferences towards investment in mobile assets, and especially in human capital. The population movements in Poland after WWII provide a unique setting to test this – notoriously hard-to-isolate – mechanism. We provided a host of results that suggests that indeed, forced migration caused an increase in educational investment among the affected Poles, relative to all other Poles. This interpretation is not only supported by our main results, but also by our findings for preferences with respect to children's education and material possession.

7 Conclusion

Forced migration is an issue in both historical 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 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 are very dramatic experiences, but

standard control variables, the coefficient on Kresy origin is -0.015 with a standard error of 0.112.

their long-run effects depend on the circumstances that await those displaced at the 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. We find that the children, grand-children, and great-grand-children of forced Kresy migrants have significantly higher average education levels than all other Poles. This result holds in a border sample around the former Kresy border, and it is robust to a host of controls. We examined several possible interpretations. The most likely is that uprootedness shifted forced migrants' preferences away from investment in physical assets and towards investment in human capital. This interpretation is also supported by survey evidence that respondents with Kresy ancestors report assign a higher importance to their children's education and less importance to physical possessions. The shift in preferences towards education offers a glimmer of hope for descendants of those who experience 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 foster 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.

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.

Bharadwaj, P., A. I. Khwaja, and A. Mian (2015). The Economics of the Diaspora: Discrimination and Occupational Structure. *International Migration 53*(4), 90–106.

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 (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.

Brenner, R. and N. M. Kiefer (1981). The Economics of the Diaspora: Discrimination and Occupational Structure. *Economic Development and Cultural Change 29*(3), 517–534.

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*.

Ciesielski, S. (1999). *Przesiedlenie ludnosci polskiej z Kresow wschodnich 1944-1947*. Warszawa: Wydawnictwo Neriton.

Cinnirella, F. and R. Schueler (2018). The Cost of Decentralization: Linguistic Polarization and the Provision of Education.

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

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica 78*, 1863–1903.

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 Second Generation 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.

Gelman, A. and G. Imbens (2014). Why High-order Polynomials Should not be Used in Regression Discontinuity Designs. NBER Working Paper 20405.

Halicka, B. (2013). *Polens Wilder Westen. Erzwungene Migration und die kulturelle Aneignung des Oderraumes 1945-1948*. Ferdinand Schöningh Verlag, Paderborn.

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

Jacob, B. A. (2004). Public Housing, Housing Vouchers, and Student Achievement – Evidence from Public Housing Demolitions in Chicago. *American Economic Review 94*(1), 233–258.

Kersten, K. (1986). *Narodziny systemu wladzy. Polska 1943-1948*. Paris: Wydawnictwo Libella.

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

Magocsi, P. R. (1996). *A History of Ukraine. The Land and Its Peoples*. Toronto: University of Toronto Press.

Murard, E. and S. Sakalli (2018). Mass Refugee Inflow and Long-run Prosperity: Lessons from the Greek Population Resettlement. Working Paper.

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.

Sequeira, S., N. Nunn, and N. Qian (2017). Migrants and the Making of America: The Short and Long-Run Effects of Immigration during the Age of Mass Migration. NBER Working Paper 23289.

Snyder, T. (2003). *The Reconstruction of Nations. Poland, Ukraine, Lithuania, Belarus 1569-1999*. New Haven and London : Yale University Press.

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.

Wirtschaft und Statistik (1956). Deutsche Bevölkerungsbilanz des 2. Weltkrieges. *Wirtschaft und Statistik 10*, 493–500.

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

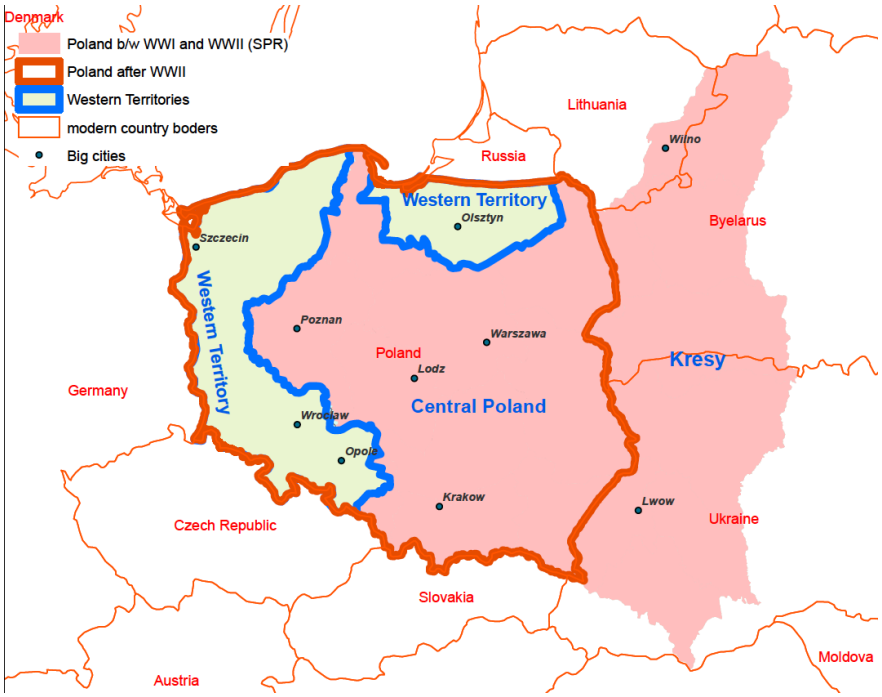


Figure 1: Map of Poland's Territorial Change after WWII

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. Poles from Kresy were forced to leave – the vast majority was resettled to the emptied Western Territories.

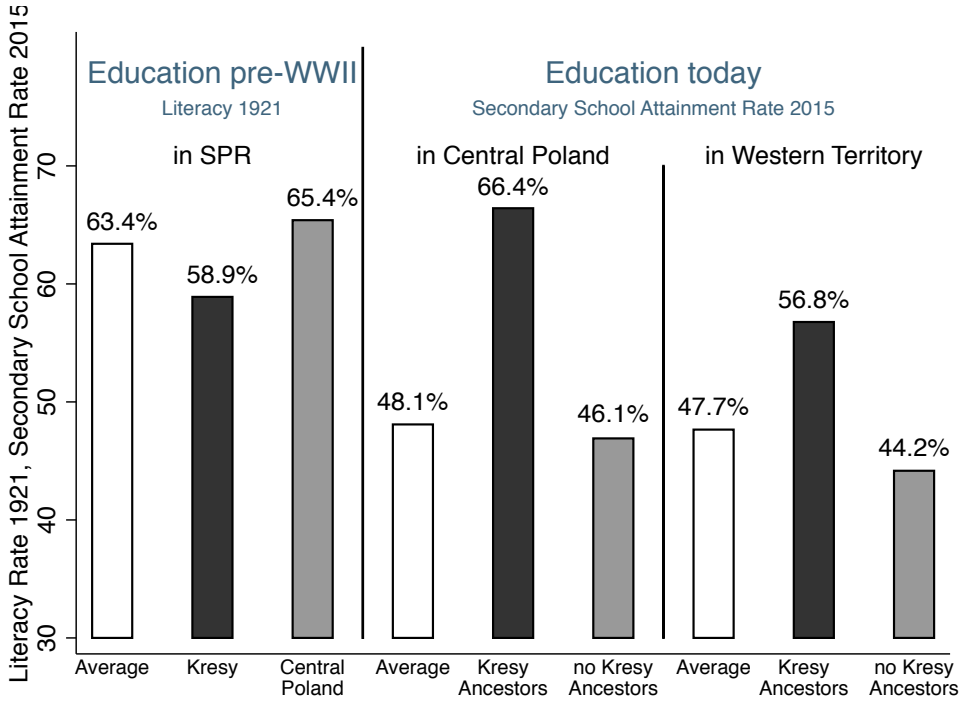


Figure 2: Overview of Historical and Contemporaneous Patterns in Education

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 the whole of the Second Polish Republic (SPR), which consisted of Kresy (Eastern Borderlands) and Central Poland (CP). Literacy rates were lower in Kresy than in CP. For today's Poland, the figure shows the secondary school attainment rate for those residing in CP and in Western Territories (WT). The average is nearly identical in the two regions (and 48.1% in CP and 47.7% in WT). However, in both CP and WT, those with Kresy ancestors are more educated.

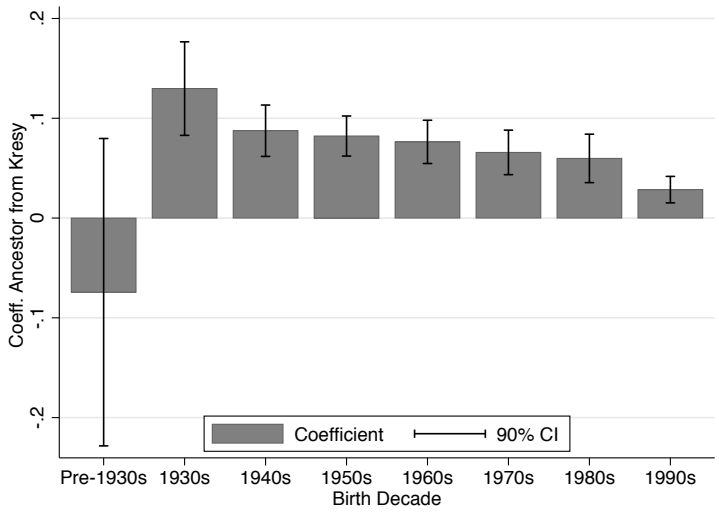


Figure 3: Ancestors from Kresy and Eduction, by Birth Cohort

Note: The figure visualizes the results of regressing ln(years of education) on Kresy ancestry for different birth cohorts. Each bar corresponds to the coefficient on 'Ancestor from Kresy'; regressions include our standard controls (see Table 2) and county fixed effects. The pre-1930 birth cohort was at least 16 years old at the end of WWII and was thus not of schooling age anymore after being forced to leave the Kresy territories. The regressions are run using the full Diagnoza sample for 2015 (see Table 3 for similar regressions) for respondents from all over Poland.

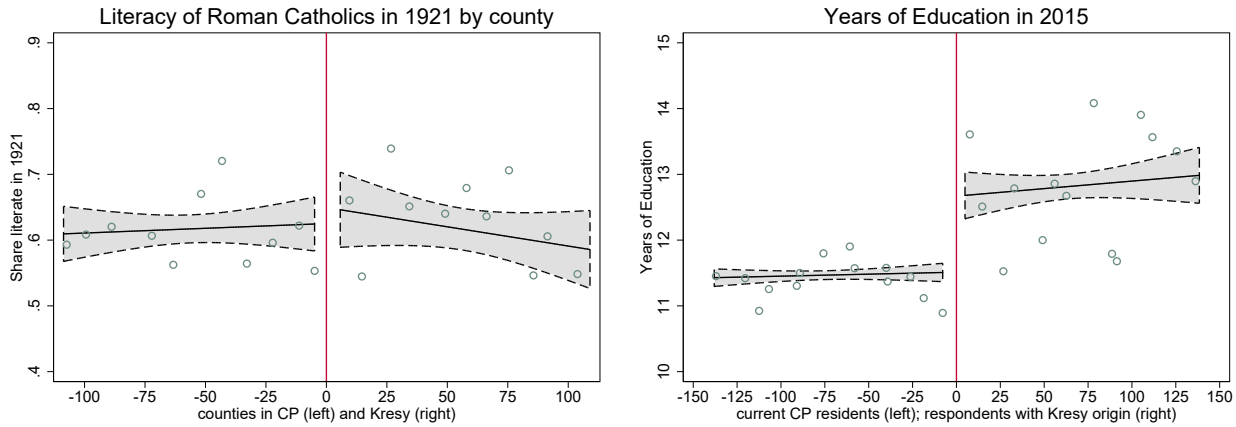


Figure 4: Border Sample Around the Kresy-Poland Border

Note: The figure uses only respondents (from the 2015 Diagnoza survey) with roots in the area of less than 150 km around the Kresy-Poland border. The left panel shows that there is no difference in literacy in 1921 around the Kresy border. The right panel tracks individuals with roots near the Kresy border by including i) individuals from the Diagnoza survey with ancestors from Kresy who lived within less than 150 km to the east of the border, and ii) individuals without Kresy ancestors who live (today) within less than 150 km to the west of the border. Dots correspond to data aggregated into 10km bins (for visualization), while the lines are based on all underlying observations, with the shaded area representing 90% confidence intervals.

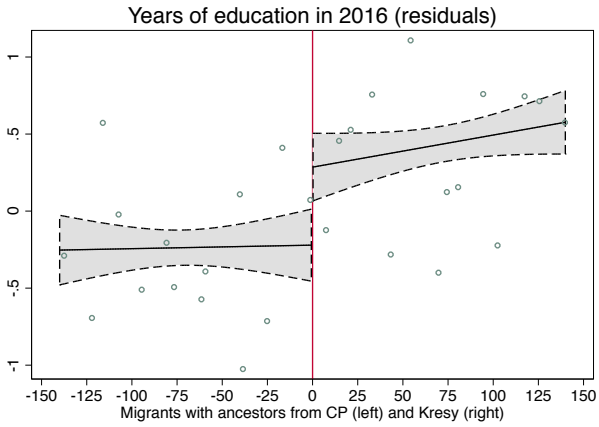


Figure 5: Kresy Border Sample: Ancestor Survey

Note: The figure uses respondents from our Ancestor Survey, i.e., individuals who live in Western Territories today. Among these, we restrict the sample to people with ancestor roots in the area of less than 150 km around the Kresy-Poland border. Underlying the figure is a regression of years of education on our standard controls (respondent's gender, age, age^2, as well as indicators for rural places and urban counties) and, most importantly municipality fixed effects. Dots correspond to residuals from this regression aggregated into 10km bins (for visualization), while the lines are based on all underlying observations, with the shaded area representing 90% confidence intervals. The corresponding regression results are presented in Table A.4.

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 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 Survey

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

Sample:	(1) All	(2) Rural	(3) Urban	(4) Central Poland	(5) Western Territories	(6) Kresy Border Sample†
Panel A. Dep. Var.: Years of education						
Ancestor from Kresy	0.849***	0.719***	0.917***	0.937***	0.752***	1.386***
	(0.073)	(0.118)	(0.093)	(0.111)	(0.098)	(0.305)
Western Territory	-0.640	-0.108	-0.664			-0.277
	(0.478)	(0.554)	(0.478)			(0.248)
Mean Dep. Var.	11.73	10.95	12.51	11.75	11.65	11.72
Observations	29,161	14,677	14,484	21,854	7,307	5,446
Panel B. Dep. Var.: Level of degree (4 categories)						
Ancestor from Kresy	0.257***	0.227***	0.272***	0.263***	0.246***	0.336***
	(0.023)	(0.040)	(0.027)	(0.033)	(0.031)	(0.096)
Western Territory	-0.011	0.045	-0.024			-0.082
	(0.165)	(0.168)	(0.164)			(0.079)
Mean Dep. Var.	2.47	2.21	2.74	2.48	2.46	2.48
Observations	29,104	14,656	14,448	21,807	7,297	5,440
Panel C. Dep. Var.: Secondary education dummy						
Ancestor from Kresy	0.114***	0.111***	0.119***	0.114***	0.111***	0.109**
	(0.011)	(0.019)	(0.013)	(0.016)	(0.015)	(0.047)
Western Territory	0.023	0.153**	0.019			-0.026
	(0.073)	(0.078)	(0.074)			(0.037)
Mean Dep. Var.	0.48	0.36	0.61	0.48	0.48	0.49
Observations	29,104	14,656	14,448	21,807	7,297	5,440
Panel D. Dep. Var.: Higher education dummy						
Ancestor from Kresy	0.089***	0.062***	0.102***	0.116***	0.064***	0.160***
	(0.010)	(0.015)	(0.013)	(0.015)	(0.013)	(0.039)
Western Territory	-0.098	0.071**	-0.090			-0.032
	(0.098)	(0.035)	(0.098)			(0.033)
Mean Dep. Var.	0.19	0.12	0.28	0.20	0.19	0.20
Observations	29,104	14,656	14,448	21,807	7,297	5,440
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 respondent level; standard errors are clustered at the household level. * p<0.1, ** p<0.05, *** p<0.01.

† Column 6 uses only respondents with roots in the area of less than 150 km around the Kresy-Poland border. These include i) individuals from the Diagnoza survey with ancestors from Kresy who lived within less than 150 km to the east of the border, and ii) individuals without Kresy ancestors who live (today) within 150 km to the west of the border.

‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for Western Territories, rural places and urban counties. Column 6 also includes a quadratic polynomial in latitude and longitude to the Kresy-Poland border.

Table 3: Education and Kresy ancestors in 2015 Diagnoza – across cohorts

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Birth Decade:	pre-1930	1930s	1940s	1950s	1960s	1970s	1980s	1990s
Age in 1945:	16+	6-15	<5	-	-	-	-	-
Age in 2015:	86+	76-85	66-75	56-65	46-55	36-45	26-35	16-25
Panel A: Dep. Var.: Years of education								
Ancestor from Kresy	-0.668	1.356***	0.940***	0.964***	1.022***	0.899***	0.852***	0.330***
	(0.826)	(0.278)	(0.179)	(0.153)	(0.172)	(0.185)	(0.197)	(0.095)
Mean Dep. Var.	7.61	9.44	10.50	11.57	12.27	13.07	13.95	11.85
R-squared	0.66	0.44	0.31	0.24	0.23	0.30	0.27	0.63
Observations	519	2,083	3,360	5,405	4,435	4,152	3,837	4,386
Panel B: Dep. Var.: Secondary education dummy								
Ancestor from Kresy	0.032	0.167***	0.144***	0.137***	0.146***	0.095***	0.079***	0.041*
	(0.094)	(0.035)	(0.028)	(0.024)	(0.027)	(0.026)	(0.024)	(0.022)
Mean Dep. Var.	0.20	0.35	0.40	0.43	0.47	0.58	0.75	0.47
R-squared	0.59	0.42	0.28	0.23	0.23	0.26	0.23	0.51
Observations	523	2,085	3,361	5,402	4,436	4,150	3,840	4,383
Panel C: Dep. Var.: Higher education dummy								
Ancestor from Kresy	-0.066	0.112***	0.090***	0.106***	0.139***	0.124***	0.060*	0.028*
	(0.077)	(0.033)	(0.025)	(0.022)	(0.025)	(0.027)	(0.031)	(0.017)
Mean Dep. Var.	0.06	0.13	0.15	0.15	0.18	0.29	0.42	0.10
R-squared	0.51	0.27	0.18	0.15	0.21	0.27	0.27	0.31
Observations	523	2,085	3,361	5,402	4,436	4,150	3,840	4,383
Respondent county FE	✓	✓	✓	✓	✓	✓	✓	
Controls‡	✓	✓	✓	✓	✓	✓	✓	

Notes: The table shows that the results from Table 2 hold across different age cohorts. Regressions are run at the respondent 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 Western Territories, rural places and urban counties.

Table 4: Labor market outcomes

Dependent variable: Individual labor market outcomes, as indicated in table header. Diagnoza Sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable:	ln(HH income)		White Collar Job		Entrepreneur		Unemployed	
Ancestor from Kresy	0.098**	0.051	0.103***	0.035***	0.024***	0.021**	-0.021***	-0.014*
	(0.039)	(0.039)	(0.014)	(0.012)	(0.008)	(0.009)	(0.007)	(0.008)
ln(years of education)		0.789***		0.947***		0.075***		-0.117***
		(0.047)		(0.028)		(0.009)		(0.011)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	8.47	8.46	0.46	0.46	0.06	0.07	0.08	0.09
R-squared	0.20	0.22	0.26	0.40	0.06	0.06	0.06	0.07
Observations	17,764	15,923	13,475	13,463	18,348	16,454	18,348	16,454

Notes: The table shows that descendants of Kresy migrants have move favorable labor market outcomes. Regressions are run at the respondent 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 Western Territories, rural places and urban counties.

Table 5: Attitudes towards Education and Material Possessions

Dependent variable: Individual-level outcomes, as indicated in table. Diagnoza sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Aspiration for education of own children (1 low – 5 high)		Main condition for success in life?				% Assets not owned for non-financial reasons†	
			Material goods		Freedom			
Ancestor from Kresy	0.151***	0.130*	-0.073***	-0.059***	0.020***	0.020***	0.038***	0.028***
	(0.051)	(0.072)	(0.013)	(0.013)	(0.006)	(0.006)	(0.008)	(0.008)
ln(years of education)		1.458***		-0.191***		-0.008		0.108***
		(0.151)		(0.015)		(0.006)		(0.007)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	4.31	4.18	0.56	0.56	0.05	0.05	0.62	0.62
R-squared	0.21	0.27	0.12	0.13	0.08	0.08	0.18	0.20
Observations	8,470	4,747	22,087	22,050	21,623	21,587	37,611	29,019

Notes: The table shows that descendants of Kresy migrants have stronger preferences for the education of their children, value material goods less, value freedom more, and chose to own fewer assets (even if they could afford them). Regressions are run at the Diagnoza respondent 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 Western Territories, rural places and urban counties.
† Diagnoza asks about the possession of 20 different assets (e.g., vacation house, ebook, home theatre, boat). For those assets not possessed, respondents are asked if this is for financial reasons. The dependent variable in columns 7-8 is the number assets not owned for non-financial reasons, divided by the number of all non-possessed assets.

Table 6: Results from the Ancestry Survey: Education in Western Territories

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Years of Education				Secondary	Higher
Notes on sample:			rural	urban		
Panel A: Respondent County Fixed Effects						
Share of Ancestors, Kresy	1.010***	0.808***	0.688***	0.792***	0.113***	0.069***
	(0.132)	(0.137)	(0.239)	(0.168)	(0.021)	(0.017)
Share of Ancestors, WT		-1.011***	-0.581*	-1.272***	-0.162***	-0.131***
		(0.191)	(0.323)	(0.245)	(0.031)	(0.024)
Share of Ancestors, abroad		-1.170	-3.635**	-0.300	-0.048	0.005
		(0.832)	(1.509)	(0.893)	(0.110)	(0.097)
Share of Ancestors, rural		-0.479***	-0.473	-0.528***	-0.059**	-0.036*
		(0.161)	(0.364)	(0.179)	(0.024)	(0.020)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	12.43	12.45	11.40	12.96	0.47	0.22
R^2	0.28	0.29	0.32	0.27	0.20	0.21
Observations	3,716	3,668	1,110	2,558	3,668	3,668
Panel B: Respondent Municipality Fixed Effects						
Share of Ancestors, Kresy	0.940***	0.740***	0.503**	0.830***	0.095***	0.067***
	(0.138)	(0.143)	(0.253)	(0.176)	(0.022)	(0.019)
Share of Ancestors, WT		-0.989***	-1.016***	-1.080***	-0.165***	-0.125***
		(0.194)	(0.345)	(0.250)	(0.033)	(0.025)
Share of Ancestors, abroad		-0.626	-2.986**	0.209	0.025	0.038
		(0.649)	(1.372)	(0.719)	(0.095)	(0.098)
Share of Ancestors, rural		-0.562***	-0.352	-0.566***	-0.076***	-0.045**
		(0.157)	(0.331)	(0.184)	(0.024)	(0.020)
Controls‡	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	12.43	12.45	11.40	12.96	0.47	0.22
R^2	0.38	0.38	0.47	0.34	0.30	0.29
Observations	3,716	3,668	1,110	2,558	3,668	3,668

Notes: The table uses data from our 2016 Ancestry 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 respondent 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, dummies for six age groups, as well as indicators for respondents living in rural places and urban counties. Excluded category is ancestors from Central Poland.

Table 7: Ancestor-level regressions for education outcomes

Dependent variable: Education at the Respondent Level in 2016

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Years of Education		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 reports the effect of forced migration (ancestors from Kresy) on education across generations. 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, dummies for six age groups, as well as indicators for respondents living in rural locations and urban counties. Excluded category is ancestors from Central Poland.

Table 8: Results from Ancestry Survey, by origin of ancestors

Dependent variable: Years of Education in 2016

	(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
Ancestor from Kresy	0.503***	0.431***	0.670***	0.445***	0.347***	0.617***
	(0.092)	(0.110)	(0.160)	(0.111)	(0.132)	(0.182)
Ancestor from rural area	-0.508***			-0.487***		
	(0.099)			(0.106)		
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	13.04	12.84	13.64	12.98	12.80	13.52
R^2	0.29	0.32	0.30	0.28	0.32	0.31
Observations	11,548	8,598	2,950	10,237	7,669	2,568

Notes: The table uses data from our 2016 Ancestry 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 respondent level; robust standard errors indicated in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age^2, dummies for six age groups, indicators for respondents living in rural locations and urban counties, and indicators for ancestors from rural areas as well as from Western Territories. Excluded category is ancestors from Central Poland.

Table 9: Education today and historically 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	0.073***	-0.031***	0.058***	-0.041***	0.114***	-0.002
	(0.015)	(0.004)	(0.018)	(0.005)	(0.025)	(0.005)
Ancestor from rural area	-0.062***	-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	0.23	0.12	0.25	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	0.145***	-0.031***	0.139***	-0.031***	0.152**	-0.030***
	(0.030)	(0.003)	(0.034)	(0.004)	(0.072)	(0.008)
Ancestor from rural area	-0.041	0.003				
	(0.033)	(0.004)				
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.58	0.16	0.57	0.16	0.63	0.15
R^2	0.31	0.30	0.32	0.32	0.56	0.53
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 – on average – 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, dummies for six age groups, indicators for respondents living in rural locations and urban counties, and an indicator for ancestors from Western Territories. Excluded category is ancestors from Central Poland.

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

Dep. Var.: Difference in education, variable indicated in table header

Dep. Var.:	(1) Years of Education	(2) Secondary Degree	(3) Years of Education	(4) Secondary Degree	(5) Years of Education	(6) Secondary Degree
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 addresses the possibility of individual selection of voluntary migrants from Central Poland to Western Territories (which would affect the composition of the control group in our Ancestry Survey results). The table shows that respondents in WT who are descendants of migrants from Central Poland are *better* educated than a reasonable comparison group – people who still live in the places of their ancestors' origin in Central Poland. Thus, our results on Kresy origin are if anything downward biased due to the high education in the control group. Regressions are run at the ancestor level; standard errors clustered by respondent. * p<0.1, ** p<0.05, *** p<0.01.

Table 11: Other Potential Channels: Congestion, Returns to Schooling, Out-Migration

Dep. Var.: as indicated in table header. Data from Diagnoza.

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Years of Education		log(HH income)		Intend to go abroad	
Sample	WT	WT	all	WT	all	all
Ancestor from Kresy	0.761***	0.777***	0.067	0.013	0.003	0.003
	(0.079)	(0.083)	(0.044)	(0.050)	(0.006)	(0.006)
% Polish speakers, 1900 (std)	0.287***					
	(0.069)					
% Polish (std) × Kresy	0.102					
	(0.074)					
Share Autochthons, 1950 (std)		0.202**				
		(0.079)				
Autochthons (std) × Kresy		0.133				
		(0.094)				
Years education (std)			0.211***	0.246***		-0.001
			(0.013)	(0.024)		(0.002)
Years edu (std) × Kresy			-0.025	0.014		0.001
			(0.030)	(0.041)		(0.005)
Controls‡	✓	✓	✓	✓	✓	✓
Region FE	✓	✓				
County FE			✓	✓	✓	✓
Mean Dep. Var.	11.65	11.65	8.46	8.40	0.06	0.06
R-squared	0.30	0.30	0.22	0.21	0.09	0.09
Observations	7,307	7,307	15,933	3,875	22,097	22,067

Notes: The table examines three alternative mechanisms that may explain the education advantage of people with Kresy ancestors: congestion due to the presence of autochthons (cols 1-2); differential returns to education (cols 3-4); differential out-migration (cols 5-6). None of these appear to confound the coefficient on Kresy. 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, dummies for six age groups, as well as indicators for Western Territories, rural places and urban counties.

Online Appendix

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

Sascha O. Becker
University of Warwick

Irena Grosfeld
Paris School of Economics

Pauline Grosjean
UNSW

Nico Voigtländer
UCLA

Ekaterina Zhuravskaya
Paris School of Economics

A Summary statistics and additional results

In this section, we show descriptive statistics and describe additional results, following the ordering of the main text.

A.1 Descriptive statistics

Table A.1 shows summary statistics describing the origin of ancestors in Diagnoza and our own Ancestry Survey. In Diagnoza, 27.3% of respondents in Western Territories declare to have at least one ancestor from Kresy. In Central Poland, this share is 6.0%.

In our own Ancestry Survey, using the representative sample (i.e. excluding the oversample of those with Kresy origin), 31.3% declare to have an ancestor from Kresy in the generation of the youngest adult in 1939. Looking at all ancestors from the generation of the youngest adult in 1939, 23.6% of ancestors are from Kresy, 15.9% from Western Territories, 60.4% from Central Poland, and 2.6% from abroad. Looking at urban vs rural origin of ancestors, 75.9% are from rural areas.

Taking the an ancestor-level perspective, the last panel additionally reveals information about the reference generation about which respondents report: 21.5% of data points are for the parent generation, 54.5% for the grandparent generation, and 24.0% for the greatgrandparent generation.

shows summary statitics for the education in Diagnoza and our Ancestry Survey. Diagnoza collects information both for education degrees, which we group in four categories (primary education, incomplete secondary education, completed secondary education, and higher education), and – separately – years of education. in our own Ancestry Survey, we collect data on the same four education category. We infer years of education by using the average years of education for each of the four education categories in Diagnoza rounded to the nearest integer. Note that the means and medians of years of education for each of the four education categories, rounded to the nearest integers, are the same in Diagnoza.

Table A.3 displays information from the 1921 Polish Census, separately for Kresy and Central Poland. The share of Roman Catholics in Kresy (24.1%) is considerably lower than in Central Poland (79.7%). The 1921 Census provides literacy rates by religion. This allows us to compute the literacy rate of Roman Catholics in Kresy and CP, overall, as well as for urban and rural areas separately. The descriptives suggest that literacy rates of Roman Catholics were lower in Kresy locations. The difference between Kresy and CP are more pronounced in rural areas.

A.2 Additional results

Table A.4 complements the graphical evidence from Figure 4. Both figure and table are our most demanding specifications: we identify the effect of ancestral origin for individuals living in the same county (column 1) or even within the same municipality (column 2 through 5) whose ancestors originate from within a certain distance from the Kresy border, using a two-dimensional RDD a la Dell (2010) with quadratic polynomials in latitude and longitude of ancestral origin. To the extent that, before WWII, the Kresy border did not exist, ancestors that originate close to an inexistent border are likely to be (more) comparable even in unobserved dimensions. We run several specifications to probe the robustness of what is displayed in Figure 4. In columns (1) to (3), we use years of education as the outcome variable. In column (1), ancestors are from within 150km from the Kresy border, and we employ county fixed effects at respondents' location of residence. In column (2), and all subsequent columns we use municipality fixed effects, and the same bandwidth of 150km. In column (3), we use a smaller bandwidth of 100km. In columns (4) and (5), we return to the 150km bandwidth, and look at secondary education dummies and higher education dummies as outcomes. Results in all 5 specifications are strongly statistically significant, suggesting that our main results are not driven by unobserved differences across ancestors' location of origin.

In Table A.5, we probe robustness of our results at the ancestor level with respect to divesity at the place of origin. We run regressions where we interact the main effect of Kresy ancestry with (standardized) county-level variables at the place of origin. The idea of this exercise is to see whether our education effects vary by the conditions at the place of origin in which their ancestors grew up. For instance, since Roman Catholics only made up about a quarter of the Kresy population, most of the ancestors from this areas will have grown up being a Roman Catholic minority, while in some counties they might have constituted a majority, and it is conceivable that the education effect varies along those lines. Alas, none of the interaction terms displayed in the table turns out to be statistically significant,[1] which we read as evidence that the education effects

[1] In Section B, we describe results from a Monte Carlo simulation showing that statistical significance in these regressions is properly measured.

we measure are universal across socio-economic conditions at ancestors' place of origin.

Table A.6 is an exercise similar to Table A.5. Here, we use geographic features at origin as interaction terms. Since pre-WWII Poland was still dominated by agriculture, it might be the case that moving to WT was a blessing or a curse in terms of change in conditions between origin and destination locations. Also here, find no single coefficient on an interaction term to be significant at the 10% level.

The poster displayed in Figure A.1 was used in Central Poland to entice voluntary migration into WT. The title reads "The land is waiting." The text below the picture reads: "The State Repatriation Office is assigning farms in Opole and Lower Silesia. The regional inspectorates (offices) will provide all necessary information." The government also used promotional films to describe WT as a land of opportunity. In the following video – -- the narrator says: "Western Territories are waiting for the Polish settler. Wheat fields, fish farms, farmsteads, agricultural warehouses, ports, towns, free country but still stray country is waiting for a host. Olsztyn, a country center in the previous East Prussia. The main institutions of economic and administrative life are already functioning. But the inflow of the Polish population is still not up to the possibilities and the needs of the cities." The video finishes with the sentence: "The future of the Polish Republic will be decided in the West" (i.e., Western Territories).

Figure A.2 illustrates that the county-level share of autochthons in the 1950 Polish Census, i.e. the local population that declared to be Polish and was not expelled from the Polish Western Territories, is highly correlated with the share of Polish speakers in the German Census of 1900, which was the last census in the German Empire which collected information on language spoken at home. This suggests that autochthons are indeed largely people with ethnic Polish ancestry. They will have had German nationality in German censuses of the inter-war period, but were no longer separately identified in German statistics until the Polish Census of 1950 counted them as autochthons.

Figure A.3 plots the self-declared intention to emigrate of Diagnoza respondents in 2015 (collapsed to the region level) against the share of people who actually emigrated from the same regions according to the 2011 Census. The high correlation suggests that intention to emigrate measures something meaningful, as in previous years the same regions indeed saw larger realized emigration. It gives further support to the evidence presented in Table **??** which shows that the intention to emigrate does not differ for those with Kresy ancestors.

Figure A.4 probes the quality of ancestry information in Diagnoza 2015. Remember that Diagnoza asks respondents about whether they have any ancestor from Kresy. Hypothetically, if Kresy migrants moved to either CP or WT but their descendants still live in the exact same location, we would expect the share of Diagnoza respondents with Kresy ancestry in a location to be very

highly correlated with the share of Kresy migrants captured in the 1950 Polish Census where this information was collected at the region (voivodship) level across all of Poland, and at the county (powiat) level in Western Territories. Of course, since descendants of Kresy migrants may well live in a different location from their ancestors, the link between historic (1950) Kresy migrant shares and modern-day Kresy migrant shares cannot be expected to be perfect, but should still show a clear correlation between the two. The left panel of Figure A.5 uses region-level variation for all of Poland and shows a strong correlation. Interestingly, the historic share of Kresy migrants in Warsaw was quite low but today is substantially higher. This suggests that urban centres like Warsaw over time absorbed some descendants of Kresy migrants. The right panel of A.5 focuses on Western Territories and uses shares at the county (powiat) level. Naturally, to the extent that there is more migration of descendants between 1950 and 2015 across county borders than across region borders, here the correlation between historic and modern Kresy migrant shares is less tight, but still clearly visible.

Figure A.5 probes the quality of ancestry information in our own Ancestry Survey in 2016. 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 0.69 with a standard error of 0.04 and R^2 of 0.59. The right panel repeats this exercise, but using only migrants from Kresy. The regression coefficient is 0.38 with a standard error of 0.09 and R^2 of 0.19.

Figure A.6 displays the origin of ancestors in our Ancestry Survey.

Figure , Panel A shows data on the stock of migrants who have arrived in the Polish Western Territories by month, during the first two years of mass migration. The data start in December 1945 and show that by the, 1.5 mio. migrants had moved into Western Territories. That stock continued to graw, reaching more than 4 mio. migrants by the end of 1947. Panel B shows the share of Kresy migrants in that stock over time, separately for urban and rural destinations. There are two important take-aways: the Kresy migrants is somewhere between 40% and 50% throughout this two-year window, in both urban and rural destinations, suggesting that Kresy migrants and 're-settlers' from CP (the official label used by the Polish authorities) arrived in parallel throughout the whole period. This implies that a potential concern that CP migrants moved into WT more quickly, generating a potential congestion effect for Kresy migrants is not warranted, as both groups arrive at the same time.

B Monte Carlo simulations to compare results from respondent-level vs ancestor-level regressions

In Section 5.2, we run regressions at ancestor level. We do so because it is a convenient way to analyze whether local characteristics at the exact place of origin of a specific ancestor matter for education outcomes of their descendants.

It is important to note that econometrically respondent-level and ancestor-level regressions are not equivalent. In particular, in a basic equation: $Y_i = \alpha Kresy_{a_{(i)}} + \psi' \mathbf{A}_{a(i)} + \phi' \mathbf{X}_i + \eta_{Locality_{(i)}} + \varepsilon_{a_{(i)}}$, the parameter α equals β from equation 1 only in the case when the origins of different ancestors of the same respondent are perfectly correlated. The lower the correlation between the origins of ancestors of the same respondent, the lower the parameter α. If the origins of ancestors were uncorrelated, the parameter α of the ancestor-level regressions would have been equal to the effect of the share of ancestors with Kresy origin divided by the mean number of ancestors per respondent, i.e., α/N, where N is the average number of ancestors per respondent. In reality, the origins are highly correlated across ancestors, so the point estimates of α at the ancestor-level and β at the respondent-level are fairly similar in magnitude. The correlation between the Kresy origins of spouses (e.g., of the mother and father or of the paternal grandmother and paternal grandfather) is over 90%, the correlation between Kresy origins of grandparents from the mother's and father's side, e.g., of the fathers of the parents of the respondent, is over 30%; and the correlation between the origins of the most distant ancestors, i.e., FFF and MMM is 7%.

Importantly, our Monte Carlo simulations suggest that the level of statistical significance is similar in (i) respondent-level regressions and (ii) ancestor-level regressions clustered at the respondent level.

This is true not only for the main effects of Kresy ancestry, but also for the interaction terms between Kresy ancestry and interaction terms capturing local characteristics (at origin or destination). The latter is particularly important because interaction terms between Kresy origin and location-specific averages are generally statistically insignificant. If the t-statistics of the interation terms were systematically underestimated, we would falsely attribute an absence of effects, when in fact those interaction might be statistically significant. However, our Monte Carlo simulations show that t-statistics for the (potentially biased) coefficients in the ancestor-level regressions and in the corresponding t-statistics in the respondent-level regressions are very similar, suggesting that our statistical inference is correct.

Table A.1: Summary statistics for variables describing origin of ancestors

	Obs	Mean	Std. Dev.	Min	Max
Diagnoza: Western Territories					
Ancestor from Kresy	7,363	0.273	0.445	0	1
Diagnoza: Central Poland					
Ancestor from Kresy	21,927	0.060	0.237	0	1
Ancestry Survey: Respondent level, representative sample					
Ancestor from Kresy	2,846	0.313	0.464	0	1
Share of ancestors from Kresy	2,846	0.236	0.381	0	1
Share of ancestors from WT	2,847	0.159	0.341	0	1
Share of ancestors from CP	2,847	0.604	0.441	0	1
Share of ancestors from rural areas	2,903	0.759	0.377	0	1
Share of ancestors from abroad	2,963	0.026	0.137	0	1
Ancestry Survey: Ancestor level, whole sample					
Ancestor from Kresy	11,928	0.324	0.468	0	1
Ancestor from CP	11,933	0.515	0.500	0	1
Ancestor from WT	11,933	0.160	0.367	0	1
Ancestor from rural area	12,435	0.751	0.433	0	1
Ancestor's gender	13,223	0.496	0.378	0	1
Ancestor code	13,223	6.283	3.432	2	15
Generational gap	13,223	2.026	0.674	1	3
Parent	13,223	0.215	0.411	0	1
Grandparent	13,223	0.545	0.498	0	1
Great-grandparent	13,223	0.240	0.427	0	1

Notes: The table shows summary statistics for ancestry variables in Diagnoza 2015 and our Ancestry Survey 2016.

Table A.2: Summary statistics for education variables

	Obs	Mean	Std. Dev.	Min	Max
		Diagnoza: Western Territories			
Educational degree	7,372	2.455	1.027	1	4
Secondary education	7,372	0.477	0.499	0	1
Higher education	7,372	0.191	0.393	0	1
Education years	7,357	11.664	3.330	0	25
		Diagnoza: Central Poland			
Educational degree	21,959	2.476	1.022	1	4
Secondary education	21,959	0.481	0.500	0	1
Higher education	21,959	0.196	0.397	0	1
Education years	21,930	11.764	3.350	0	28
		Ancestry Survey: Respondent level, representative sample			
Educational degree	3,169	2.671	1.0041	1	4
Secondary education	3,169	0.4781	0.500	0	1
Higher education	3,169	0.208	0.406	0	1
Education years	3,169	12.529	3.189	7	17

Notes: The table shows summary statistics for education variables in Diagnoza 2015 and our Ancestry Survey 2016.

Table A.3: Literacy rates of Poles in Kresy and Central Poland parts of the SPR

1921 Polish Census:	**Kresy**	**Central Poland**
Share of Roman Catholics	24.1	79.7
Share of literate Roman Catholics, total	58.9	65.4
Share of literate Roman Catholics, urban	73.6	74.1
Share of literate Roman Catholics, rural	55.4	63.2

Notes: The table shows that average literacy among Roman catholics in 1921 in Kresy and Central Poland.

Table A.4: Education in Western Territories: Ancestors Originating Near Kresy Border

Dependent variable: as indicated in column header

	(1)	(2)	(3)	(4)	(5)
Dep. Var.:	Years of education (based on 4 categories)			Secondary	Higher
Notes on sample:	$< 150km$	$< 150km$	$< 100km$	$< 150km$	$< 150km$
Ancestor from Kresy	0.681*	0.946**	1.334**	0.115**	0.126**
	(0.355)	(0.392)	(0.527)	(0.058)	(0.055)
Controls‡	✓	✓	✓	✓	✓
County FE	✓				
Municipality FE		✓	✓	✓	✓
Mean Dep. Var.	12.72	12.72	12.66	0.54	0.24
R^2	0.30	0.43	0.53	0.41	0.36
Observations	3,291	3,291	1,949	3,291	3,291

Notes: The table uses data from our 2016 Ancestry Survey in Western Territories, using only ancestors from within given distance from Kresy border. Regressions are run at the respondent level; robust standard errors indicated in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

‡ Controls include respondents' gender, age, age2, dummies for six age groups, as well as indicators for rural places and urban counties, an indicator for the generation of the ancestor, rural location of the ancestor. All columns control for a quadratic polynomial in latitude and longitude of ancestor's location of origin.

Table A.5: No heterogeneous effects w.r.t. origin diversity:

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Years of education (based on 4 categories)							
Ancestor from Kresy	0.586***	0.577***	0.507***	0.557***	0.576***	0.546***	0.553***	0.493***
	(0.131)	(0.144)	(0.105)	(0.109)	(0.097)	(0.107)	(0.097)	(0.187)
Share Rom. Cath., 1931 (std)	0.059							
	(0.123)							
Rom. Cath., 1931 (std) * Kresy	-0.065							
	(0.147)							
Share Polish speakers, 1931 (std)		0.037						
		(0.146)						
Polish speakers, 1931 (std) * Kresy		-0.043						
		(0.179)						
Share Russian speakers, 1931 (std)			0.193					
			(0.229)					
Russian speakers, 1931 (std) * Kresy			-0.094					
			(0.235)					
Literacy rate, 1931 (std)				-0.008				
				(0.086)				
Literacy rate, 1931 (std) * Kresy				0.015				
				(0.097)				
Urbanization rate, 1931 (std)					0.046			
					(0.060)			
Urbanization rate, 1931 (std) * Kresy					-0.087			
					(0.057)			
Literacy rate, 1921 (std)						0.019		
						(0.082)		
Literacy rate, 1921 (std) * Kresy						-0.043		
						(0.094)		
Literacy rate Rom. Cath., 1921 (std)							0.027	
							(0.072)	
Literacy rate Rom. Cath., 1921 (std) * Kresy							-0.034	
							(0.085)	
Share Rom. Cath., 1921 (std)								0.077
								(0.114)
Rom. Cath., 1921 (std) * Kresy								-0.175
								(0.166)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	13.15	13.15	13.15	13.15	13.15	13.15	13.15	13.15
R^2	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
Observations	8,794	8,794	8,794	8,794	8,614	8,772	8,772	8,772

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, age2, dummies for six age groups, as well as indicators for rural places and urban counties.

Table A.6: No heterogeneous effects w.r.t. geographic features at origin

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)
	Years of education (based on 4 categories)			
Ancestor from Kresy	0.587***	0.475***	0.585***	0.555***
	(0.100)	(0.118)	(0.104)	(0.096)
Land suitability for wheat at origin (std)	-0.043			
	(0.082)			
Land suit. for wheat (std) * Kresy	0.015			
	(0.096)			
Annual temperature at origin (std)		0.027		
		(0.089)		
Annual temperature (std) * Kresy		-0.169		
		(0.115)		
Precip.-evatranspiration ration at origin (std)			-0.015	
			(0.065)	
Precip.-evatranspiration ration (std) * Kresy			-0.052	
			(0.099)	
Ruggedness at origin (std)				0.036
				(0.048)
Ruggedness (std) * Kresy				-0.091
				(0.083)
Controls‡	✓	✓	✓	✓
County FE	✓	✓	✓	✓
Mean Dep. Var.	13.15	13.15	13.15	13.15
R^2	0.29	0.29	0.29	0.29
Observations	8,794	8,794	8,794	8,794

Notes: The table shows that the coefficient on Kresy ancestry does not vary systematically with geographic characteristics 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, age2, dummies for six age groups, as well as indicators for rural places and urban counties.



Figure A.1: Migrants from CP were invited to come and explore available resources.

Note: The poster's title reads "The land is waiting." The text below the picture reads: "The State Repatriation Office is assigning farms in Opole and Lower Silesia. The regional inspectorates (offices) will provide all necessary information."

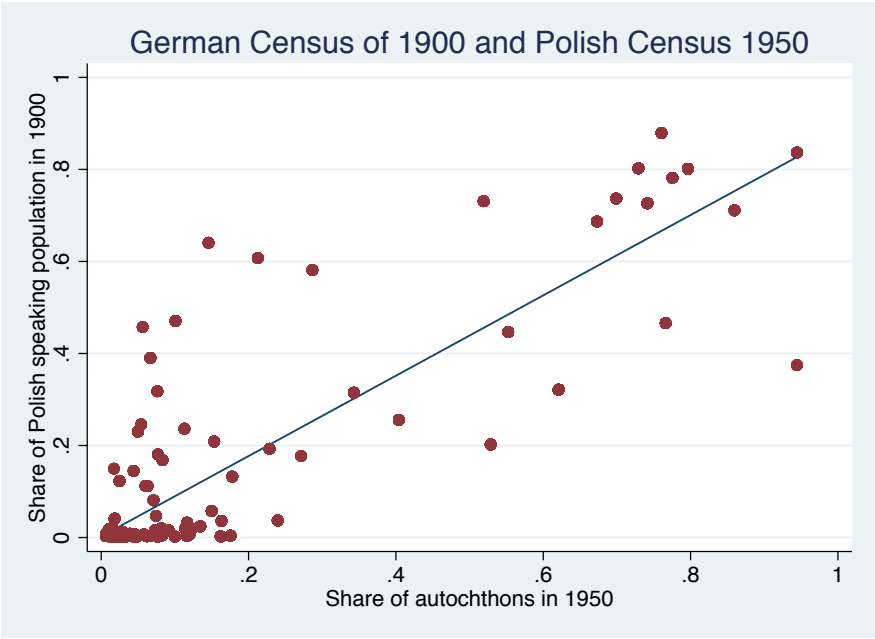


Figure A.2: Two alternative measures of the share of autochthons across WT counties

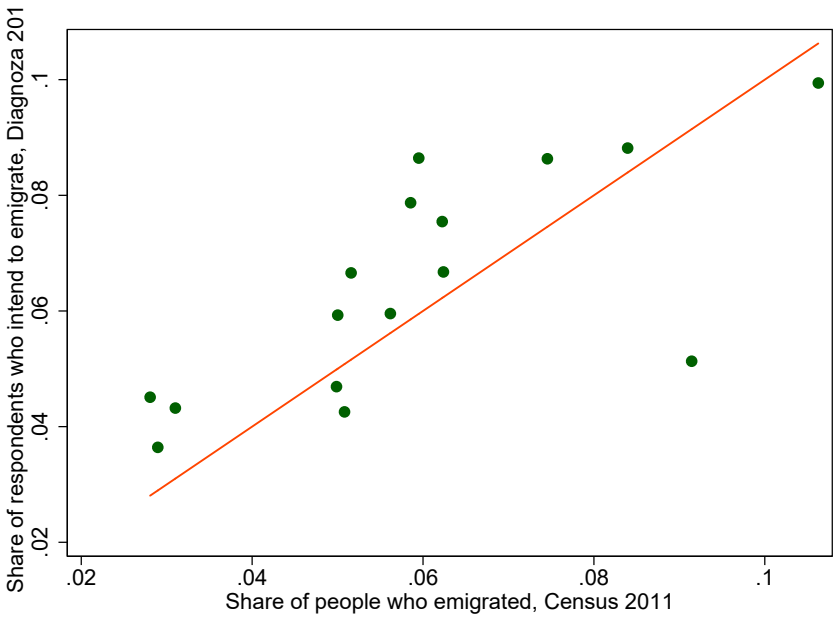


Figure A.3: Present intention to emigrate vs. Past emigration rates

Note: The figure plots the share of respondents who intend to emigrate (Diagnoza 2015) against the share of people who emigrated (Census 2011). Coef=0.65, SE=.18, R-sqrd=0.53. 45-degree line in red color.

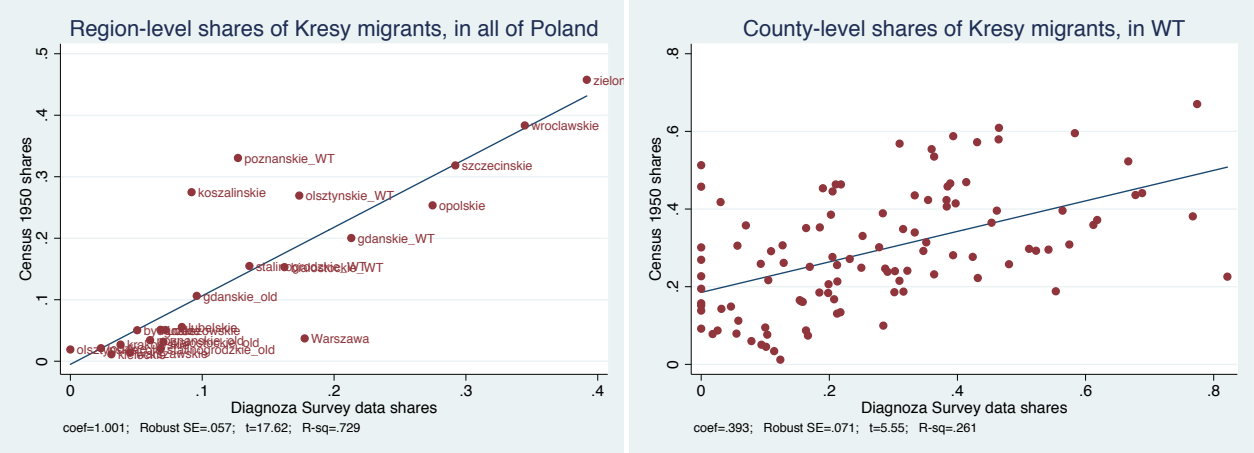


Figure A.4: Data quality check of Diagnoza Survey

Note :The left panel 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 (with separate observations for the parts of regions that were split by the border of Western Territories). The regression coefficient is 1.00 with a standard error of 0.057 and R^2 of 0.74. The right panel of 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 0.26.

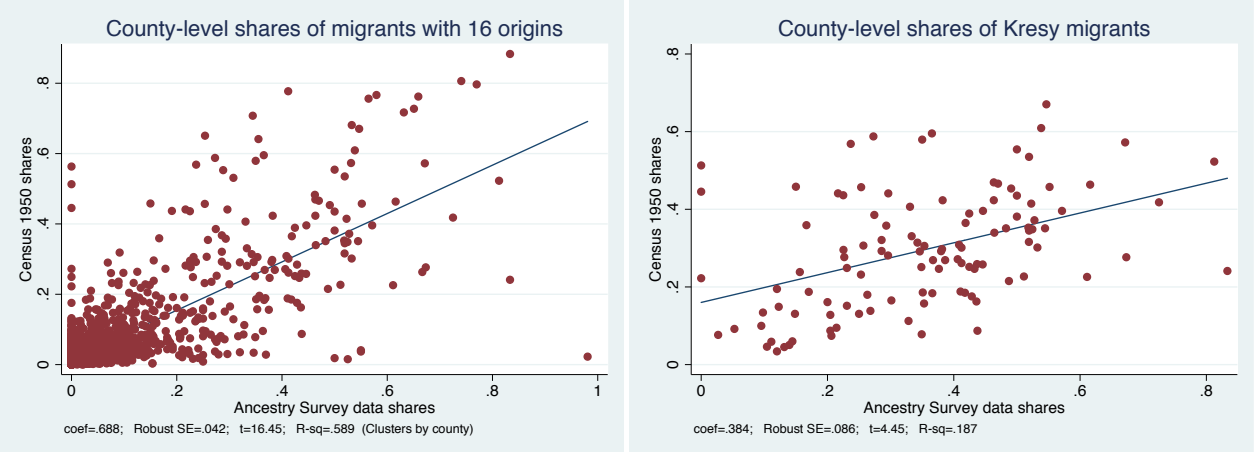


Figure A.5: Data quality check of our Ancestry Survey

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 0.69 with a standard error of 0.04 and R^2 of 0.59. The right panel repeats this exercise, but using only migrants from Kresy. The regression coefficient is 0.38 with a standard error of 0.09 and R^2 of 0.19.

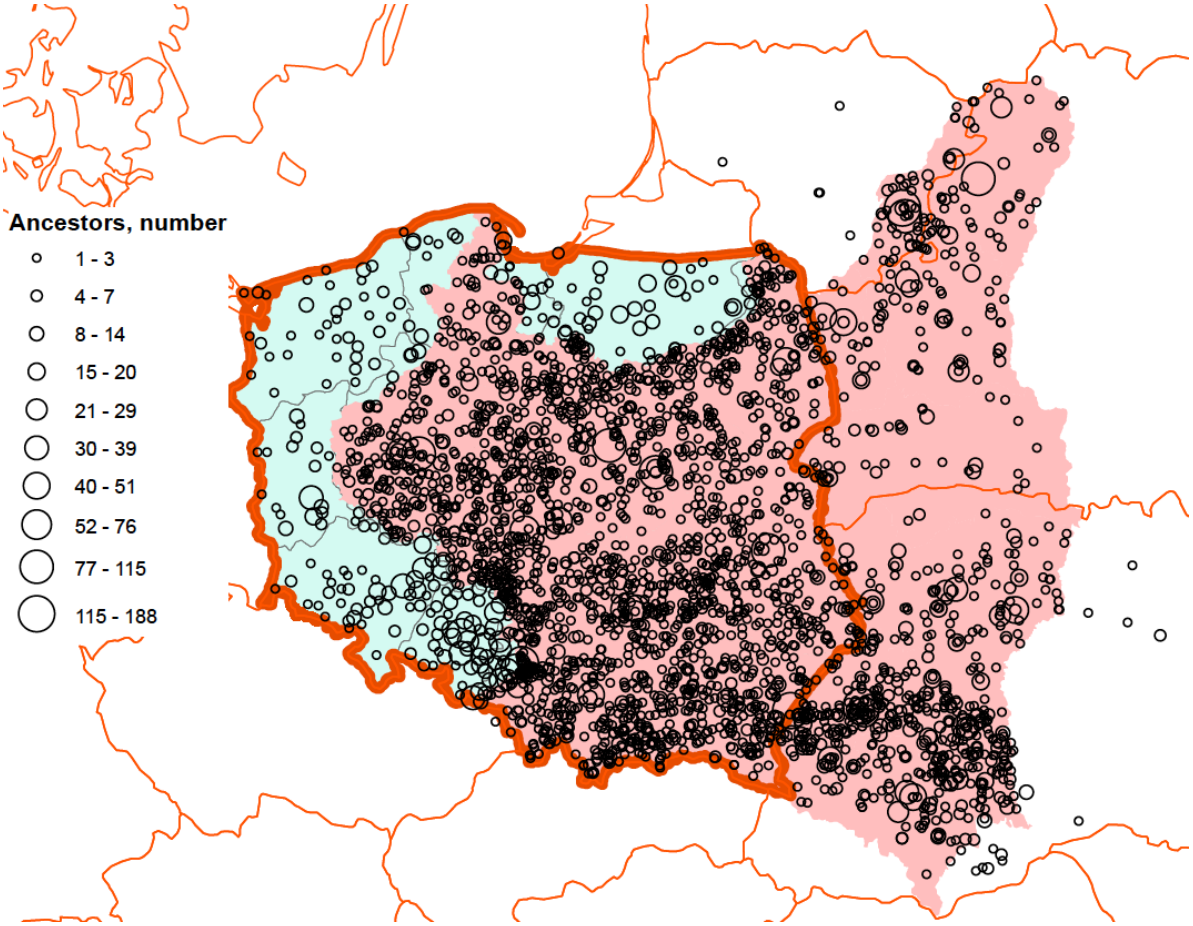
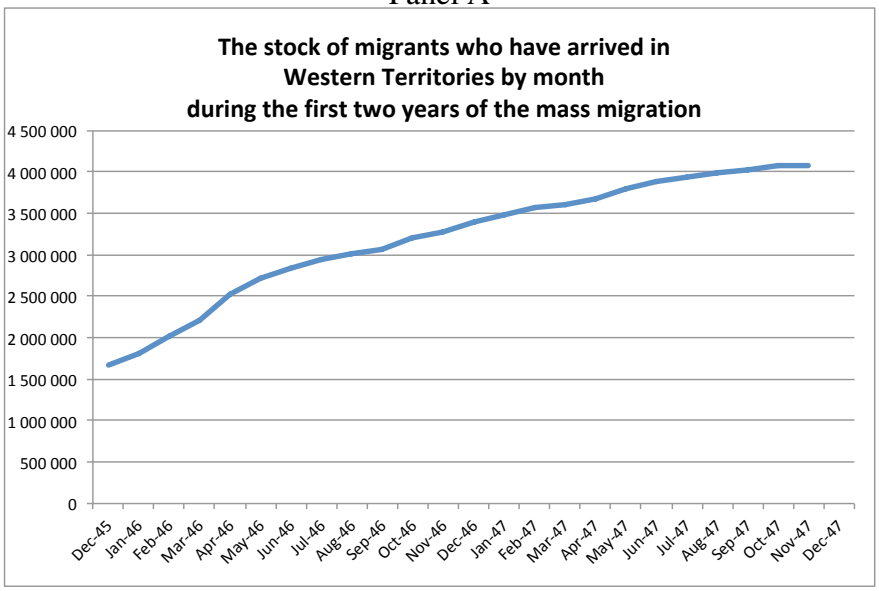


Figure A.6: Origin of ancestors in Ancestor Survey.

Note: The Figure displays the origin of ancestors in our Ancestor Survey.

Panel A



Panel B

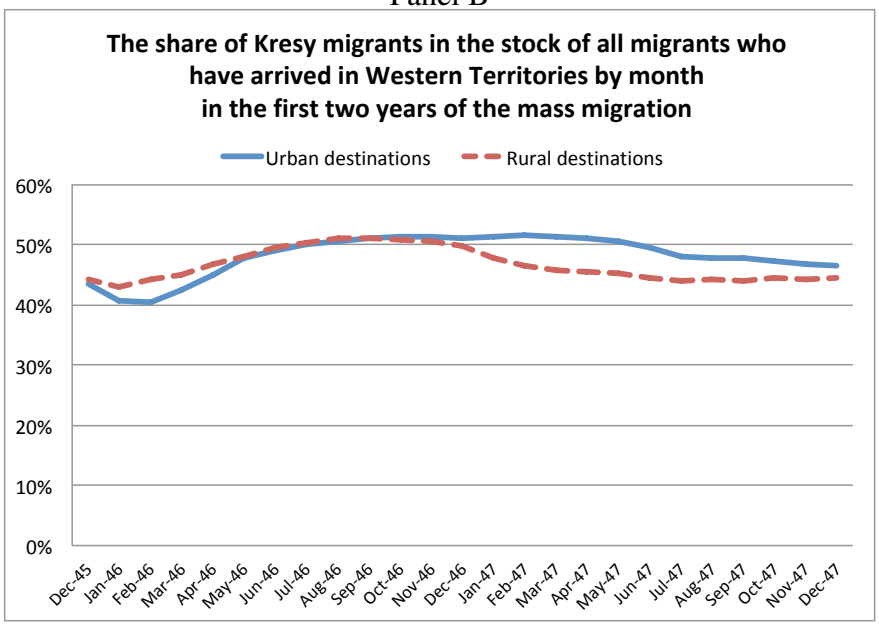


Figure A.7: The timing of arrival of migrants to WT

Note: The registry of migrants accounts for re-settlers from Central Poland and forced migrants from Kresy. The data come from the Document of the Ministry of Recovered Territories, No. 1661 (The Central Archives of Modern Records in Warsaw).