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: September 25, 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 Kresy border and include fixed effects for the destination of migrants. Since Kresy migrants were of the same ethnicity and religion as other Poles, we bypass confounding factors of other cases of forced migration. We show that labor market competition with natives and selection of migrants are also unlikely to drive our results. Survey evidence suggests 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 Boston University, the Chicago Booth Miniconference on Economic History, the NBER Political Economy Meeting, the Conference on 'Deep-Rooted Factors in Comparative Development' at Brown University, the 'Workshop in Political Economy and Economic Policy' at QMU London, the Oxford-Warwick-LSE (OWL) Workshop, the 'Culture, Institutions and Prosperity' conference in Paris, and seminars at Barcelona GSE, Bristol, Cambridge, Chicago Harris, DIW Berlin, EEA-ESEM 2018, Frankfurt, Harvard, PSE, Simon Fraser, UPF, and Warwick. We thank Samuel Bazzi, Luis Candelaria, and Giampaolo Lecce for insightful discussions, Ilona Kawalec from CBOS for outstanding collaboration, and Vladimir Avetian for excellent research assistance. 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).

"And so it happened that ... the marshall came: 'Leave' — 'But where should I go?'— 'To Poland.' And I say: 'I am in Poland.' And he says: 'This is not Poland anymore.' "[1]

1 Introduction

Forced migration is a life-changing experience, leaving deep scars in the memory of expellees. Does the experience affect also subsequent generations? In his bestselling autobiographical novel, Amos 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"*(p. 172; Oz, 2005). The idea that forced migration may affect preferences for education has been attributed by Stigler and Becker (1977) to Chicago economist Reuben Kessel. This argument has not only been made for Jews, but more generally – for example, in the academic literature 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 hard to test. Even for the most prominent case – that of the Jews – Botticini and Eckstein (2012) have convincingly challenged the idea that expulsion and discrimination are the main drivers of their educational lead.[2] 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, and religion. In addition, labor market competition with locals often affects educational choices of migrants.

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 when Polish frontiers were moved westward. 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 – the *Western Territories* (WT) – 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. In the East, Poles were forced to leave 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 who were

[1] Testimony cited in an exhibition of the Polish History Museum devoted to forced migrants from Kresy. See Appendix I for detail and sample photographs.

[2] They argue that 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.

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 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 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 the reversal in the education of Poles from Kresy and their descendants: Before WWII, when Poland consisted of the Kresy territories and Central Poland (CP), Poles in the former had, if anything, lower literacy rates. In contrast, in today's Poland, people with ancestors from Kresy have substantially higher rates of secondary education.[3] This receives further support when we examine the educational advantage of Kresy descendants by birth cohorts. 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.[4]

The Diagnoza Survey contains only information about ancestors from Kresy, but not about other ancestors. This may lead to recall bias if more educated respondents are more likely to remember the location of origin of their ancestors in general – and thus ancestors from Kresy in particular. To address this, we conducted an additional Ancestry Survey in 2016 in the Western Territories, where the majority of Kresy migrants were transported after WWII. We asked a representative sample of about 4,000 respondents about the origin location of *all* their ancestors (from the generation of youngest adults in 1939). Overall, we obtained the detailed location of almost 12,000 ancestors from all over Poland, as well as from Kresy. Using these data, we show that our results are not affected by recall bias. In addition, the Ancestry Survey allows us to compare the education levels of the descendants of forced migrants from Kresy, of voluntary migrants from Central Poland, and of Poles who had already lived in WT before the war (autochthons). 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 a stronger long-term impact on

[3] We use the share of people with a secondary degree in Figure 2 because it is comparable to literacy rates in 1921 in terms of its nationwide average. Among the respondents in our surveys, the rate of primary education is above 99%, without any meaningful variation. In our empirical analysis, we use years of schooling as the main outcome variable.

[4] The results shown in Figure 3 use fixed effects for respondents' county of residence today. These absorb potential local differences in the education system and in the labor market environment.

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 150 kilometers of either side of the Kresy border. At the same time, we include municipality fixed effects for the location of today's survey respondents. We find that among respondents who live in the same town or village today, those whose ancestors lived in Kresy, a few kilometers to the East of the Kresy border, have significantly higher education today than those whose ancestors lived in Central Poland, a few kilometers to the west of the Kresy border.

We examine several possible mechanisms that may drive our main result. These can be broadly structured into four categories: First, pre-existing differences – people from Kresy may have had higher education or different preferences for education already before WWII. Second, selection – people from either Kresy or from other parts of the country differentially selected into specific locations or occupations. Third, differential access – after moving, people from Kresy may have been given either preferential or restricted access to schooling or employment opportunities. Finally, preferences – the loss of physical possessions due to forced migration led to a change in attitudes towards the accumulation of human capital as opposed to physical assets. We present detailed historical and empirical evidence that the first three mechanisms are unlikely to account for the observed education premium of the descendants of forced Kresy migrants. In contrast, the fourth mechanism – a shift in preferences towards investment in human capital – is the most likely explanation for our findings.[5]

We support this interpretation 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 physical assets, relative to the number of physical assets they can afford. Historical narratives from the time of expulsions corroborate our survey evidence, suggesting a change in preferences towards education. The Western Institute in Poznan (*Instytut Zachodni*) collected memoirs written by re-settlers in Western Territories in the 1950s, some of which were subsequently published and analyzed by historians. For example, the memoir by a forced migrant from Kresy, who came from a simple peasant family, reads: *"In Western Territories, there was a specific situation. People did not attach great importance to material wealth. After all, nobody had it at that time ... most of the people who came here were still living in the memories of places of their origin and of material things that had belonged to their families for generations. In a new life situation, the cult of new values emerged, i.e., values that are indestructible, that cannot be lost,*

[5]At the end of the appendix, we provide a 'Guide to Mechanisms' with summary tables that schematize the evidence related to each possible mechanism.

and that die with the man – the cult of knowledge, of skills, which can resist cataclysms" (Bieniasz (1987), as cited in Halicka (2015), p. 262). The former president of Poland (2010-15) Bronisław Komorowski emphasizes in an interview how these values were passed on in families. *"I was born near Wrocław [the former German Breslau], in the house of grandparents Komorowski who had come there from Vilnius [in Kresy]... At home, nobody attached any importance to the material side, because everything that was valuable had been lost"* (Gazeta Wyborcza, 3 June 2017).

Our interpretation is consistent with recent evidence pointing to how preferences can adjust to shocks to environmental or institutional conditions and persist in subsequent generations. A robust body of evidence has described how individual preferences change in response to exposure to violence, natural disasters, or economic shocks.[6] Recent evidence suggests that these effects persist in future generations. Zhang (2018) documents systematic differences in preferences for competition among current Chinese high school students, depending on whether their grandparents or great-grandparents were exposed to state-imposed gender-egalitarian policies in the 1950s. Going back many more generations, Galor and Özak (2016) document systematic variations in time preferences among present-day populations related to changes in agro-climatic conditions during the Columbian exchange, more than five centuries ago.

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, and socio-economic effects on migrants themselves and on their descendants.[7] This literature studies both voluntary and forced migration.[8] Key drivers of forced migration are natural disasters, international wars, and civil wars.[9] Finally, a large body

[6] See Blattman and Miguel (2010) for a review of the literature on exposure to violence, and Voors, Nillesen, Verwimp, Bulte, Lensink, and van Soest (2012), Bauer, Cassar, Chytilová, and Henrich (2014), Cassar, Grosjean, and Whitt (2013), or Jakiela and Ozier (2018) for more recent contributions. On natural disasters, see Cameron and Shah (2015), Cassar, Healy, and von Kessler (2017), and Hanaoka, Shigeoka, and Watanabe (2018). On economic shocks, c.f. Giuliano and Spilimbergo (2014), Fisman, Jakiela, and Kariv (2015), and Malmendier and Nagel (2016).

[7] See the comprehensive discussions in Borjas (2014), Card and Peri (2016) and Dustmann, Schönberg, and Stuhler (2016) for short-run effects; for evidence of the long-run effects of migrants at their destinations see Hornung (2014), Peters (2017), and Murard and Sakalli (2018). Dustmann, Frattini, and Lanzara (2012) provide an overview of the literature on second-generation immigrants. Katz and Rapoport (2005) build a model that formalizes how forced migration can lead to a shift away from investing in physical capital toward investing in human capital.

[8] 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 Omar Mahmoud (2014) use forced migration to identify the effect of migration on economic outcomes at the destination.

[9] 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 (2018) exploit exogenous variation in mobility caused by public housing demolitions in Chicago. While neither paper finds effects on educational attainment, displaced children have better labour market outcomes as adults.

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).[10]

Our focus is on the long-term effects of forced migration after WWII, 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 West Germany. They find that migrant children tend to acquire more education than their native peers. The main mechanism behind this finding 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 education or other outcomes for later generations.

Relative to the existing literature, we make several contributions. First, we test the prominent hypothesis, untested by the previous literature, that uprootedness leads to human capital investment. Second, we analyze the hitherto unstudied 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. Third, we break new ground by studying the long-run effects of forced migration on the descendants of migrants over several generations. This is of high policy relevance in a world with large waves of forced displacement. Finally, our results suggest that caution is warranted in the prominent approach that uses forced migration as an instrument to estimate the effect of voluntary migration – this instrument may have a direct effect on outcomes via a change in preferences.

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 the two surveys. Section 5 examines possible mechanisms, and Section 6 concludes.

[10] 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.

2 Historical Background

2.1 The Change of Poland's Borders

Poland's Borders before 1945

Poland's borders have seen several changes over the last 500 years. The Polish Lithuanian Commonwealth (PLC) was established in 1569 when the Polish Kingdom and the Grand Duchy of Lithuania formed a union that lasted for over 200 years. In 1795, Poland lost its statehood as its territory was split among three European empires: Russia, Prussia, and Austro-Hungary. No sovereign Polish state existed until 1918; this period of Polish history is known as the 'Partitions of Poland.' At the end of World War I, the independent Polish state was recreated as the Second Polish Republic (SPR). Poland ceased to exist again as an independent state at the beginning of WWII, when Nazi Germany and the Soviet Union invaded the SPR in September 1939, splitting it according to the Molotov-Ribbentrop Pact.

Poland's Borders after 1945

At the end of Word War II, an independent Poland reemerged within new borders that moved Poland 200 kilometers to the West. 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). At the same time, Poland lost the Eastern Borderlands, known as *Kresy*, located to the east of the Curzon line (more detail below). The Kresy territory was divided among the Soviet Republics of Lithuania, Belorussia, and Ukraine. 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 by a thick red line.[11]

Henceforth, we refer to the part of Poland that belonged to the Second Polish Republic before WWII and continued to be Polish after WWII as 'Central Poland.' Thus, the territory of Poland before WWII was comprised of Central Poland and Kresy, whereas the Polish territory after WWII is comprised of Central Poland (CP) and WT. The 1931 Polish Census – the last census of the SPR – counts about 3 million ethnic Poles in Kresy. Before WWII, according to the 1939 German census, 8.8 million people lived in areas that after WWII became the Polish Western Territories. Almost 90% of them declared to be 'German,' 10% were Poles, and about 1% Jews (Dziewanowski, 1977).

[11]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.

Arbitrariness of the Kresy border of 1945

The Kresy border (i.e., the post-WWII Eastern border of Poland) was established roughly along the Curzon line after many discussions between Stalin and the Allies. The Curzon line, in turn, had been suggested as an armistice line in a note by British Foreign Secretary Lord Curzon during the 1920 Polish-Soviet conflict – a suggestion that was then disregarded by both Poland and the Soviet Union. The 1921 Treaty of Riga instead provided Poland with land that – on average – was about 250 kilometers eastward of the Curzon line. The Curzon line also did not correspond to the border between Germany and the Soviet Union according to the terms of the Molotov-Ribbentrop Pact of 1939; after the military defeat of Poland in September 1939, the Soviet Union annexed territories extending well to the west of the Curzon line – as far as Lublin and Warsaw. Nor did the Curzon line correspond to any natural border: There is no discontinuity in geo-climatic characteristics such as precipitation, temperature, elevation, terrain ruggedness, or in suitability to various crops (see Appendix V).

After recapturing Eastern Poland from Germany in 1944, the Soviets unilaterally declared the new border between Poland and the Soviet Union approximately along the Curzon line, to which the Allies ultimately conceded at the Yalta conference. Historians of Poland agree that the post-WWII border between Poland and the USSR, which we henceforth refer to as Kresy border, was arbitrary. For example, Davies (1981, p. 493) writes: *"All decisions regarding the Polish frontiers were taken ad hoc[...] No attempt to trim the frontiers to the wishes of the population ever succeeded, [...] it was decided in 1944–5 to trim the population to the requirements of arbitrary frontiers."*

Poles in Kresy and Central Poland before WWII

In the context of our study, a relevant question is whether Poles from Kresy were exposed to radically different experiences than Poles from other regions already before WWII. In the two periods when Poland was a sovereign state – PLC in 1569-1795 and SPR in 1918-39 – Poles had the same rights in all parts of the country. Namely, Poles who lived in what later became Kresy and Poles who lived in what later became Central Poland had exactly the same status (Davies, 1981). In contrast, during the Partitions of Poland, the living conditions and the rights of Poles differed across the three empires (e.g., Davies, 1981; Grosfeld and Zhuravskaya, 2015). The Russian and the Austro-Hungarian Partitions stretched over parts of Kresy and parts of Central Poland. *Within* these two Partitions, Poles had the same rights irrespective of whether (or not) they lived in Kresy.[12]

[12]Below, in Section 5.5, we show that our results hold when we restrict the sample to ancestors who lived within the former Russian Partition of Poland (which covered about three quarters of the Kresy territory and one-half of the territory of Central Poland). The Prussian Partition did not include any part of Kresy, and the Austro-Hungarian Partition covered about one-quarter of Kresy.

Overall, Poles in Kresy faced differential discrimination (as compared to Poles in other parts of Poland) only once – when they were forced to move from Kresy at the end of WWII.[13]

2.2 Post-WWII Mass Population Movements

As a result of the change of borders after WWII, mass population movements occurred. At the end of WWII, an estimated 2.5-3.4 million Germans (who had not fled as the Red Army advanced), and 1 million Poles were still located in the Western Territories (Dziewanowski, 1977). The remaining Germans were expelled from WT 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 following the change of the borders, according to which Poles from Kresy were forced to resettle within the new Poland, while Ukrainians, Belorussians, and Lithuanians had to leave Poland and resettle in the USSR. These mass movements of people began in 1944 and were largely completed by 1948 (e.g., Schechtman, 1962; Eberhardt, 2003).

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 of Ukraine, Belorussia, and Lithuania (Ciesielski, 1999). The official language of these agreements did not explicitly specify that the ethnic groups in question were to be expelled from the two respective sides of the Curzon line. However, historians agree that the members of these groups had no viable alternative but to move – this was also driven by the Polish and Soviet authorities seeking to quickly create irreversibility, by moving populations according to the new frontiers (c.f. Davies, 1981; Kersten, 1986).

Forced migration from Kresy territories

By 1950, 2.1 million Poles had been forced to move from the Kresy territories. The Polish State Repatriation Bureau tried to ensure an orderly movement of Poles from Kresy directly to WT. However, this was hard to implement because of the war-related devastation, destruction of infrastructure, and the lack of adequate transport. Approximately one quarter of Kresy migrants settled

[13] A related question is whether Poles in Kresy and Central Poland were differentially exposed to ethnic conflict. During most of the history of PLC, Poles coexisted relatively peacefully with other religious and linguistic groups that lived in the territory of Poland (Davies, 1981). Occasionally, however, PLC did experience open ethnic conflicts. The most violent event in PLC broke out during Bogdan Khmelnitsky's uprising in the middle of the 17th century, when Ukrainians turned against Poles and Jews. Thereafter, there was no major episode of ethnic violence until 1939, with the notable exception of a series of anti-Jewish pogroms perpetrated by Poles following the death of Polish leader Józef Piłsudski in 1935 (these pogroms took place mostly in the Southern part of Central Poland). The largest conflict involving Poles (which was not carried out by the Nazi occupiers) erupted towards the end of WWII. In 1943-44, under the occupation by Nazi Germany, the Ukrainian Insurgent Army perpetrated mass killings of Poles, known as Volhynian massacres (Snyder, 2003). One of the objectives of Ukrainian nationalists was similar to the one later implemented by Stalin – to purge Poles from Ukraine. This was an unprecedented event, which foreshadowed the forced expulsions by Stalin after WWII.

in Central Poland, many of whom had family ties there. The aim of the Polish authorities was to resettle Kresy deportees in those places in the 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. Each family was allowed up to two tons of luggage.

Even though historians agree that Poles were forced to move from Kresy, not everybody left during the post-war population exchange.[14] Some of the remaining Poles in Kresy 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 high levels of animosity between Poles and Ukrainians at the end of WWII (e.g., Ciesielski, 1999). In all three Soviet republics, pressure on the urban Polish population to leave was high. We exploit the urban vs. rural and Ukraine vs. rest-of-Kresy variation below.

Voluntary migration from Central Poland

Despite war-related destruction, land, housing, infrastructure, and capital stock were abundant after the expulsion of the German population from WT. Before the war, these territories had been densely populated. This made the Western Territories an attractive destination for voluntary migrants from Central Poland, who sought a better fortune than in their homeland: Deprivation and poverty were the main drivers of migration from Central Poland (Zaremba, 2012, p. 97). The flow of migrants from CP started as early as in the spring of 1945. Some of this early voluntary migration was spontaneous (mostly from the neighbouring Polish areas, sometimes on foot, or by horse carts and trucks), some was triggered by an advertising campaign organised by the Polish authorities that promoted a move to WT in order to populate the newly acquired land as quickly as possible. The campaign advertised the Western Territories as the land of abundant resources. Figure A.3 in the appendix presents an example of a poster dating back to this campaign. As a result of this campaign, voluntary migrants came to WT from all over Central Poland, many of whom traveled long distances by train (Zaremba, 2012).

[14]The Socialist propaganda tried to suggest that "repatriates" returned voluntarily to their "mother country." As Ther (1996, p. 783) points out, the contrary was true. Kresy Poles regarded Eastern Poland as their *mała ojczyzna* (homeland). *"These 'repatriates' did not return to their home country but were forcibly relocated to the former territories of a foreign country."* Ther goes on to explain why the Socialist propaganda was convenient also for Western politics: *"One possible explanation for the success of Eastern propaganda can perhaps best be described as 'bad conscience.' Since Winston Churchill and Franklin D. Roosevelt had agreed to the expulsion of the Eastern Poles without even consulting the Polish government, the West was prone to accept a rosy version of Polish postwar history."*

Aggregate statistics on mass population movements

The first reliable post-WWII population census in Poland took place in 1950.[15] In addition to statistics typically collected during censuses, it provides information about the mass movements of the Polish population by asking about the place of residence before September 1, 1939.[16] Table 1 reports aggregate statistics from the 1950 census about the origin of the Polish population, separately in the Western Territories and in Central Poland. Of the total 24.6 million Polish population in 1950, 23% (5.6m) lived in WT. Within the Western Territories, about 50% (2.8m) came from Central Poland, 28% (1.6m) came from Kresy, and 20% (1.1m) were autochthons, i.e., Poles who had lived in WT 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% (about 583,000) came from Kresy. Very few inhabitants of CP came from WT or from abroad (0.1% and 0.3%, respectively).

Importantly, post-WWII Poland was largely an ethnically and religiously homogeneous country, composed of ethnic Poles of Roman Catholic faith that differed only in their pre-WWII region of residence. The Western Territories, as the rest of Poland, were also ethnically and religiously homogenous: According to the 1950 census, 96% of WT's 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 had come to WT right after the war), and Ukrainians, who were forced migrants from CP to WT (during the so-called "Operation Vistula" – c.f. Snyder, 1999).

The arrival of migrants to the Western Territories

Upon arrival to Western Territories, Poles (irrespective of whether they came from Kresy or Central Poland) were allocated land, housing, and capital that expelled Germans had left behind. In rural areas, this primarily meant houses, land plots, and agricultural machines; in urban areas – apartments, townhouses, shops, and office buildings. Initially, the Polish administration was very weak and operated under conditions of chaos, confusion, and lack of rules. There was no register of available properties, and people were more or less free to find and claim a place. During this first period, the capital goods left by Germans were distributed on a first come, first serve basis. *"The Polish settlers were searching houses that were available and not reserved for other Poles...*

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

[16] In particular, respondents indicated if they had lived within the post-WWII Polish borders, and if so, in which region (voivodeship). If in 1939, respondents had lived outside the borders of post-WWII Poland, they had to indicate the country in which their 1939 place of living was located in 1950. Thus, forced Kresy migrants indicated that they lived in the USSR before the war.

they were registering them with the local administration and – if there were no counter-arguments – could settle there" (Halicka, 2015, p. 203). When institutions and the Polish administration became stronger, authorities began to organize the distribution of land and capital. The arrival of migrants in WT coincided with the land reform in 1944-48. Migrants to rural areas typically got lots of 8-10 hectares per family; larger estates were parceled out among several families (Davies, 1981, p. 559). The peasants became owners of their land for an equivalent of a one-year harvest payable in several installments. Large farms of more than 100 hectares in WT (and more than 50 hectares in Central Poland) were transformed into State Agricultural Farms. After 1956, de-collectivization reduced the number of collective farms but they remained important in WT. The houses and flats left by Germans were nationalized, and settlers got lifetime rental contracts.

Forced Kresy migrants and voluntary migrants from Central Poland arrived to the Western Territories at the same time. They were treated equally upon arrival because all migrants helped to achieve the main objective of the Polish authorities – to populate the Western Territories as quickly as possible (Schechtman, 1962, p. 213). As the deputy minister of Public Administration wrote to the Central Party Committee in May 1945: "*The assessment of the organizational capacity of the Polish Nation abroad, and the security of our Western borders, will depend upon our capacity to populate and develop the area in the West*" (cited in Halicka, 2015, p. 184). The Ministry of Recovered Territories collected statistics on the rates of arrival of migrants by month during 1946 and 1947. Figure A.4 in the appendix visualizes these data, showing that the share of migrant inflow from Kresy was about 40-50% throughout this two-year window. By the end of 1947, the Kresy migration of the first repatriation wave came to an end.

2.3 Uncertainty Perceived by Kresy Migrants and its Connection with Education

Historical and journalistic accounts of re-settlements into Western Territories suggest that forced migrants perceived a higher degree of uncertainty than other settlers or autochthons. The fate of the Western Territories was viewed as uncertain by its inhabitants because of the lack of a legal guarantee for the Polish-German border.[17] The prominent Polish sociologist Zdzisław Mach describes this in an interview to the leading Polish newspaper *Gazeta Wyborcza*: *"Settlers did not feel that the land they found was given to them forever. Until the 70s it was not certain that the Western Territories would remain part of Poland. ...Władysław Gomułka [the first Communist Party secretary] ...did not invest in the Western Territories because at heart he was not sure what would happen to them... It is not a random expression that the first generation of re-settlers were living*

[17]Until 1950, a mere memorandum of the Potsdam Conference guided the demarcation of the border along the Oder-Neisse line. In 1950, East Germany and socialist Poland signed the first bilateral treaty legalizing the Oder-Neisse line. In 1970, a similar treaty was signed between West Germany and Poland. The final treaty was signed by Poland and the unified Germany in 1990, and it was ratified by the Polish Sejm and the German Bundestag in 1991.

on suitcases. They never felt sure and secure..." (*Gazeta Wyborcza*, Dec 29, 2010). The perceived uncertainty differed depending on past experiences: Because of their traumatic experience of the previous expulsion, settlers from Kresy were particularly worried that Germany would take over the Western Territories (see, e.g., Zaremba, 2012).[18] Magdalena Grzebałkowska – a journalist and author of *"1945. War and Peace,"* a book based on testimonies of descendants of re-settlers to the Western Territories – was herself born in WT; her grandparents had been forced to move from Kresy. In her book, she reflects on her own experiences growing up: *"As a child, I was worried that if something is postgerman, at some point it may become 'postpolish.' Unconsciously, I inherited the fear of my ancestors-settlers that the place where I live is given to us just for a moment"* (Grzebałkowska, 2015, p.72).[19] In an interview with the authors of this paper (conducted on May 9, 2018), Grzebałkowska confirmed that the experience of forced migration had an important effect on the perception of uncertainty, which in turn is related to education decisions: *"Unlike migrants from Central Poland who always had an option of going back to Central Poland (and some actually did go back),... forced Kresy migrants got the 'one-way-ticket' and lost everything... when you lost everything, it seems worth investing in yourself, getting more education."*

There were no systematic studies of educational attainment of re-settlers in the Western Territories by ancestors' origin before our paper. One important reason is that studies of forced migration from Kresy were prohibited in socialist Poland. Nevertheless, the scattered evidence that is available underlines the focus on education – even among the first generation of Kresy migrants. For example, the sociologist Irena Turnau assembled data on schooling in Wrocław (the former German Breslau) in 1948. She found that children of Kresy migrants were over-represented among secondary school students, and even more so among students in higher education.[20]

3 Data

We use numerous data sets for modern-day and historical Poland in our analysis. To capture modern-day educational attainment at the individual level, we use two surveys that also ask questions about the history of migration of respondents' ancestors in the aftermath of WWII. We complement these surveys with aggregate (regional and county-level) data from historical censuses that describe population characteristics in Poland before and after WWII. We describe each of

[18] A popular saying illustrates how Kresy migrants perceived the ambiguous status of the Polish borders before the 1970s: *"One atomic bomb, and we will again return to Lviv... A second one... we will be back to Vilnius."* (see, e.g., Zaremba, 2012).

[19] "Postgerman" (*poniemieckie*) is an actual word in Polish, which refers specifically to land and assets in the Western Territories that were taken over by Poles from Germans after WWII.

[20] In Wrocław overall, 22% of the total population was born in Kresy. Among secondary school students, 27.5% had roots in Kresy, and among students in higher education, 36.5% came from Kresy (numbers reported in Turnau, 1960, pp. 31-33). These numbers have to be interpreted with caution because they combine different surveys.

these data sources in turn.

3.1 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.[21] We commissioned the addition of several questions to the 2015 wave, which inquired whether any of the ancestors of the respondent came from Kresy and if so, from which exact location.[22] 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 ancestors from Kresy. Our primary interest is investment in human capital. Thus, we focus on the education of respondents, using years of education as our main variable of interest. We also use two dummies for educational attainment: i) for having (at least) secondary education and ii) for (at least) completed higher (tertiary) education. 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. In addition, to measure the actual choices in accumulation of physical wealth, we construct a variable that measures the share of physical assets that respondents chose *not* to own, despite the fact that they could afford them.

A drawback of the Diagnoza Survey is that it only includes information on ancestors from Kresy, but not for ancestors from other areas, such as from Central Poland. Our Ancestry Survey fills this gap.

3.2 Ancestry Survey

In 2016, we conducted our own survey in the Western Territories, which had seen the largest inflow of Kresy migrants after WWII. We asked a professional survey company to draw a representative sample of the population in the Western Territories (3,169 respondents), as well as an additional representative sample of people in WT with Kresy origin (900 respondents).[23] 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

[21] For further detail on the survey see http://diagnoza.com/index-en.html (accessed on September 24, 2018).

[22] The main question was "Is there anybody in your household who himself or his parents or grandparents 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. 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 Ancestry Survey (see below).

[23] 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 the appropriate sampling weights to adjust for oversampling.

parents, we asked where 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 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. Overall, the 4,069 respondents gave information about 13,223 ancestors. The vast majority of respondents knew the exact name of the locality of origin of their ancestors (and not only the broad region of origin), even when the generation of youngest adults in the family was the great-grandparents. This highlights the importance of the mass population movements in the family histories of Poles.[24] Overall, we were able to identify and geo-code the place of residence of 11,928 out of 13,223 ancestors.

We report summary statistics for the Diagnoza Survey and for our Ancestry Survey in Appendix II, Tables A.1 and A.2. Figure A.5 in the appendix displays the origin of ancestors in our Ancestry Survey.

3.3 Historical Censuses

Post-WWII Polish Census 1950 – The Polish Census in 1950 contains information on population movements. It asked in which Polish region or in which country people lived before WWII (according to post-WWII borders – thus, people who lived in Kresy in 1939 had to answer 'USSR'). In the Western Territories, this information is available by county (powiat) of residence; in Central Poland, it is available by region (voivodeship) of residence, and for Kresy there is no further detail (since all of Kresy falls under 'USSR').

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 and shares of population with different languages and religions by locality, but without cross-tabulations of the data. The 1921 Census, in contrast, has literacy rates by religious denomination, allowing us to measure the literacy rates among Roman Catholics. This is a close proxy for the literacy of ethnic Poles because in the SPR, only Poles were Roman Catholics; other groups had other religious affiliations, such as Orthodox Christians, Greek Catholics, and Jews.

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

[24] In our survey, we were able to monitor the interview process, and we were impressed by how survey respondents engaged with the questionnaire. Most 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.

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

3.4 Benchmarking Survey Data against Historical Census Data

While we have no way to confirm the accuracy of ancestors' location provided by *individual* respondents, we can benchmark the survey responses against the information on post-WWII migration given by the 1950 Polish census. The latter provides both the pre-WWII location (see above) and the location in 1950; these two pieces of information can be used to construct migration movements in 1939-50. We compare these with population movements implied by the survey responses – i.e., the difference between respondent location in 2015/16 and the location of origin of their ancestors in 1939. We find that the ancestry information from the two surveys lines up well with the 1950 Census data – despite the fact that the former includes post-1950 movements, while the latter does not. Appendix III provides further detail and presents the graphs summarizing the data comparisons in Figures A.6 and A.7.

4 Empirical Results: Forced Migration from Kresy and Education

Our analysis relates modern-day education outcomes to the place of origin of respondent's ancestors. We use our two individual-level data sets – the Diagnoza Survey and our Ancestry Survey. 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 and no indication of origin of ancestors from regions other than Kresy. Our Ancestry Survey fills this gap, by collecting information on all ancestors from the generation that was affected by the post-WWII population transfers. One caveat is that the Ancestry Survey is run only in the Western Territories (where most Kresy migrants resettled). This potentially raises concerns about selection of voluntary migrants to WT. We discuss this in detail in Section 5. Overall, Diagnoza and our Ancestry Survey can be seen as complements: The former allows us to compare descendants of forced Kresy migrants to all other Poles, so that selection of the 'control group' is not an issue. The latter includes more detailed information on ancestors by focusing on the area that saw the largest inflow of migrants – Western Territories. The main results in both surveys are almost identical, suggesting that neither missing detail on non-Kresy ancestors in Diagnoza, nor selection of voluntary migrants in the Ancestry Survey, confound our results.

In both surveys, we estimate the following regression 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, $Kresy_i$ is a dummy variable that takes on value one if *any* ancestor was from Kresy. When using our own Ancestry Survey, we can compute $Kresy_i$ as the *share* of i's 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 the 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 fixed effects for counties (*powiat*) or municipalities (*gmina*). There are 377 counties and 1,726 municipalities in the Diagnoza sample, and 115 counties and 407 municipalities in our Ancestry Survey. Because in Diagnoza several respondents may come from the same household, we cluster the error term ε_i at the household level.

4.1 Diagnoza Survey Results

Using the Diagnoza Survey, Table 2 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 commonly used outcome variable 'years of education.' In columns 1 and 2 we examine the full sample, with approximately 25,700 respondents (out of whom almost 3,000 had Kresy ancestors). The results in column 1 show that Kresy ancestry is associated with 0.97 additional years of schooling, relative to an average of 11.95 years. When we include county fixed effects in column 2, the coefficient on Kresy ancestry is almost unchanged (0.93 extra years of schooling). This suggests that our results are not affected by spatial sorting of migrants, or by local characteristics such as labor markets or land quality. In what follows, we refer to column 2 (i.e., including county fixed effects) as our baseline specification. Next, columns 3 and 4 restrict the sample to respondents in rural and urban areas, respectively. The coefficient on Kresy is somewhat larger in urban areas (where average educational attainment is also higher). In addition, the results are broadly similar for respondents in Central Poland and the Western Territories (columns 5 and 6). In other words, the descendants of forced migrants from Kresy enjoy an educational advantage everywhere in Poland.

In Panels B and C the dependent variable is an indicator for secondary and higher education, respectively. In our baseline specifications in column 2, we find that descendants of Kresy migrants are 12.7 percentage points more likely to finish secondary education (relative to a mean of 51%), and 9.9 percentage points more likely to graduate from college (relative to a mean of 21%). Thus, in *relative* terms, the association between Kresy origin and education is strongest for higher educa-

[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. In all regressions that use education as the outcome variable, we exclude respondents who were still students by the time of the surveys in 2015 and 2016.

tion. Table A.3 in the appendix shows that the results from columns 2-6 are almost identical when using municipality (instead of county) fixed effects. Thus, even when comparing Poles who live in the same town or city today, those with Kresy ancestors have a substantial education advantage.

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 kilometers on each side of the border between Kresy and Central Poland. 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 kilometers to the east of the Kresy border.[26] As for the area within 150 kilometers 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.

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. In contrast, the right panel of Figure 4 shows that there is a sharp discontinuity at the border, with today's education jumping by about one year. This confirms that Kresy descendants have substantially higher education levels, even in a subset of individuals with ancestors from close to the Kresy border.[27]

Column 7 in Table 2 complements Figure 4, presenting regression results for the border sample. 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. The regressions in column 7 include all our standard controls, as well as a quadratic polynomial in latitude and longitude to capture unobservables that may vary around the Kresy border (Dell, 2010).[28] Turning to the results, we find positive and significant coefficients for Kresy ancestors for

[26] Whenever a respondent gave the location of more than one Kresy ancestor (see footnote 22), we make a conservative choice – using the maximum distance to the Kresy border among all ancestors.

[27] In Appendix V (Figures A.8 and A.9), we show that the Kresy border was unrelated to natural boundaries. This complements the historical discussion on the arbitrariness of the Kresy border in Section 2.1.

[28] Following the argument in Gelman and Imbens (2014) that cubic and higher-order polynomials can yield misleading estimates, we use a second order polynomial. Note that we do not include fixed effects for respondents' location,

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 lived in 1939. This may be a subsample with particularly vivid memories of the forced migration experience, augmenting the long-run effects on education.

Note that our Diagnoza border sample analysis has an important shortcoming: Individuals with Kresy roots now largely live in the Western Territories, far away from the historical Kresy border. Our border analysis compares them to individuals who still live close to the Kresy border today (to its west). That is, we compare respondents who potentially live far apart today, although their ancestors were likely living close to each other. We address this limitation in Section 4.3, using the more detailed data from our Ancestry Survey.

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). For years of schooling in this cohort, we find a small negative (and insignificant) coefficient on Kresy ancestry. The same is true for higher education in Panel C. Panel B shows a very small and insignificant positive coefficient on Kresy ancestry for secondary education. In other words, in the cohort 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 educational differences or by selection of educated migrants from Kresy.

Columns 2-8 in Table 3 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 highly significant throughout and 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 14% higher years of schooling.[29] For later birth cohorts, the coefficient size

because these would absorb the variation in distance to the west of the Kresy border. This is because we use today's location of respondents from CP (i.e., those within 150 km to the west of the Kresy border) as a proxy for their ancestors' place of living. Below, we address this issue by using data from our Ancestry Survey, which includes many respondents whose ancestors lived in CP close to the Kresy border, but who themselves live scattered throughout the Western Territories today.

[29]Historical accounts suggests that the supply of schools was well organized as early as 1946, even in the Western

declines continuously. This makes sense in the context of our hypothesis that forced migration led to a shift in preferences towards education: The intergenerational transmission of preferences is not one-to-one, even when taking into account local peer effects and assortative mating of parents (c.f. Dohmen, Falk, Huffman, and Sunde, 2012).[30]

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 3-4 show that people with Kresy ancestors are more likely to have white collar occupations; at the same time, they are less likely to be unemployed (columns 5-6). These results remain statistically significant even after we control for education, but the coefficients on Kresy origin become smaller in magnitude.

4.2 Ancestry Survey Results – Respondent Level

We now turn to our Ancestry Survey, which has information on the origin of *all* ancestors in a respondent's family tree, for the generation of the youngest adults at the beginning of WWII. We use this information to compare the descendants of forced migrants from Kresy with descendants of voluntary migrants from Central Poland, and with autochthons. We also perform numerous robustness checks of our results. Compared to Diagnoza (which covered all of Poland), a limitation of the Ancestry Survey is that it only covers respondents who live in the Western Territories – where most migrants from Kresy were transferred to. This potentially raises concerns about selection of other people who voluntarily migrated to WT (i.e., the 'control group' in our regressions). We conduct various analyses to show that such selection is unlikely to confound our results. Before moving to these checks and the interpretation of our results in Section 5, we show that our main results also hold in the Ancestry Survey.

We use the detailed information on ancestor origins in our Ancestry Survey to compute, for each respondent, the *share* of ancestors from Kresy (average 23.6% in the representative sample of WT population), as well as from Central Poland (avg. 60.5%), autochthons from 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).

[30]Note that among the 1990s birth cohort, many respondents were still in school/university by the time of the Diagnoza Survey in 2015. Since we exclude respondents who were still students in all regressions where education is the dependent variable, the mean of the education variables is lower in column 8, and the coefficients on Kresy need to be interpreted with caution.

Territories (avg. 15.9%), and from abroad – other than USSR (avg. 1.3%). We use the share of ancestors from Kresy as our main explanatory variable in equation (1) and add the shares of autochthons and Poles who lived abroad in 1939 as controls. Note that the share of ancestors from Central Poland is thus the reference group. In addition, we also control for the share of each respondent's ancestors who came from rural *origin* locations to capture possible differences between migrants from rural and urban areas.

Column 1 in Table 5 shows that the share of ancestors from Kresy is associated with significantly higher levels of education among respondents. This coefficient 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.[31] The coefficient on Kresy changes very little when we control for the shares of ancestors from other origin locations (column 2). The negative coefficients on the share of ancestors from WT show that in the Western Territories (where the Ancestry Survey was conducted) autochthons have lower education levels as compared to the control group – descendants of migrants from Central Poland. 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 migrants, voluntary migrants, non-migrants. 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. We find that the results are remarkably similar for both subsamples, and, if anything, slightly larger in urban areas, which confirms the Diagnoza results from Table 2. Finally, columns 5 and 6 show that the share of Kresy ancestors is also significantly related to the probability of finishing secondary and higher education.

While Panel A in Table 5 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 not confound our results.

4.3 Ancestry Survey Results – Ancestor Level

In the analysis above, we used our Ancestry Survey at the respondent level. We gain complementary insights by using the data at the ancestor level, where each ancestor a of each respondent i is

[31] The similarity in coefficients can be explained by the fact that the majority (52.4%) of respondents with any Kresy ancestor in the representative sample of our Ancestry Survey had *all* ancestors from Kresy. When using a dummy for any ancestor from Kresy in the specification of column 1 in Table 5 (i.e., replicating the Diagnoza specification), we obtain a coefficient of 0.913 (with standard error 0.106).

a separate observation. We estimate the following equation:

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

where Y_i is respondent i's education, as above, and $Kresy_{a_{(i)}}$ indicates whether ancestor a of respondent i came from Kresy. In addition to all standard controls for respondents' demographics ($\mathbf{X}_i$) and destination location fixed effects $\eta_{Locality_{(i)}}$, we control for ancestor characteristics $\mathbf{A}_{a_{(i)}}$: dummies for whether ancestor a is a parent, grandparent, or great-grandparent of respondent i.[32] $O_{a_{(i)}}$ denotes characteristics at the origin location of ancestor a, such as whether a came from a rural area. We also include dummies indicating whether the ancestor was an autochthon or came from abroad, which leaves origin from Central Poland 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 that education of the respondent does not vary across ancestors.[33]

Table 6 shows that our results are robust to the estimation at the ancestor level. Odd columns include county fixed effects; even columns include municipality fixed effects. We find positive and significant coefficients on the Kresy dummy throughout. As one should expect, the point estimates are somewhat smaller in the ancestor-level estimation than in the respondent-level estimation (see footnote 33 and Appendix Section IV for detail).

Border Sample based on Ancestry Survey

Our Ancestry Survey also allows us to perform a particularly restrictive border sample analysis, complementing the results in Section 4.1. The Ancestry Survey includes information on ancestors from *both* sides of the Kresy border (see Figure A.10 in the appendix). This enables us to compare people who live in the same town or village in WT today, but have ancestors from the different sides of the Kresy border.

Figure 5 illustrates the border effect for years of education. As in Section 4.1, we restrict the sample to people with ancestor origin within less than 150 kilometers of the Kresy border. However, in contrast to the Diagnoza analysis presented above, we use fixed effects for the current

[32] Among all ancestors (who were the youngest adults in 1939 in the respondents' family history), 23% were the parents; 55% were the grandparents; and 22% were the great-grandparents.

[33] It is important to note that, econometrically, respondent-level and ancestor-level regressions are not equivalent. In Appendix IV we present Monte Carlo simulations comparing the results of ancestor-level and respondent-level regressions. First, we show that the point estimate of the parameter of interest in the ancestor-level regression, γ, is smaller than the point estimate of the parameter of interest in respondent-level regressions, β from equation (1). The relationship between the two parameters depends on the correlation between indicator variables for Kresy origin of different ancestors of the same respondent. Second, we show that the level of significance in the respondent-level and the ancestor-level regressions is similar irrespective of the correlation among ancestor origins of the same respondent, as long as this correlation is positive (as is the case in our data). In other words, statistical inference in both types of regressions is the same.

municipality of respondents in addition to our standard controls. The figure confirms and extends our previous border results: When comparing people who live in the same municipality in the Western Territories today, those whose ancestors were expelled from just a few kilometers to the east of the Kresy border have significantly higher education than those whose ancestors lived a few kilometers to the west of the Kresy border. Table A.4 in the appendix provides the corresponding regression results, in which we use our three measures of education as dependent variables and control for a quadratic polynomial of latitude and longitude of ancestor's locations in addition to the current residence municipality fixed effects and the standard controls. We also show robustness to restricting the sample to 100 kilometers around the Kresy border. In all cases, the dummy for Kresy origin is positive and statistically significant.

A limitation of the border analysis in our Ancestry Survey is that migrants from the west of the Kresy border (i.e., from Central Poland) may have been selected. To address this concern, the two border analyses from Diagnoza and our Ancestry Survey serve as complements: They show that the descendants of forced migrants from the east of the Kresy border are more educated than *both* 'stayers' from the west of the border (Figure 4) and than the descendants of (voluntary) migrants who left the area to the west of the Kresy border (Figure 5). The magnitude of the Kresy coefficients is also very similar in both analyses (see Tables 2 and A.4). Thus, in combination, the two border samples suggest that selection of voluntary migrants is unlikely to confound our results. Nevertheless, we discuss the possibility of selected voluntary migrants at length in the next section.

5 Mechanisms

What explains the higher education among the descendants of forced migrants? In this section, we consider several potential mechanisms. We begin by showing evidence for the most likely mechanism – a shift in preferences towards investment in education as opposed to physical capital. We then continue by discussing three alternative mechanisms: Pre-existing differences, selection of migrants, and differential treatment at the destination (such as access to job opportunities) as well as various other possible channels such as congested labor markets and differential out-migration, fertility, or returns to schooling. None of these alternative mechanisms appear to play an important role. The table at the end of the appendix summarizes our findings for each (potential) mechanism, with references to tables and sections in the main text.

5.1 Preferences for Education vs. Ownership of Physical Assets

In Table 7, we examine differences in attitudes toward education and material possessions between descendants of forced migrants and other Poles. In the first two columns, we use a question from

Diagnoza about respondents' aspiration for the education of their children. The outcome variable is an indicator that takes on value one for respondents with the highest aspiration.[34] People with Kresy ancestors score 8.6 percentage points higher, relative to a mean of 61%. Remarkably, this result holds even after we control for the respondent's own education (column 2). In other words, among people with the same years of schooling (who also live in the same county), those with Kresy ancestors have significantly stronger preference for the education of their children. We elaborate on this finding by performing a Sobel-Goodman mediation test. This analysis (not reported in the table) examines the extent to which a mediating variable (respondents' own education) carries the influence of an explanatory variable (Kresy ancestors) to a dependent variable (aspiration for education of children). We find that only 13.7% of the effect of Kresy ancestry on "high aspiration" are mediated by own education. The remaining 86.3% constitute a "direct effect," i.e., independent of a respondent's own education. This finding suggests that our results are predominantly driven by a change in preferences towards education among people with Kresy ancestors, as opposed to a mechanism whereby "educated parents have educated children."

In columns 3-6 of Table 7 we examine answers to the question: "What is the main condition for success in life?" We construct dummies for two categories: "possession of material goods" and "freedom."[35] 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., house, apartment, vacation home, garden land plot, ebook reader, 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 of assets *not* owned for non-financial reasons (i.e., assets that the household could afford, but chooses not to purchase), divided by the number of all non-possessed assets.[36] Consistent with the results on stated preferences from columns 3 and 4, we find that Kresy migrants own fewer assets, even if they could afford them. In sum, the results in Table 7 lend support to the interpretation that

[34]The survey question was: 'What level of education would you like your children to attain?' The answer included five categories, and we create a dummy for the highest category. Results are robust to using the full categorical variable instead of the dummy for the highest score. Note that the sample is smaller because this question is not answered when children have already finished their education.

[35]For each category, the dummy takes on value one when the respondent answered: "definitely yes," "yes," or "rather yes." The dummy equals zero for the answers "neither yes nor no," "rather no," "no," and "definitely no."

[36]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.

forced migration shifted preferences towards investment in education, and away from material possessions. In what follows, we turn to alternative possible mechanisms.

5.2 Pre-Existing Differences?

Could our results be driven by differences of Poles from Kresy before WWII? Possible examples include pre-existing differences in education, in preferences for schooling, and in socio-economic or geographic characteristics. We show that these are unlikely to affect our findings.

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 have presented evidence that makes this unlikely. Figure 2 shows that in 1921, Roman Catholics (i.e., Poles) in Kresy had a literacy rate of 58.9%, as compared to 65.4% in Central Poland. Table A.5 in the appendix shows that this pattern also holds when we differentiate between rural and urban areas. 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 is also confirmed by our cohort analysis in Figure 3 and in the border sample in Figure 4 (left panel).

Did Poles in Kresy already have higher preferences for education before WWII?

Pre-existing differences in preferences for education are unlikely to drive our results. As discussed in Section 2.1, there were no differences in access to education in Kresy vs. CP before WWII (all belonged to Poland then), and there was also no discrimination of Poles in Kresy. Thus, if Poles from Kresy had pre-existing preferences for education, these should have materialized in higher literacy rates before WWII. If anything, the contrary is true, as we have shown above. In addition, the Kresy border was arbitrary (see Section 2.1 and Appendix V). It is thus unlikely that pre-existing cultural differences would jump at the Kresy border. Consequently, our border samples in Figures 4 and 5 help to address possible unobserved differences that may have affected the education decisions of migrants.[37]

Socio-Economic and Geographic Characteristics at Origin Locations

To what extent do characteristics of migrants' origin locations affect the relationship between education and Kresy origin? To analyze this, we use our Ancestry Survey and enrich specification (2)

[37] Another factor that potentially could confound our results is the difference in the way imperial powers treated Poles during the Partitions of Poland *before 1918* (see Section 2.1). To address this, we use the historical fact that *within* the Partitions, Poles were treated equally no matter if they lived in Kresy or Central Poland. Below in Section 5.4, we replicate our main result in a sub-sample of respondents whose ancestors came from the Russian Partition (see odd columns in Panel B of Table 9). The Russian Partition covered about three quarters of Kresy and one half of Central Poland, which makes this restriction of the sample possible.

by adding a host of socio-economic and geographic controls (all measured at origin locations), as well as their interaction with the Kresy origin dummy. Specifically, using the 1931 Polish Census, we interact Kresy origin with the share of Roman Catholics, the shares of native Polish, Ukrainian, and Russian speakers, the literacy rate, and the urbanization rate. We also use the share of literate Roman Catholics from the Polish Census of 1921. Going beyond the population characteristics, we 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 were key features of the economic environment. Tables A.6 and A.7 in the appendix show that neither the variables in levels nor their interaction terms with Kresy origin are statistically significant. In addition, the coefficients (all based on standardized variables) are typically an order of magnitude smaller than the coefficient on Kresy origin, while the latter maintains its magnitude and significance from our baseline ancestor regressions in Table 6 (columns 1 and 2). We interpret this as evidence that the effect of uprootedness is driven by forced migration itself, and not by specific circumstances at the place of origin.[38]

5.3 Selection of Migrants from Kresy?

Could selection of forced migrants from Kresy drive our results? We discuss the possibilities of selection at the origin and selection into destinations.

Were forced migrants from Kresy selected at the origin?

Selection *at the origin* is highly unlikely among Kresy migrants, as it contradicts 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). In particular, the pressure on Poles to leave was lower in rural areas in the Belorussian and Lithuanian parts of Kresy. In Ukraine, in contrast, rising animosity between Poles and Ukrainians at the end of WWII led to an (almost) complete exodus of Poles even from rural areas. In what follows, we explore this variation 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 origin, and the county of origin).

If selection of Poles from Kresy affects our results, the coefficient on Kresy origin should vary depending on how much scope for selection a given ancestor's region of origin offered. Table 8 tests whether this is the case using our main outcome variable – years of education. We cre-

[38] Among the interaction results, the following is worth highlighting: Columns 2-4 of Table A.6 show that our main result is not affected by the share of Poles, Ukrainians, or Russians at the ancestors' origin locations. Moreover, the interaction between Kresy and each of these shares is small, negative, and insignificant. This suggests that a possible pre-existing animosity between Poles and other ethnicities does not affect our results.

ate different subsamples depending on ancestors' locations of origin. Regressions are run at the ancestor level as outlined by equation (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 urban and rural origin locations, respectively. The point estimates are slightly higher for the urban origin sample than for the rural origin sample. In other words, our results are stronger for locations from which the expulsion of Poles was nearly universal. One potential concern is that the estimate in the urban origin sample (column 2) could be inflated if more educated urban migrants from Kresy (see Table A.5) were displaced to rural areas in WT. If these (former) city dwellers passed on their taste for education, we would compare their well-educated descendants to the less educated rural population in WT. We address this possibility in column 4, restricting the sample to those cases in which *both* ancestors and descendants are from urban areas. The effect of Kresy is almost unchanged. In column 5, we also report the results for the subsample of rural origins and destinations. The coefficient on Kresy is smaller (as one would expect, given the lower average education in rural areas), but it remains significant at the 10% level.

In columns 6-8 in Table 8 we repeat the above analysis, now restricting the sample to ancestors from the Ukrainian part of Kresy, where expulsions were universal. The coefficient in column 6 (for both urban and rural origin 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 7 and 8 show a similar pattern as columns 2 and 3: Coefficients are highly significant for both rural and urban ancestors; and they are somewhat larger in the urban origin subsample. In sum, it is unlikely that our findings are driven by selection of Kresy migrants at the origin.

Selection of forced Kresy migrants into destinations?

Even if selection from origin locations in Kresy 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 the Western Territories, about one quarter moved to Central Poland. For example, if the most capable Kresy migrants moved to the Western Territories, our results within WT would be biased. In addressing this concern, we begin by noting that the results from Table 2 (columns 5 and 6) show that the coefficients on Kresy ancestry are, if anything, larger in CP than in WT. Next, we present an additional check: We restrict the Diagnoza sample to respondents with Kresy origin. Within this subsample, we can compare the level of education of those who live in CP (overall 1,314 respondents) with those who live in WT (2,008 respondents). Table A.8 in the appendix shows that respondents with Kresy origin are somewhat *less* educated in the Western

Territories than in Central Poland.[39] Overall, these results suggest that selection of Kresy migrants into different areas of Poland is not driving our results.

5.4 Selection of Voluntary Migrants?

In our results for Poland overall (i.e., using the Diagnoza Survey), selection of the *control group* (i.e., voluntary migrants) is not an issue – the control group comprises 'all other Poles.' However, our Ancestry Survey was conducted only in the 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 biased upward if the control group of less educated individuals was more likely to migrate from CP to WT after WWII. In what follows, we perform several analyses that show that this is unlikely to confound our findings.

Regional selection of voluntary migrants from Central Poland to the Western Territories?

We first examine the possibility of regional selection – migrants from Central Poland coming from areas with historically lower education. 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 counties of their origin. This allows us to compare the historical literacy rates in the counties of origin of migrants from Kresy and from Central Poland. Table 9 presents the results of regressions at the ancestor level, with secondary education as the contemporaneous measure for education in odd columns (see footnote 3), and with historical literacy in even columns. Panel A uses literacy of Roman Catholics from the 1921 Polish Census that covered all of the Second Polish Republic; Panel B uses literacy of Poles in the Polish language from the 1897 Russian Empire Census, covering the Russian partition of Poland, which after 1918 became a part of the SPR.[40] Column 1 replicates our main results in the subsamples for which the historical literacy data at the ancestors' origins are available: We find that in both samples, respondents with Kresy ancestors have significantly higher secondary education than respondents with ancestors from Central Poland who live in the same county today. Column 2 uses historical literacy rates as the dependent variable. The coefficient on the Kresy dummy in this regression shows the average difference in historical literacy rates between counties in Kresy and in Central Poland from which respondents' ancestors originated. Because we use respondent county fixed effects, we compare historical lit-

[39] The reason for this difference is probably migration of highly skilled individuals with Kresy background to urban centers such as Warsaw and Krakow in Central Poland. Indeed, people with Kresy origin show a particularly high education advantage in these areas (see Table A.8).

[40] The number of observations in Panel B 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. Note also that neither of these historical censuses cover the Western Territories (which belonged to Germany).

eracy rates at the origin of ancestors whose descendants today live in the same counties in WT. According to the results in column 2, Kresy ancestors came on average from locations with a 3 percentage point *lower* literacy rate. Columns 3-6 show that a similar pattern of 'reversal of education' holds when we restrict the sample to ancestors from rural origin locations or to those from urban origins.

Individual selection of voluntary migrants from Central Poland to the Western Territories?

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 Western Territories, whereas educated Poles *from the same origin counties* stayed in Central Poland. In what follows, we show that this type of selection is unlikely to drive our results.

Negative selection of Central Polish migrants into WT would imply that the control group in our Ancestry Survey has too low education, biasing the coefficient on Kresy origin upward. To examine directly whether there was negative individual selection, we would need historical individual-level data on the education of voluntary migrants and stayers in Central Poland. These are not available. However, we can check whether the (potential) selection concern matters *for our results*: If one were worried about negative selection of migrants from CP, 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: We show that respondents in WT with ancestors from Central Poland (i.e., voluntary migrants) are actually slightly *more* educated than a reasonable comparison group – today's respondents in those counties in CP where the voluntary migrants' ancestors originated from.

To implement this check, we focus on respondents whose ancestors moved from CP to WT. From our Ancestry Survey, we know their county of origin in Central Poland. We also know the education level today in these origin counties from respondents in the Diagnoza Survey.[41] Using the combined information, we construct the following variable for each respondent i:

$$\Delta Edu(i) = Edu^{WT}(i) - E\left[Edu^{CP}_{county}\left(a_{(i)}\right)\right] \tag{3}$$

where $Edu^{WT}(i)$ is today's education of respondent i living in WT, whose ancestors came from CP. The term $Edu^{CP}_{county}\left(a_{(i)}\right)$ denotes the average education today in the CP county of origin of ancestor a of respondent i. $E\left[\,\cdot\,\right]$ is the average education across origin counties of all ancestors

[41] We only use Diagnoza respondents in Central Poland without any ancestors from Kresy. Similarly, we restrict the subsample from our Ancestry Survey to those respondents who have only ancestors from Central Poland.

of respondent i. Since we only look at descendants of migrants from CP, all these counties are in Central Poland.

Table 10 presents the results for the null hypothesis that $\Delta Edu(i) = 0$ for secondary education and for higher education.[42] Columns 1 and 2 show positive differences, i.e., that descendants of CP migrants who now live in WT have on average slightly *higher* education than their 'cousins' in their ancestors' origin counties in CP. This result could be driven by migration from rural areas in CP to cities in WT: Since education is higher in urban areas, destinations would tend to show higher education than origin locations.[43] To account for this possibility, we restrict the sample to individuals for whom *both* origin and destination locations were urban (columns 3 and 4) or rural (columns 5 and 6). In all cases, the differences are small and statistically insignificant.[44] This suggests that the positive differences shown in columns 1 and 2 are in part driven by rural-to-urban migration.[45] Another possible explanation for the positive $\Delta Edu(i)$ in columns 1 and 2 is that CP migrants from rural areas who came to WT cities may have been positively selected. Ultimately, we cannot differentiate between selection among historical migrants and other potential mechanisms that may drive the observed (small) educational gap.[46] Nevertheless, the results from Table 10 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 on average, if anything, somewhat *better* educated than their closest comparison groups. Thus, our Ancestry Survey results tend to *underestimate* the effect for Kresy origin in the Western Territories.

[42]The definition of years of education is different across the two surveys. In Diagnoza, this variable is the self-reported number of years spent in educational institutions. In contrast, in our Ancestry Survey years of education are imputed using four educational categories. While years of education are comparable for different observations within each survey, they are not directly comparable between the two data sources. As $\Delta Edu(i)$ entails the comparison of values across the two surveys, we do not use years of education in this analysis.

[43]Note that this concern is specific to the analysis in Table 10, which compares individuals *across* locations and therefore does not use location fixed effects. In contrast, all our main results hold with municipality fixed effects, which absorb (among many others) average differences across urban vs. rural areas.

[44]Note that, in contrast, our main results hold in the urban-to-urban and rural-to-rural subsamples (see columns 4 and 5 in Table 8).

[45]In fact, if we restrict the sample to respondents in urban areas of WT with ancestors from rural CP areas, we – unsurprisingly – obtain significantly positive differences.

[46]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). 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.

5.5 Differential Treatment at Destinations and Other Potential Channels

In what follows, we examine whether our findings may be affected by differential access to labor markets (congestion) or by other potentially different characteristics of Kresy migrants after their displacement.

Congestion

The previous literature (as discussed in the introduction) showed that migrants who lack access to local land resources (which are held by entrenched locals) often opt for education in order to get access to non-agricultural jobs. This is unlikely to affect our results for several reasons. First, the Western Territories were largely empty after WWII, and the idea of the resettlement was to populate this 'empty space.' Second, as we described in Section 2.2, migrants from Kresy and CP arrived to WT at the same time (see Figure A.4). 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.

While destination fixed effects in our previous regressions capture any direct effect of congestion on education, it is still possible that congestion affected Kresy migrants differentially. We test for this channel by using interactions between Kresy ancestry and the population of autochthons in the respondent's county of residence. Autochthons were a minority in WT, but their share varied across localities. We use two alternative measures of autochthons' presence in counties in WT: the share of Polish native speakers in 1900 (from the German 1900 Census) and the share of autochthons in 1950 from the Polish Census.[47]

Columns 1 and 2 in Table 11 report the results, adding an interaction term between the Kresy origin of respondents and the county-level autochthon share to specification (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. We use the Diagnoza Survey focusing on respondents in the Western Territories, with years of education as dependent variable.[48] The coefficient on the share of autochthons is positive and significant, both when measured by the share of Polish speakers in 1900 (column 1) and when using the 1950 Polish Census (column 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. Im-

[47]The share of autochthons in 1950 in the median county was 6.5%, and the mean, 15%. For Polish speakers in 1900, the median was 1.4% (mean 13%). Figure A.11 in the appendix shows that the two measures are highly correlated. Also, there is ample variation, with some counties having more than 90% autochthons, while others had close to zero.

[48]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.

portantly, we 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 0.1 extra years of schooling among people with Kresy ancestors – relative to a direct Kresy coefficient of almost 0.8. These results suggest that differential congestion effects for Kresy migrants are unlikely to drive our findings.[49]

Returns to schooling

Could our results be driven by differential returns to schooling for Kresy migrants? We shed light on this question in columns 3 and 4 in Table 11. We use log household income as dependent variable and are interested in the interaction term between Kresy origin 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 the Western Territories.[50] This suggests that differential returns to schooling do not affect our results.

Out-migration

Columns 5 and 6 in Table 11 examine whether differential migration from Poland to other countries (after Poland's EU accession in 2004) 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?" We find no relationship between Kresy ancestry and the intent to emigrate (column 5). In addition, the interaction term between education and Kresy origin is also small and insignificant (column 6). If the respondents who intend to emigrate have similar characteristics as those who had left already, these results make it unlikely that education and Kresy origin drove emigration in a fashion that would confound our results. As we do not observe directly the people who emigrated, we provide indirect evidence in support of this underlying assumption. The Polish Census in 2011 included a question: "How many members of your household have emigrated?" The response to this question is publicly available at the regional level. In Figure A.12 in the appendix we show that there is a

[49] Another type of congestion effect could theoretically have affected Kresy migrants if they had systematically arrived in the Western Territories after migrants from Central Poland. However, this is not consistent with the historical evidence. As shown in Figure A.4 in the appendix, migrants from CP and from Kresy arrived in parallel throughout the years after WWII.

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

strong positive relationship between the actual out-migration and the intent to emigrate reported in Diagnoza. This validates our use of the latter as a proxy for emigration from Poland.

Differential fertility

Columns 7 and 8 in Table 11 study the possibility that differential fertility may confound our results. For example, Kresy migrants may have chosen lower fertility to remain more flexible in an environment that they perceived as highly volatile (see Section 2.3). Fewer offspring could then have enabled higher investment in each child's human capital. Over time, this may have translated into stronger preferences for education. We find that Kresy origin is uncorrelated with the number of children per household member, which is the closest proxy for measuring fertility in our data. While this does not exclude the possibility that differential fertility played a role initially, it makes it unlikely that this channel is still at play for the younger generations in our data. In addition, note that in the differential-fertility interpretation, preferences for education would develop later, with lower fertility being the initial driver. In contrast, the historical evidence discussed in Section 2.3 suggests that preferences shifted immediately, as a direct result of uprootedness.

Communities

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 descendants' education, this may have reinforced 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 for migrants moving as whole communities (which we describe in Appendix Section VI). Using this measure, Table A.9 in the appendix shows that controlling for whether ancestors moved as a community does not affect our main results.

Recall bias

A potential worry in using survey data about ancestral origin is 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. If education leads to a higher probability of remembering ancestors (and thus, ancestors from Kresy), then our results would be biased. In the Diagnoza survey, we cannot control for this potential bias. In contrast, in our Ancestor Survey, recall bias is less of a concern, because it should affect both our 'treatment group' of Kresy ancestors as well as the 'control group' of ancestors from other areas. We can use our Ancestry Survey to check whether there is differential recall bias for people with ancestors from Kresy, i.e., whether remembering (any) ancestor locations is correlated with Kresy origin. We construct, for each respondent, the share of ancestors with missing information

on their location of origin (which is low – only 12% on average). We then show that i) the share of ancestors with missing information is uncorrelated with Kresy origin, and ii) controlling for this share does not affect our results. We describe how we built this variable in Appendix Section VI and present the results in Table A.10.

5.6 Summary of Mechanisms

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 on education and 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. Our results suggest that, indeed, forced migration caused an increase in educational investment among the affected Poles and their descendants, relative to all other Poles. Further, our findings in Table 7 suggest that this education premium is driven by a shift in preferences away from material possessions and towards education.

6 Conclusion

Forced migration is an important 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 natural disasters. While the immediate experience of expulsion is dramatic, the long-run effects on the displaced and their descendants are less clear. Such long-term effects of forced migration are difficult to distinguish from confounding factors. We collected novel individual-level data to 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. We find that the children, grandchildren, and great-grandchildren of forced Kresy migrants have significantly higher average education levels than all other Poles. This result holds in border samples around the Kresy border and is robust to a host of controls. We also show that descendants of forced migrants value the education of their children more, and assign a lower importance to material possessions. We examine several possible interpretations of these results and conclude that the most likely is that uprootedness shifted forced migrants' preferences away from investment in physical assets and toward investment in portable human capital.

The observed emphasis on 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 forced migrants and their children. While the international aid community does consider education as an important factor contributing to the reduction of economic and social marginalization of refugees (G20, 2017; UNICEF, 2017), our results show that the benefits of providing schooling for forced migrants may be even higher – and more persistent – than previously thought.

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, M., A. Cassar, J. Chytilová, and J. Henrich (2014). War's Enduring Effects on the Development of Egalitarian Motivations and In-Group Biases. *Psychological Science 25*(1), 47–57.

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. Khwaja, and A. Mian (2015). Population Exchange and its Impact on Literacy, Occupation and Gender – Evidence from the Partition of India. *International Migration 53*(4), 90–106.

Bieniasz, Z. (1987). Droga do mądrości. In Z. Linkowski (Ed.), *Wiosna na Rumowisku i Inne Wspomnienia Pionierów*. Gorzów Wielkopolski: Wydawnictwo regionalne.

Blattman, C. and E. Miguel (2010). Civil War. *Journal of Economic Literature 48*(1), 3–57.

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. Omar 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.

Cameron, L. and M. Shah (2015). Risk-Taking Behavior in the Wake of Natural Disasters. *Journal of Human Resources 50*(2), 484–515.

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.

Cassar, A., P. Grosjean, and S. Whitt (2013). Legacies of Violence: Trust and Market Development. *Journal of Economic Growth 18*(3), 285–318.

Cassar, A., A. Healy, and C. von Kessler (2017). Trust, Risk, and Time Preferences After a Natural Disaster: Experimental Evidence from Thailand. *World Development 94*, 90–105.

Chyn, E. (2018). Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Labor

Market Outcomes of Children. *American Economic Review forthcoming*.

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

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

Eberhardt, P. (2003). *Ethnic Groups and Population Changes in Twentieth Century Eastern Europe: History, Data and Analysis*. London and New York: Routledge.

Fisman, R., P. Jakiela, and S. Kariv (2015). How did Distributional Preferences Change during the Great Recession? *Journal of Public Economics 128*(C), 84–95.

G20 (2017). Education and Skills Development in the Context of Forced Migration. Technical report, G20 Insights.

Galor, O. and O. Özak (2016, October). The Agricultural Origins of Time Preference. *American Economic Review 106*(10), 3064–3103.

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

Giuliano, P. and A. Spilimbergo (2014). Growing up in a Recession. *Review of Economic Studies 81*(2), 787–817.

Grosfeld, I. and E. Zhuravskaya (2015). Cultural vs. Economic Legacies of Empires: Evidence from the Partition of Poland. *Journal of Comparative Economics 43*(1), 55–75.

Grzebałkowska, M. (2015). *1945. Wojna i pokój*. Warszawa: Agora.

Halicka, B. (2015). *Polski Dziki Zachód. Przymusowe Migracje i Kulturowe Oswajanie Nadodrza 1945-1948*. Krakow: TAiWPN Universitas.

Hanaoka, C., H. Shigeoka, and Y. Watanabe (2018). Do Risk Preferences Change? Evidence from the Great East Japan Earthquake. *American Economic Journal: Applied Economics 10*(2), 298–330.

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.

Jakiela, P. and O. Ozier (2018). The Impact of Violence on Individual Risk Preferences: Evidence from a Natural Experiment. *Review of Economics and Statistics Forthcoming*.

Katz, E. and H. Rapoport (2005). On human capital formation with exit options. *Journal of Population Economics 18*(2), 267–274.

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

Malmendier, U. and S. Nagel (2016). Learning from Inflation Experiences. *Quarterly Journal of Economics 131*(1), 53–87.

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.

Oz, A. (2005). *A Tale of Love and Darkness*. Vintage, NY City.

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.

Schechtman, J. B. (1962). *Postwar Population Transfers in Europe 1945-1955*. Philadelphia: University of Pennsylvania Press.

Semrad, A. (2015). Immigration and Educational Spillovers: Evidence from Sudeten German Expellees in Post-War Bavaria. Munich Discussion Paper No. 2015-7.

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. (1999). To Resolve the Ukrainian Question once and for all: The Ethnic Cleansing of Ukrainians in Poland, 1943-1947. *Journal of Cold War Studies 1*(2), 86–120.

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

Stigler, G. J. and G. S. Becker (1977). De Gustibus Non Est Disputandum. *American Economic Review 67*(2), 76–90.

Ther, P. (1996). The Integration of Expellees in Germany and Poland after World War II: A Historical Reassessment. *Slavic Review 55*(4), 779–805.

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.

Turnau, I. (1960). *Studia nad strukturą ludnościową polskiego Wrocławia*. Poznań: Instytut Zachodni.

UNICEF (2017). Education Uprooted: For Every Migrant, Refugee and Displaced Child, Education. Technical report, UNICEF.

Voors, M., E. Nillesen, P. Verwimp, E. Bulte, R. Lensink, and D. van Soest (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review 102*(2), 941–964.

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

Zhang, J. (2018). Culture, Institutions, and the Gender Gap in Competitive Inclination: Evidence from The Communist Experiment in China. *Economic Journal Forthcoming*.

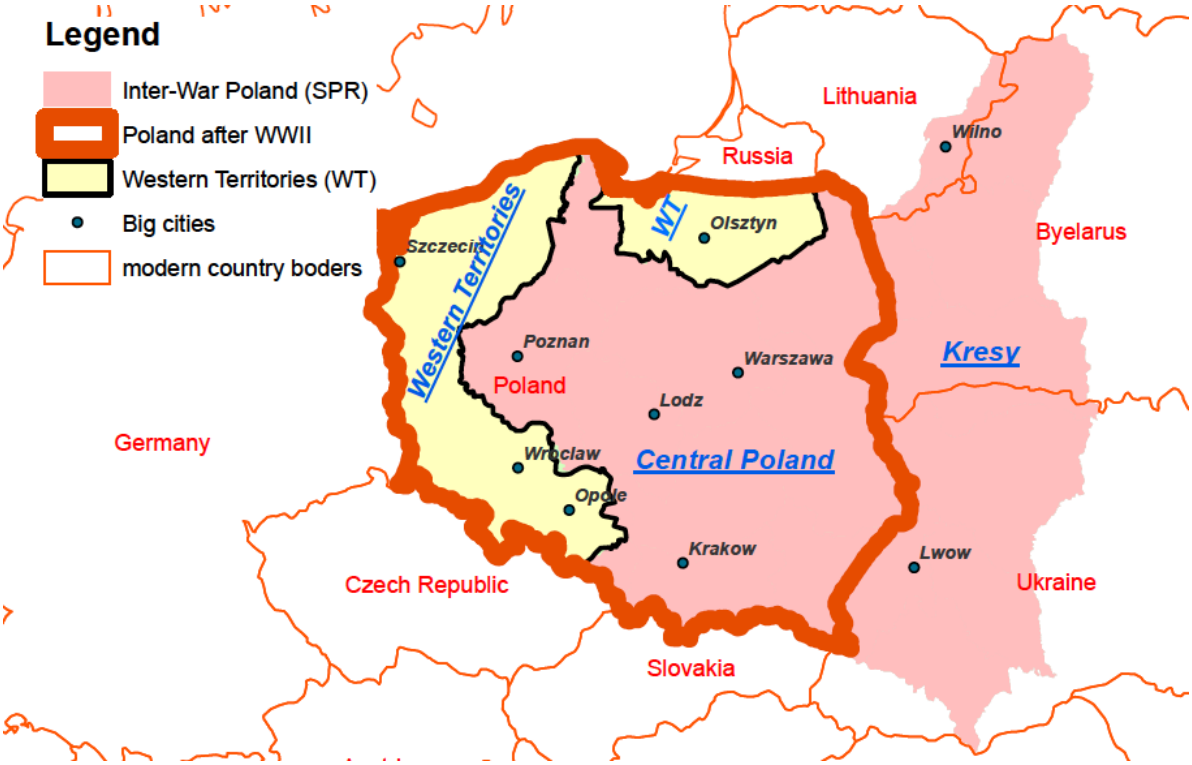


Figure 1: Map of Poland's Territorial Change after WWII

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 and north (Western Territories) became Polish. Poles from Kresy were forced to leave – the vast majority was resettled to the emptied Western Territories.

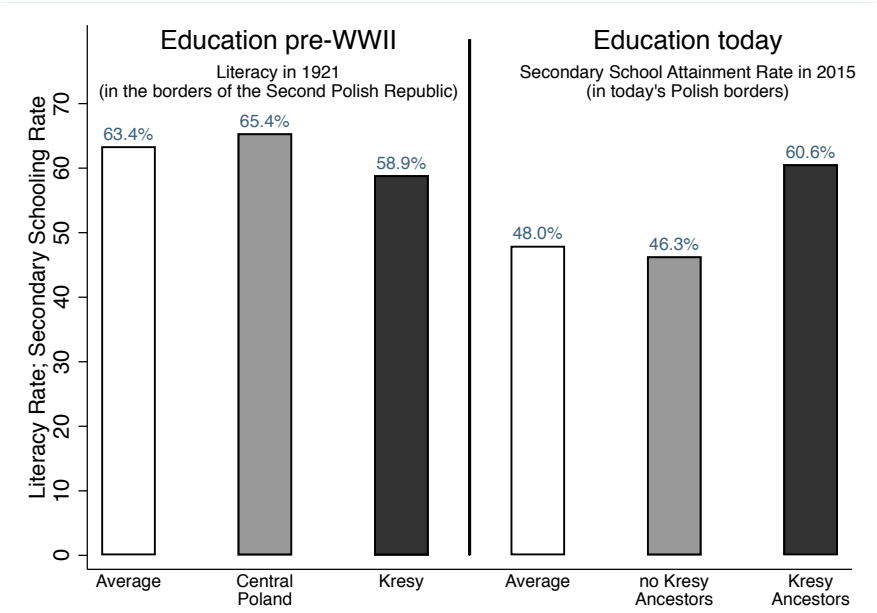


Figure 2: Overview of Historical and Contemporaneous Patterns in Education

Note: The figure shows the reversal in education for forced migrants and their descendants: People at the origin location of forced migrants (Kresy) had lower education before WWII, while descendants of forced Kresy migrants today have higher educational attainment. The data are from the 1921 Polish Census and the 2015 Diagnoza Survey. For 1921, the figure displays literacy rates of Roman Catholics (i.e., ethnic Poles) in the whole of the Second Polish Republic, 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 on average, for people without Kresy ancestors (25,972 respondents), and for people with Kresy ancestors (3,318 respondents).

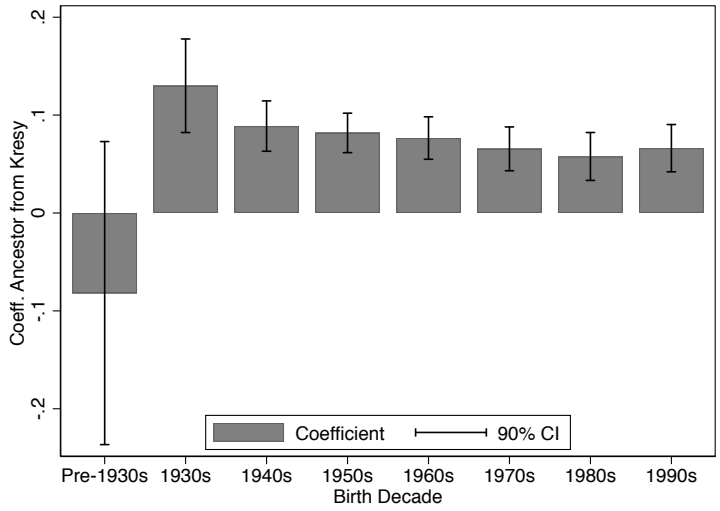


Figure 3: Ancestors from Kresy and Eduction, by Birth Cohort

Note: The figure visualizes the results of regressing ln(years of education) on Kresy ancestry for different birth cohorts. The underlying regressions include our standard controls (see note to Table 2) and respondent county fixed effects. Each bar corresponds to the coefficient on 'Ancestor from Kresy.' The pre-1930 birth cohort was at least 16 years old at the end of WWII and was above schooling age at the time of forced migration. The regressions are run using the Diagnoza sample for 2015 (see Table 3 for similar regressions). Respondents who were still students by the time of the survey in 2015 are excluded.

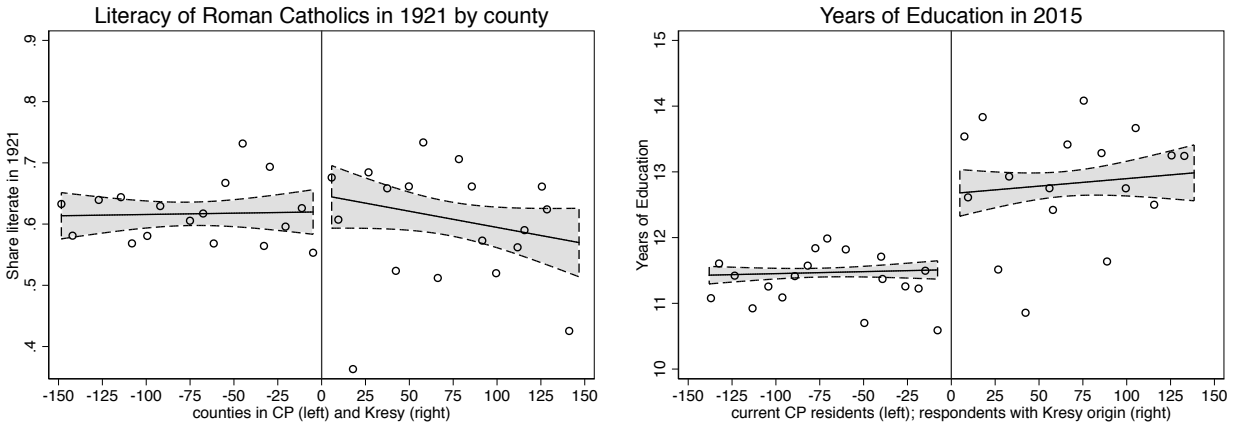


Figure 4: Kresy Border Sample: 1921 Census and Diagnoza Survey

Note: The figure uses only respondents (from the 2015 Diagnoza Survey) with roots in the area of less than 150 km around the border between Kresy and Central Poland. 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 8 km (5 miles) bins for visualization, while the lines are based on all underlying observations, with the shaded area representing 90% confidence intervals.

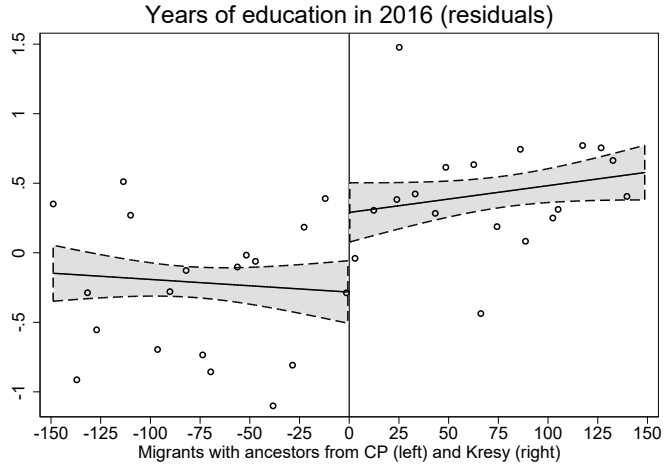


Figure 5: Kresy Border Sample: Ancestry Survey

Note: The figure uses respondents from our Ancestry Survey, i.e., individuals who live in the Western Territories today. Among these, we restrict the sample to people with ancestor roots in the area of less than 150 km around the border between Kresy and Central Poland. Underlying the figure is an ancestor-level regression, as in specification (2), of years of education on our standard controls (see note to Table 5) and on respondents' municipality fixed effects. Dots correspond to residuals from this regression (aggregated into 8 km (5 miles) 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 in the appendix.

TABLES

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

	Western Territories (WT)	Central Poland (CP)	Share of 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. Data are from the 1950 Polish census. 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: Forced Migration from Kresy and Education

Dependent variable: Individual-Level education, as indicated in each panel

Sample:	(1) All (no FE)	(2) All	(3) Rural	(4) Urban	(5) Central Poland	(6) Western Territories	(7) Kresy Border Sample†
Panel A. Dep. Var.: Years of education							
Ancestor from Kresy	0.973***	0.930***	0.805***	1.007***	1.022***	0.830***	1.386***
	(0.077)	(0.078)	(0.127)	(0.099)	(0.118)	(0.104)	(0.305)
Mean Dep. Var.	11.95	11.95	11.12	12.78	11.98	11.87	11.72
Observations	25,719	25,719	12,816	12,903	19,255	6,464	5,446
Panel B. Dep. Var.: Secondary education dummy							
Ancestor from Kresy	0.139***	0.127***	0.121***	0.133***	0.128***	0.122***	0.109**
	(0.011)	(0.012)	(0.021)	(0.014)	(0.017)	(0.016)	(0.047)
Mean Dep. Var.	0.51	0.51	0.38	0.64	0.51	0.51	0.49
Observations	25,720	25,720	12,819	12,901	19,249	6,471	5,440
Panel C. Dep. Var.: Higher education dummy							
Ancestor from Kresy	0.102***	0.099***	0.071***	0.114***	0.130***	0.070***	0.160***
	(0.011)	(0.011)	(0.016)	(0.014)	(0.017)	(0.014)	(0.039)
Mean Dep. Var.	0.21	0.21	0.13	0.30	0.21	0.21	0.20
Observations	25,720	25,720	12,819	12,901	19,249	6,471	5,440
Respondent county FE		✓	✓	✓	✓	✓	
Controls‡	✓	✓	✓	✓	✓	✓	✓

Notes: The table shows that individuals whose ancestors were expelled from the Kresy territories have significantly higher levels of education today. Regressions are run at the respondent level using data from the 2015 Diagnoza Survey; standard errors are clustered at the household level. * p<0.1, ** p<0.05, *** p<0.01.

† Column 7 uses only respondents with roots in the area of less than 150 km around the Kresy 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 7 also includes a quadratic polynomial in latitude and longitude.

Table 3: Kresy Ancestors and Education – Across Cohorts

Dependent variable: Individual-level education, as indicated in each panel

	(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	-0.669	1.311***	0.950***	0.959***	1.025***	0.891***	0.855***	0.788***
	(0.840)	(0.277)	(0.179)	(0.154)	(0.172)	(0.185)	(0.197)	(0.179)
Mean Dep. Var.	7.60	9.44	10.49	11.58	12.27	13.06	13.91	12.61
R-squared	0.66	0.44	0.31	0.24	0.23	0.30	0.27	0.45
Observations	516	2,075	3,347	5,391	4,418	4,129	3,771	2,016
Panel B: Dep. Var.: Secondary education dummy								
Ancestor from Kresy	0.025	0.165***	0.144***	0.135***	0.146***	0.094***	0.078***	0.132***
	(0.095)	(0.035)	(0.028)	(0.024)	(0.027)	(0.026)	(0.024)	(0.041)
Mean Dep. Var.	0.20	0.35	0.40	0.43	0.47	0.58	0.75	0.62
R-squared	0.58	0.42	0.27	0.23	0.23	0.26	0.23	0.38
Observations	520	2,078	3,348	5,388	4,419	4,127	3,774	2,018
Panel C: Dep. Var.: Higher education dummy								
Ancestor from Kresy	-0.059	0.108***	0.091***	0.105***	0.140***	0.124***	0.066**	0.073**
	(0.076)	(0.032)	(0.026)	(0.022)	(0.025)	(0.027)	(0.031)	(0.035)
Mean Dep. Var.	0.06	0.13	0.15	0.15	0.18	0.29	0.42	0.15
R-squared	0.52	0.27	0.18	0.15	0.21	0.27	0.27	0.35
Observations	520	2,078	3,348	5,388	4,419	4,127	3,774	2,018
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 using data from the 2015 Diagnoza Survey; 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

Dep. var.: Individual labor market outcomes, as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	ln(HH income)		White Collar Job		Unemployed	
Ancestor from Kresy	0.098**	0.051	0.103***	0.035***	-0.021***	-0.014*
	(0.039)	(0.039)	(0.014)	(0.012)	(0.007)	(0.008)
ln(years of education)		0.790***		0.947***		-0.117***
		(0.047)		(0.028)		(0.011)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	8.47	8.46	0.46	0.46	0.08	0.09
R-squared	0.20	0.22	0.26	0.40	0.06	0.07
Observations	17,763	15,922	13,474	13,462	18,347	16,453

Notes: The table shows that descendants of Kresy migrants have more favorable labor market outcomes. Regressions are run at the respondent level using data from the 2015 Diagnoza Survey; 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: Forced Migration from Kresy and Education in Western Territories: Ancestry Survey

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	1.013***	0.810***	0.678***	0.787***	0.110***	0.070***
	(0.134)	(0.139)	(0.245)	(0.170)	(0.021)	(0.018)
Share of Ancestors, WT		-1.027***	-0.542	-1.310***	-0.159***	-0.131***
		(0.196)	(0.332)	(0.253)	(0.032)	(0.024)
Share of Ancestors, abroad		-1.161	-3.615**	-0.224	-0.017	0.000
		(0.855)	(1.500)	(0.929)	(0.112)	(0.099)
Share of Ancestors, rural		-0.469***	-0.462	-0.532***	-0.064***	-0.035*
		(0.164)	(0.368)	(0.183)	(0.024)	(0.020)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	12.42	12.44	11.38	12.96	0.47	0.22
R^2	0.28	0.29	0.32	0.28	0.20	0.21
Observations	3,636	3,588	1,084	2,504	3,588	3,588
Panel B: Respondent Municipality Fixed Effects						
Share of Ancestors, Kresy	0.938***	0.741***	0.519**	0.834***	0.091***	0.068***
	(0.141)	(0.147)	(0.257)	(0.180)	(0.023)	(0.019)
Share of Ancestors, WT		-0.989***	-0.907***	-1.107***	-0.158***	-0.124***
		(0.200)	(0.349)	(0.261)	(0.034)	(0.026)
Share of Ancestors, abroad		-0.623	-2.856**	0.294	0.056	0.031
		(0.659)	(1.348)	(0.738)	(0.095)	(0.100)
Share of Ancestors, rural		-0.553***	-0.342	-0.581***	-0.082***	-0.044**
		(0.161)	(0.337)	(0.188)	(0.025)	(0.021)
Controls‡	✓	✓	✓	✓	✓	✓
Municipality FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	12.42	12.44	11.38	12.96	0.47	0.22
R^2	0.38	0.39	0.47	0.34	0.30	0.29
Observations	3,636	3,588	1,084	2,504	3,588	3,588

Notes: The table uses data from our 2016 Ancestry Survey in the 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 the 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 6: Education Outcomes from Ancestry Survey: Ancestor-level

Dependent variable: as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Years of Education		Secondary education		Higher education	
Ancestor from Kresy	0.503***	0.464***	0.072***	0.062***	0.046***	0.043***
	(0.092)	(0.091)	(0.015)	(0.014)	(0.014)	(0.014)
Ancestor from WT	-0.912***	-0.867***	-0.156***	-0.148***	-0.128***	-0.121***
	(0.137)	(0.134)	(0.024)	(0.025)	(0.020)	(0.020)
Ancestor from abroad	1.033	1.284	0.150	0.077	0.099	0.147
	(0.948)	(0.862)	(0.134)	(0.128)	(0.177)	(0.157)
Grandparent	0.346**	0.289*	0.025	0.022	0.033	0.017
	(0.159)	(0.160)	(0.025)	(0.026)	(0.021)	(0.022)
Great-grandparent	1.040***	0.872***	0.174***	0.146***	0.109***	0.085**
	(0.228)	(0.231)	(0.038)	(0.038)	(0.035)	(0.037)
Ancestor from rural area	-0.508***	-0.524***	-0.072***	-0.074***	-0.045***	-0.044***
	(0.099)	(0.094)	(0.016)	(0.015)	(0.015)	(0.015)
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓		✓		✓	
Municipality FE		✓		✓		✓
Mean Dep. Var.	13.04	13.04	0.55	0.55	0.26	0.26
R^2	0.29	0.37	0.21	0.30	0.22	0.29
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, using the data from our 2016 Ancestry Survey at the ancestor level, as given by specification (2). Standard errors are 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 7: Attitudes towards Education and Material Possessions

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	High aspiration for education of own children[#]		Main condition for success in life? Material goods		Main condition for success in life? Freedom		% Assets not owned for non-financial reasons[†]	
Ancestor from Kresy	0.086***	0.074***	-0.073***	-0.059***	0.020***	0.020***	0.041***	0.032***
	(0.029)	(0.028)	(0.013)	(0.013)	(0.006)	(0.006)	(0.009)	(0.009)
ln(years of education)		0.495***		-0.191***		-0.008		0.124***
		(0.052)		(0.015)		(0.006)		(0.008)
Controls[‡]	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.61	0.61	0.56	0.56	0.05	0.05	0.69	0.69
R-squared	0.24	0.27	0.12	0.13	0.08	0.08	0.19	0.20
Observations	4,747	4,747	22,049	22,049	21,586	21,586	28,973	28,973

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 respondent level using data from the 2015 Diagnoza Survey; 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 respondents to rank their aspiration for education of their children on a scale from 1 to 5. The dependent variable is an indicator for the highest category. Note that the sample is smaller because this question is not answered when children have already finished their education.

† Diagnoza asks about the possession of 20 different assets (e.g., apartment, vacation house, garden land plot, ebook reader, 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 8: Main Results for Kresy Migrants from Rural vs. Urban Areas, and from Ukraine Only

Dependent variable: Years of education in 2016, at the respondent level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
"Ancestors from Kresy" includes:	All Kresy Ancestors					Only Kresy Ancestors from Ukraine		
Notes on sample:		Ancestor location:		Ancestor & descendant:			Ancestor location:	
	all	urban	rural	urban	rural	all	urban	rural
Ancestor from Kresy	0.503***	0.668***	0.430***	0.584***	0.312*	0.444***	0.614***	0.347***
	(0.092)	(0.160)	(0.110)	(0.173)	(0.184)	(0.111)	(0.182)	(0.132)
Ancestor from rural area	-0.508***					-0.487***		
	(0.099)					(0.106)		
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	13.04	13.64	12.84	13.87	11.83	12.98	13.52	12.80
R^2	0.29	0.30	0.32	0.31	0.35	0.28	0.31	0.32
Observations	11,548	2,950	8,598	2,417	3,084	10,237	2,568	7,669

Notes: The table uses data from our 2016 Ancestry Survey in the 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 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, as well as indicators for the ancestor generation (grandparents and great-grandparents, with parents being the excluded category), and for ancestors from rural areas, Western Territories, and abroad. 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 or Urban origin		Rural origin		Urban origin	
Panel A: Literacy from the 1921 Polish Census						
Ancestor from Kresy	0.073***	-0.031*	0.060***	-0.041**	0.109***	-0.002
	(0.015)	(0.017)	(0.018)	(0.020)	(0.025)	(0.023)
Ancestor from rural area	-0.069***	-0.170***				
	(0.017)	(0.012)				
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.57	0.62	0.54	0.58	0.66	0.75
R^2	0.21	0.37	0.24	0.11	0.26	0.21
Observations	9,645	9,645	7,161	7,161	2,484	2,484
Panel B: Literacy from the 1897 Russian Census						
Ancestor from Kresy	0.139***	-0.031**	0.135***	-0.031**	0.137**	-0.030**
	(0.030)	(0.014)	(0.034)	(0.014)	(0.069)	(0.014)
Ancestor from rural area	-0.040	0.003				
	(0.033)	(0.005)				
Controls‡	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.58	0.16	0.57	0.16	0.63	0.15
R^2	0.32	0.30	0.33	0.32	0.58	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): The coefficient on Kresy in even columns reflects the average difference in historical literacy rates between counties in Kresy and in Central Poland from which respondents' ancestors originated. Regressions are run at the ancestor level, using data from our Ancestry Survey. Standard errors clustered by individual respondents in odd columns and using two-way clustering by individual respondents and by county of origin in even columns. * 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 the 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

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Secondary Education	Higher Education	Secondary Education	Higher Education	Secondary Education	Higher Education
Sample:	Urban or Rural		Urban origin & destination		Rural origin & destination	
$\Delta Edu(i)$	0.027**	0.042***	0.012	0.041	-0.028	-0.005
	(0.014)	(0.011)	(0.028)	(0.027)	(0.026)	(0.016)
Observations	1,391	1,391	323	323	347	347

Notes: The table combines data from our Ancestry Survey with Diagnoza data. The table provides the results from estimating equation (3). This addresses the possibility of individual selection of voluntary migrants from Central Poland to the 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, if anything, slightly *better* educated than a reasonable comparison group – people who still live in the places of their ancestors' origin in Central Poland. * p<0.1, ** p<0.05, *** p<0.01.

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

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Analysis:	Congestion?		Return to Schooling?		Out-Migration?		Fertility?	
Dep. Var.:	Years of Education		log(HH income)		Intend to go abroad		Share Children in HH	
Sample	WT	WT	all	WT	all	all	all	# children ≥ 1
Ancestor from Kresy	0.761***	0.777***	0.067	0.013	0.003	0.003	-0.002	0.004
	(0.079)	(0.083)	(0.044)	(0.050)	(0.006)	(0.006)	(0.005)	(0.008)
% Polish speakers, 1900 (std)	0.287***							
	(0.069)							
% Polish sp (std) × Kresy	0.102							
	(0.074)							
Share Autochthons, 1950 (std)		0.202**						
		(0.079)						
Sh 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	0.11	0.33
R-squared	0.30	0.30	0.22	0.21	0.09	0.09	0.26	0.24
Observations	7,307	7,307	15,932	3,875	22,096	22,066	29,105	10,126

Notes: The table examines four alternative mechanisms that may explain the education advantage of people with Kresy ancestors: congestion due to the presence of autochthons (columns 1-2); differential returns to education (columns 3-4); differential out-migration (columns 5-6); and differential fertility (columns 7-8). None of these appear to confound the coefficient on Kresy. Regressions are run at the level of respondents in Diagnoza; standard errors clustered by country. * p<0.1, ** p<0.05, *** p<0.01. WT = Western Territories.

‡ 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

I Background

Forced Kresy migrants just before leaving Kresy and upon arrival to WT

Figures A.1 and A.2 presented below exhibit the images of forced Kresy migrants right before leaving Kresy and right after arriving to the Western Territories. The online exhibition of the Polish History Museum devoted to forced migrants provides the following testimony as a caption to the image in the first figure: *"And so it happened that ... the marshall came: 'Leave' — 'But where should I go?'— 'To Poland.' And I say: 'I am in Poland.' And he says: 'This is not Poland anymore.' "*[1]

Figure A.1: Forced Kresy Migrants before their Departure from Kresy, Hłyboka (Ukraine), 1946.

Source: The collection of Polish History Museum.

[1]Edward Jaremko (cited by S. Ciesielski, *Exit. Kresy Wschodnie – Ziemie Zachodnie*), online exhibit https://artsandculture.google.com/exhibit/mwLihxsZye49Lw?hl=pl (Accessed on May 17, 2018).

Figure A.2: Forced Migrants from Kresy with their Belongings Arriving to Bielawa, former Langenbielau (a locality in the Western Territories), 1946.

Source: Figure 29 in Zaremba (2012).

Promotional poster for voluntary migrants from Central Poland to the Western Territories

Figure A.3 displays a typical example of posters that were used by the authorities in Central Poland to entice voluntary migration to the Western Territories.

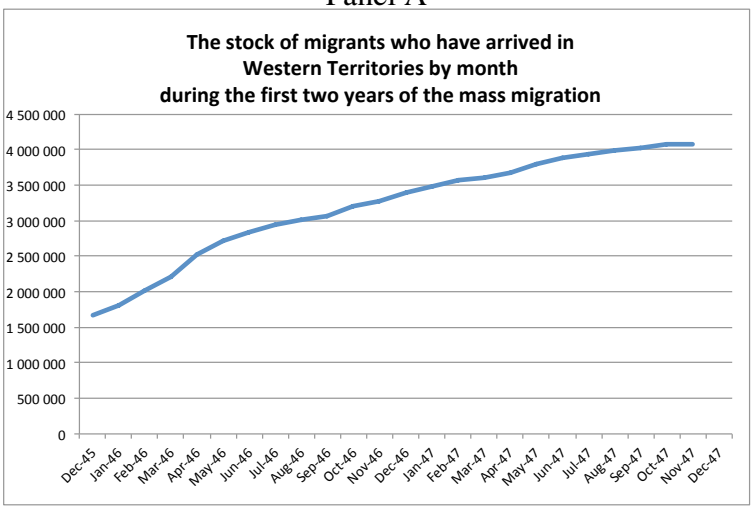
Figure A.3: Advertising to Attract Migrants from Central Poland to the Western Territories

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."

The timing of mass migrations from Kresy and Central Poland

Figure A.4 illustrates that forced migrants from Kresy and voluntary migrants from Central Poland arrived in the Western Territories (WT) at the same time. Panel A shows data on the stock of migrants who had arrived in WT by month, during the first two years of mass migration. The data start in December 1945 and show that by then, 1.5m migrants had moved into WT. That stock continued to grow steadily, reaching more than 4m migrants by the end of 1947. Panel B displays the share of Kresy migrants in that stock over time, separately for urban and rural destinations. Kresy migrants accounted for 40-50% of all migrants throughout this two-year window, in both urban and rural destinations. This suggests that Kresy migrants and 're-settlers' from CP (the official label used by the Polish authorities) arrived in parallel throughout the whole period. Thus, a potential concern that CP migrants moved into WT more quickly, generating a potential congestion effect for Kresy migrants, is not warranted.

Panel A



Panel B

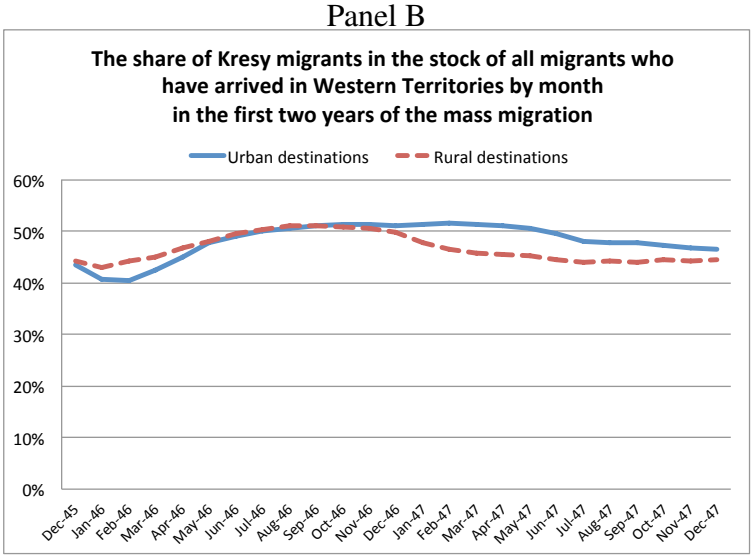


Figure A.4: The Timing of Arrival of Migrants to the Western Territories

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).

Places of origin of ancestors

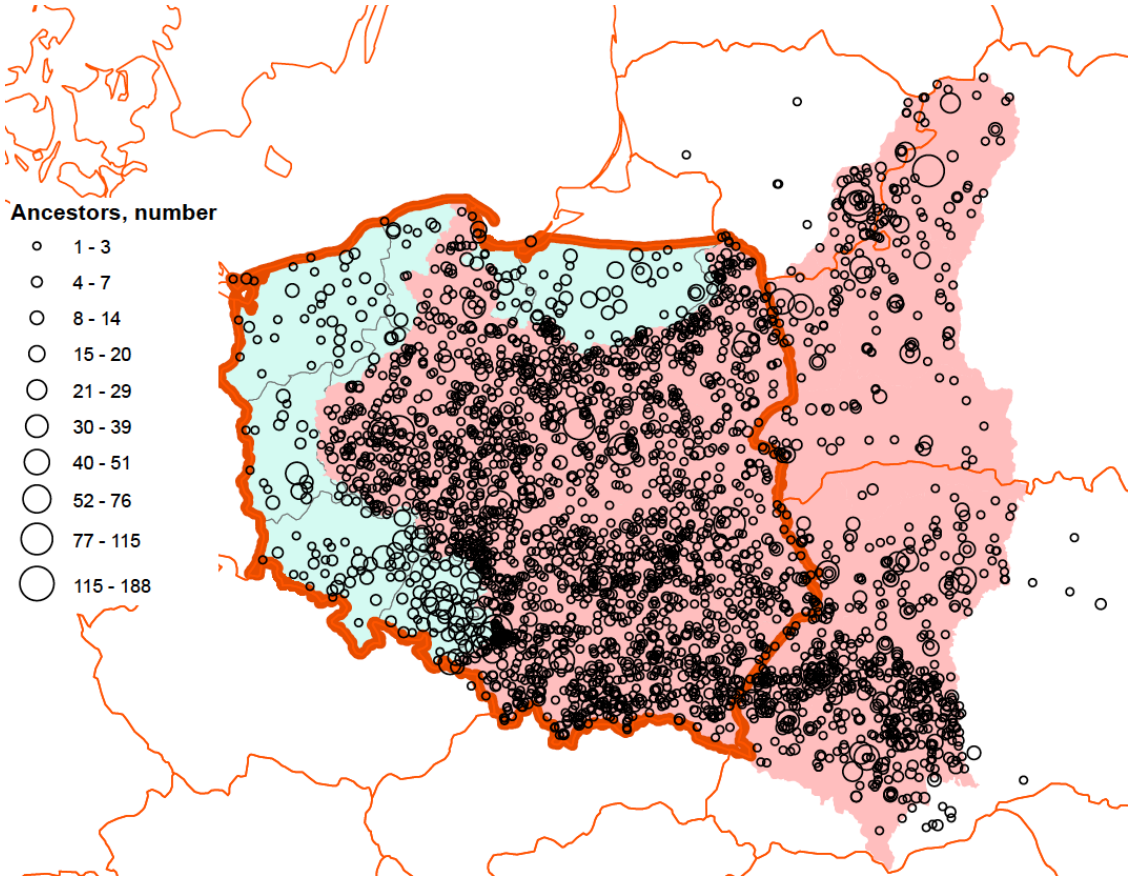


Figure A.5: Origin of Ancestors in our Ancestry Survey.

Note: The figure displays the origin of ancestors in our Ancestry Survey. The different dot sizes indicate the number of ancestors from each respective location. The different areas on the map are described in the note to Figure 1 in the paper: In the east, the former Eastern Polish territories (Kresy); in the west, the Western Territories, and in the center, Central Poland.

II Summary Statistics

Summary statistics

Tables A.1 and A.2 present summary statistics for the main explanatory and dependent variables. Table A.1 below presents summary statistics for the variables we use to measure education in both surveys. Note that in our Ancestry Survey, there is no question on the years of education (see also footnote 42 in the paper). We infer this information from the answer to the questions about the educational degrees. We consider four categories: primary education, incomplete secondary education, completed secondary education, and higher education. Information necessary to construct these variables is present in both Diagnoza and our Ancestry Survey. We impute the years of education in the Ancestry Survey by using the average years of education for each of the four

education categories in Diagnoza, rounded to the nearest integer.

Table A.1: Summary Statistics for Education Variables

	Obs	Mean	Std. Dev.	Min	Max
			Diagnoza		
Education years	29,348	11.720	3.366	0	28
Secondary education	29,289	0.480	0.500	0	1
Higher education	29,289	0.195	0.396	0	1
Ancestry Survey: Respondent level, representative sample					
Education years	2,846	12.595	3.222	7	17
Secondary education	2,846	0.491	0.500	0	1
Higher education	2,846	0.220	0.415	0	1

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

Table A.2 describes variables measuring the origin of ancestors in both surveys. In the Diagnoza survey, 11.3% of respondents have at least one ancestor from Kresy; as one should expect, the share of respondents with Kresy origin is higher in Western Territories (27.2%) than in Central Poland (6.0%). In our Ancestry Survey, in the representative sample of the WT population (i.e., excluding the oversample of respondents with Kresy origin mentioned in Section 3.1), 31.3% of respondents have at least one ancestor from Kresy in the generation of the youngest adult in 1939. The mean share of ancestors from Kresy is 23.6%, from Western Territories –15.9%, from Central Poland – 60.5%, and from abroad – 1.3%. The mean share of ancestors from rural areas is 75.7%. The third panel of the table presents summary statistics at the respondent level for the whole sample of the Ancestry Survey, including the oversample of respondents with Kresy origin. The last panel in Table A.2 summarizes data at the ancestor level for the whole sample of the Ancestry Survey. About 23% of the ancestors are from the parent generation, 55% from the grandparent generation, and 22% from the great-grandparent generation.

Table A.2: Summary Statistics for Variables Describing the Origin of Ancestors

	Obs	Mean	Std. Dev.	Min	Max
Diagnoza: Poland, full sample					
(Any) Ancestor from Kresy	29,392	0.113	0.317	0	1
Diagnoza: Western Territories					
(Any) Ancestor from Kresy	7,389	0.272	0.445	0	1
Diagnoza: Central Poland					
(Any) Ancestor from Kresy	22,003	0.060	0.237	0	1
Ancestry Survey: Respondent level, representative sample					
(Any) 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 CP	2,846	0.605	0.441	0	1
Share of ancestors from WT	2,846	0.159	0.341	0	1
Share of ancestors from rural areas	2,811	0.757	0.377	0	1
Share of ancestors from abroad	2,843	0.013	0.078	0	1
Ancestry Survey: Respondent level, whole sample					
(Any) Ancestor from Kresy	3,716	0.463	0.499	0	1
Share of ancestors from Kresy	3,716	0.367	0.434	0	1
Share of ancestors from CP	3,716	0.501	0.449	0	1
Share of ancestors from WT	3,716	0.131	0.312	0	1
Share of ancestors from rural areas	3,671	0.735	0.391	0	1
Share of ancestors from abroad	3,712	0.012	0.074	0	1
Ancestry Survey: Ancestor level, whole sample					
Ancestor from Kresy	11,928	0.324	0.468	0	1
Ancestor from CP	11,928	0.516	0.5	0	1
Ancestor from WT	11,928	0.160	0.367	0	1
Ancestor from rural area	11,548	0.745	0.436	0	1
Ancestor female	11,928	0.497	0.382	0	1
Parent	11,928	0.229	0.420	0	1
Grandparent	11,928	0.547	0.498	0	1
Great-grandparent	11,928	0.225	0.417	0	1

Notes: The table shows summary statistics for ancestry variables in Diagnoza from 2015 and our Ancestry Survey from 2016. In both surveys, we consider the samples of individuals with non-missing information about Kresy origin. For Diagnoza, we further restrict the sample to respondents with non-missing information about educational attainment, which is known for all respondents in the Ancestry Survey.

III Migration Flows Implied by Survey Data vs. Historical Census

Diagnoza survey vs. 1950 Census

In this section, we check the quality of the ancestry data from our surveys against migration flows implied by the 1950 Polish Census. The Diagnoza Survey and the 1950 Census cover all of the Polish post-WWII territory. The data in the 1950 Census is available at the regional level, providing information on where respondents lived in 1939 and in 1950. This allows us to construct migration flows. We begin with migrants from Kresy (who Migrants indicated "USSR" as their place of residence in 1939). Figure A.6 compares the results of the Diagnoza survey with the 1950 Census. The left panel displays the share people (in each region) in 1950 who had lived in Kresy in 1939, plotted against the share of respondents with ancestors from Kresy in the 2015 Diagnoza Survey. The historical and contemporaneous shares line up very well for most regions.[2]

For population in the Western Territories, the 1950 Census provides information at the more disaggregated level of counties. We can thus compute the share of Kresy migrants in each WT county in 1950. We use this information to repeat the consistency check on the Diagnoza data in the right panel of Figure A.6. 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 Figure A.6 shows a tight relationship.

Ancestry survey vs. 1950 Census

Figure A.7 repeats the above exercise using our 2016 Ancestry Survey in combination with the 1950 Census. Recall that our Ancestry Survey was conducted only in the Western Territories. Correspondingly, we use the available county-level data from the 1950 Census for WT. 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, as in Diagnoza). The 1950 Census, in turn, provides information on overall 16 origin areas (i.e., areas of residence in 1939). These include Kresy, the Western Territories, and 14 regions in Central Poland. We thus compute, for each county in WT, the share of migrants from each of these 16 origin areas in 1950. We then map the origin location data from the Ancestry Survey to the same 16 origin areas. The left panel of Figure A.7 plots the county-level origin shares from the 1950 Census against those from our Ancestry Survey. The right panel restricts attention to migrants from Kresy, plotting the share of people of Kresy origin by county from our Ancestry Survey against the same share 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 Survey and in the Ancestry Survey gave reasonable answers to the questions about their ancestral places of origin.

[2]There are a few exceptions. For instance, Warszawa (Warsaw) is considerably below the regression line. This means that, while in 1950 few people of Kresy origin lived there because the majority moved straight to the Western Territories, in 2015 the share of Warsaw survey respondents with Kresy ancestors is considerably larger. This is likely driven by the capital city's attraction of educated people – among them the descendants of Kresy migrants.

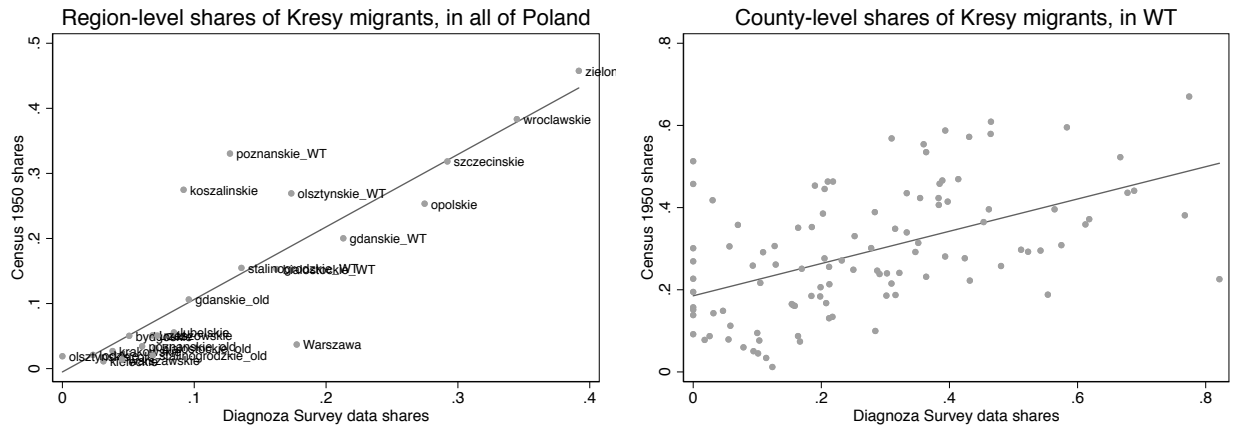


Figure A.6: Data Quality Check of Diagnoza Survey

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 regional level. 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 the Western Territories). The regression coefficient is 1.00 with a standard error of 0.057 and R^2 of 0.73. The right panel of the figure plots the county-level share of migrants from Kresy territories in the 1950 Census (y-axis) against the Kresy migrant share from the 2015 Diagnoza data. These more detailed 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.

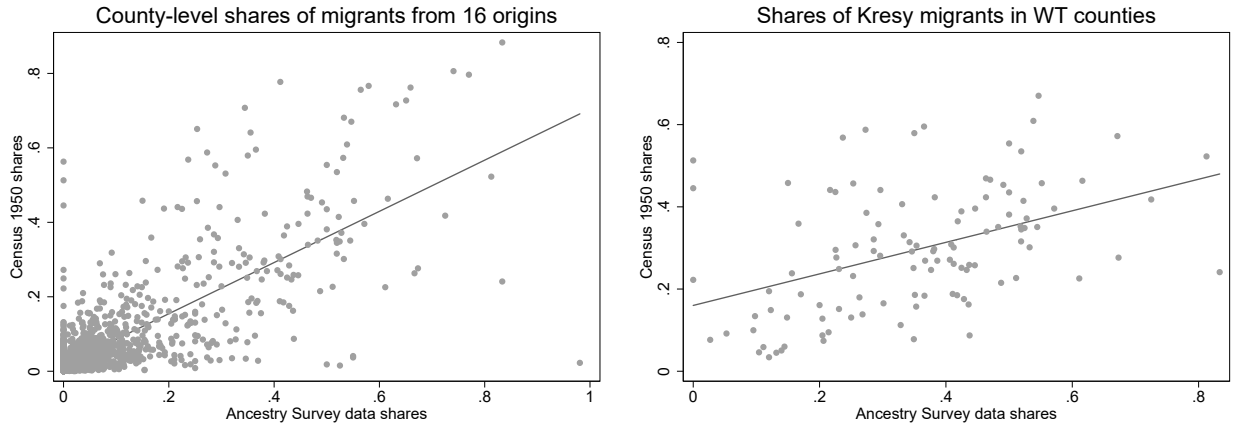


Figure A.7: Data Quality Check of our Ancestry Survey – WT Only

Note: The left panel plots the county-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.

IV Monte Carlo Simulations

In Section 4.3, we present the results of regressions estimated at the ancestor level. It is important to understand how the ancestor level results compare to the estimations at the respondent level. We run a series of Monte Carlo simulations to compare both the point estimates and the level of significance for the following two equations:

$$\text{Respondent-level: } Y_i = \beta\ Kresy_i + \phi'\ \mathbf{X}_i + \eta_{Locality_{(i)}} + \varepsilon_i, \tag{A.1}$$

$$\text{Ancestor-level: } Y_i = \gamma Kresy_{a_{(i)}} + \psi' \mathbf{A}_{a(i)} + \phi' \mathbf{X}_i + \eta_{Locality_{(i)}} + \varepsilon_{a_{(i)}} \tag{A.2}$$

Note that, in line with our specifications (1) and (2) in the paper, in the first equation above, $Kresy_i$ is respondent i's *share* of ancestors from Kresy; and in the second equation, $Kresy_{a_{(i)}}$ is a *dummy* that equals one if ancestor a of respondent i came from Kresy. In addition, we cluster the error term in the second equation at the respondent level.

The Monte Carlo Simulations yield the following results: Econometrically, the respondent-level and ancestor-level regressions are not equivalent. The estimated parameters β and γ, in general, are not equal; yet, the statistical inference, i.e., the significance of these parameter estimates, is similar.

First, we find that the parameters β and γ are equal only in the case when dummies for Kresy origin of different ancestors of the same respondent are perfectly correlated for all respondents. Formally, this means that for each respondent i, the indicators for Kresy origin of all ancestors of this respondent i in the generation of the youngest adults before the war are the same (i.e., $Kresy_{m_{(i)}} = Kresy_{f_{(i)}}$, where m and f are ancestors drawn at random from the full set of ancestors of respondent i in the considered generation, and this holds for all i).[3]

The parameter γ depends on the correlation between the indicators of Kresy origin of ancestors of the same respondent. The lower the correlation, the lower is γ (however, it is bounded below). If that correlation is zero, the parameter γ of the ancestor-level regressions is equal to the effect of the share of ancestors with Kresy origin of the respondent-level regressions (β), divided by the average number of ancestors per respondent (N), i.e., $\gamma = \beta/N$. More formally, the condition for equality of γ and β is that indicator variables for Kresy origin of any ancestor $a_{(i)}$ are i.i.d.

The parameter γ is within the interval $[\beta/N; \beta]$ as long as the correlation between indicator variables of Kresy origin of different ancestors of the same respondent is non-negative (i.e., if one ancestor drawn at random from the pool of all ancestors of all respondents has a Kresy origin, the other ancestor drawn at random from the set of ancestors of the same respondent is more likely to also be of Kresy origin than an ancestor drawn at random from the whole pool of all ancestors of all respondents).

In reality, the origins are positively correlated across ancestors of the same respondent, but this correlation is strictly below one, which means that we should expect smaller point estimates in the ancestor level regressions than in the respondent level regressions. In particular, the correlation between the dummies indicating the Kresy origin of spouses (e.g., of the mother and father or of the paternal grandmother and paternal grandfather of the same respondent) is over 90%. The

[3] If the considered generation of ancestors is parents, m and f are simply mother and father; if grandparents, these are two grandparents randomly drawn from the pool of all grandparents of the respondent i, etc.

correlation between dummies for Kresy origin 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. different great-grandparents, is 7%.

Second, the Monte Carlo simulations show that the level of statistical significance is similar between the respondent-level regressions and the ancestor-level regressions, when we cluster error terms at the respondent level. The level of significance is comparable irrespective of the level of correlation between the origins of different ancestors of the same respondent. Namely, when γ is below β, the standard errors are also proportionally smaller in the ancestor-level estimation, and therefore, statistical inference is similar.

Third, both of these facts are true not only for the estimation of the direct effects of Kresy ancestry (γ vs β), but also for the heterogeneity in the effects. In particular, when we consider an interaction term between the Kresy ancestor variables (share or dummy in the respondent-level and ancestor-level regression, respectively) and a characteristic of the place of origin of respondents ancestors (which is averaged across ancestors in the respondent-level regressions), we find that the statistical inference is similar in both cases. This is particularly important because in Section 5.2 of the main text, we show that the interactions between the characteristics of the origin locations and the dummy for Kresy origin of the respondent's ancestor are statistically insignificant.

To sum up, our Monte Carlo simulations show that t-statistics for the 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.

V Additional Evidence for the Main Result

In this section, we present additional evidence in support of our main result.

Results from Diagnoza with municipality FEs

Table A.3 replicates columns 1 to 5 from Table 2 from the paper with controls for municipality fixed effects instead of county fixed effects. The coefficient sizes are almost identical, suggesting that local unobservables do not confound our results.[4]

[4]Note that controlling for municipality FEs is a very restrictive specification in Diagnoza because there are very few respondents in smaller municipalities. As a consequence, often there is no variation in Kresy ancestry within a given small municipality (e.g., two out of two respondents having no Kresy ancestors). Thus, the coefficient on Kresy in Table A.3 is identified mostly from larger municipalities with many respondents.

Table A.3: Average Education of Individuals in the 2015 Diagnoza Survey

Dependent variable: Individual-Level Education, as indicated in each panel

	(1)	(2)	(3)	(4)	(5)
Sample:	All	Rural	Urban	Central Poland	Western Territories
Panel A. Dep. Var.: Years of education					
Ancestor from Kresy	0.903***	0.719***	0.965***	0.985***	0.807***
	(0.084)	(0.142)	(0.103)	(0.127)	(0.112)
Mean Dep. Var.	11.95	11.12	12.78	11.98	11.88
Observations	25,702	12,805	12,897	19,248	6,454
Panel B. Dep. Var.: Secondary education dummy					
Ancestor from Kresy	0.125***	0.110***	0.130***	0.127***	0.121***
	(0.012)	(0.024)	(0.014)	(0.017)	(0.017)
Mean Dep. Var.	0.51	0.38	0.64	0.51	0.51
Observations	25,703	12,808	12,895	19,242	6,461
Panel C. Dep. Var.: Higher education dummy					
Ancestor from Kresy	0.101***	0.073***	0.109***	0.130***	0.074***
	(0.012)	(0.018)	(0.015)	(0.018)	(0.015)
Mean Dep. Var.	0.21	0.13	0.30	0.21	0.21
Observations	25,703	12,808	12,895	19,242	6,461
Respondent municipality FE	✓	✓	✓	✓	✓
Controls‡	✓	✓	✓	✓	✓

Notes: The table replicates Table 2, columns 2-6, from the paper, using municipality fixed effects instead of county fixed effects. Data are from the Diagnoza Survey; standard errors clustered a 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.

Arbitrariness of the Kresy border

This subsection complements our discussion in Section 2.1 of the paper about the arbitrariness of the Kresy border and the Kresy border analysis presented in the results section. Figures A.8 and A.9 examine geo-climatic and agricultural characteristics of counties in a 150 km corridor around the Kresy border. There is no discontinuity at the Kresy border in any geo-climatic characteristic, such as mean temperature, precipitation, altitude, or terrain ruggedness. The same is true for the suitability for various major crops (barley, wheat, potato, and sunflower).

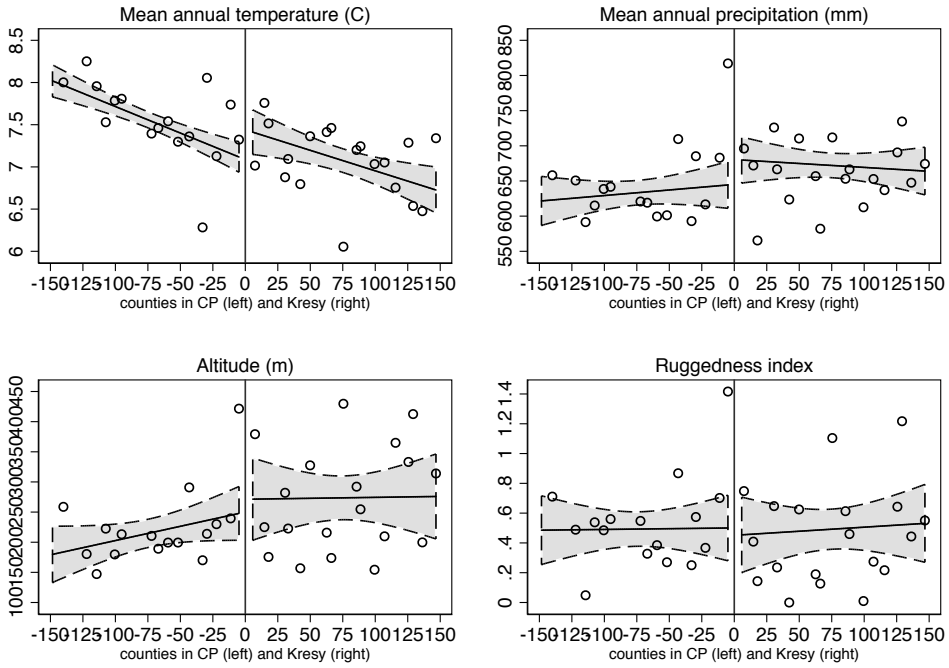


Figure A.8: Kresy Border Sample: Geo-climatic Characteristics

Note: The figure shows that there is no discontinuity around the border between Kresy and Central Poland in terms of geo-climatic characteristics. The figure uses data from FAO, averaged at the county level. Dots correspond to data aggregated into 8 km (5 miles) bins for visualization, while the lines are based on all underlying observations, with the shaded area representing 90% confidence intervals.

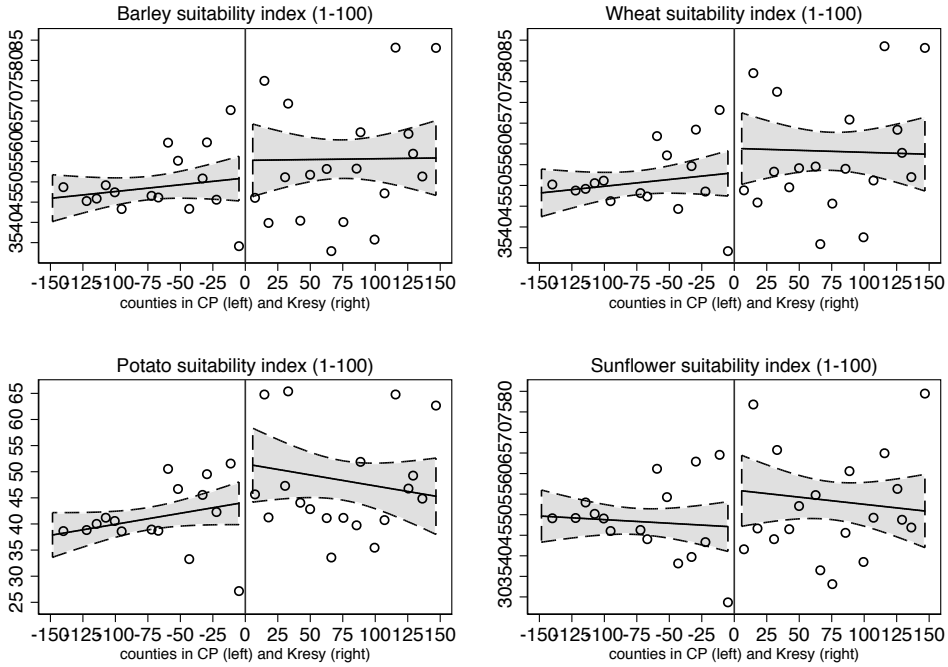


Figure A.9: Kresy Border Sample: Crop Suitability

Note: The figure shows that there is no discontinuity around the border between Kresy and Central Poland in terms of soil suitability. The figure uses data from FAO, averaged at the county level. Dots correspond to data aggregated into 8 km (5 miles) bins for visualization, while the lines are based on all underlying observations, with the shaded area representing 90% confidence intervals.

Border analysis in our Ancestry Survey – additional results

The results shown in this subsection complement our border analysis from Section 4.3 in the paper. Figure A.10 illustrates the border sample based on our Ancestry Survey data. It shows the locations of origin places for ancestors for those ancestors who came from within 150 kilometers from the Kresy border.

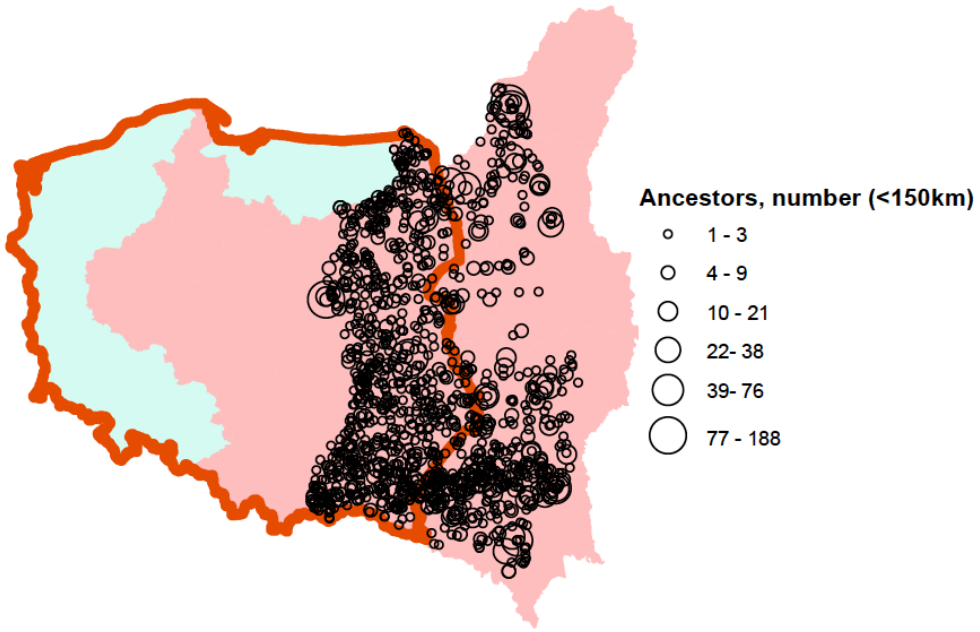


Figure A.10: Origin of Ancestors in our Ancestry Survey.

Note: The figure displays the origin of ancestors in the border sample of our Ancestry Survey – within 150km of the Kresy border. The different dot sizes indicate the number of ancestors from each respective location. The different areas on the map are described in the note to Figure 1 in the paper: In the east, the former Eastern Polish territories (Kresy); in the west, the Western Territories, and in the center, Central Poland.

Table A.4 complements the graphical evidence from Figure 4 of the main text. The table presents the results of our most demanding specifications: We identify the effect of ancestors' origin for individuals living within the same county (column 1) or even within the same municipality (columns 2 to 5) whose ancestors originate from localities close to the Kresy border using a spatial dimensional RDD that controls for a quadratic polynomials in latitude and longitude of the ancestor's origin. We estimate several specifications to illustrate the robustness of the main result displayed in Figure 4. In columns 1 to 3 of Table A.4, we use years of education as outcome variable and show that the results are robust to using samples within 150 and 100km from the Kresy border. In columns 4 and 5, we report the results for secondary and higher education dummies,

respectively. Results of all specifications are consistently strong and of similar magnitude as our main results for the Ancestry Survey in Table 5 in the paper.

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

Dependent variable: as indicated in column header

	(1)	(2)	(3)	(4)	(5)
Dep. Var.:	Years of education			Secondary	Higher
Notes on sample:	$< 150km$	$< 150km$	$< 100km$	$< 150km$	$< 150km$
Ancestor from Kresy	0.678*	0.948**	1.344**	0.116**	0.126**
	(0.354)	(0.390)	(0.523)	(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 the Western Territories, using only ancestors from within the indicated distance from the Kresy border. Regressions are run at the ancestor level; robust standard errors clustered at the respondent level indicated in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.
‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties, 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 ancestors' location of origin.

VI Additional Results on Potential Alternative Mechanisms

Table A.5 below presents data from the 1921 Polish Census, separately for Kresy and Central Poland. It shows that, on average, literacy rates among Roman Catholics were lower in Kresy than in Central Poland before WWII and that this difference was more pronounced in rural areas than in urban areas.

Table A.5: Literacy rates of Poles in Kresy and Central Poland parts of the SPR

1921 Polish Census:	Kresy	Central Poland
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

Heterogeneity in the effect depending on characteristics of origin locations

In Tables A.6 and A.7 we test for possible differential effects of Kresy origin depending on characteristics at the ancestors' place of origin. In particular, we run regressions at the ancestor level, in which we include interactions between the dummy for Kresy ancestry and (standardized) county-level characteristics of the place of origin of the ancestor, as well as characteristics of the place of origin of the ancestor themselves.[5]

Table A.6 examines the heterogeneity with respect to various measures of diversity at the origin location. In particular, we consider the following pre-WWII county level variables: the share of Roman Catholics, the share of Polish speakers, the share of Ukrainian speakers, the share of Russian speakers, the total literacy rate and the literacy rate among Roman Catholics, as well as the urbanization rate. Table A.7 considers heterogeneity with respect to land suitability for wheat (which was the main crop in pre-WWII Kresy), mean temperature, the precipitation-evatranspiration ratio, and ruggedness of the origin locations. We find no differential effects of Kresy origin on years of education with respect to any of these characteristics. This evidence suggests that the effect of Kresy origin is driven by forced migration itself, rather than by the characteristics of the origin of Kresy migrants.

[5]Since we use 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.

Table A.6: No Heterogeneous Effects with Respect to Ancestors' Origin Characteristics

Dependent variable: Years of education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ancestor from Kresy	0.542***	0.540***	0.572***	0.466***	0.508***	0.579***	0.502***	0.505***
	(0.131)	(0.141)	(0.105)	(0.105)	(0.108)	(0.096)	(0.104)	(0.098)
Share Rom. Cath., 1931 (std)	0.049							
	(0.114)							
Rom. Cath., 1931 (std) × Kresy	-0.042							
	(0.141)							
Share Polish speakers, 1931 (std)		0.028						
		(0.136)						
Polish speakers, 1931 (std) × Kresy		-0.006						
		(0.167)						
Share Ukrainian speakers, 1931 (std)			-0.001					
			(0.124)					
Ukrainian speakers, 1931 (std) × Kresy			-0.077					
			(0.126)					
Share Russian speakers, 1931 (std)				0.209				
				(0.212)				
Russian speakers, 1931 (std) × Kresy				-0.181				
				(0.213)				
Literacy rate, 1931 (std)					-0.027			
					(0.080)			
Literacy rate, 1931 (std) × Kresy					0.045			
					(0.093)			
Urbanization rate, 1931 (std)						0.047		
						(0.060)		
Urbanization rate, 1931 (std) × Kresy						-0.087		
						(0.058)		
Literacy rate, 1921 (std)							0.005	
							(0.075)	
Literacy rate, 1921 (std) × Kresy							-0.007	
							(0.093)	
Literacy rate Rom. Cath., 1921 (std)								0.015
								(0.066)
Literacy rate Rom. Cath., 1921 (std) × Kresy								0.002
								(0.085)
Controls‡	✓	✓	✓	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	13.14	13.14	13.14	13.14	13.14	13.15	13.14	13.14
R^2	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
Observations	9,706	9,706	9,706	9,706	9,667	8,613	9,645	9,645

Notes: The table uses data from our Ancestry Survey. Regressions are run at the ancestor level. The table shows that the coefficient on Kresy ancestry does not vary significantly with average characteristics of the population at the place of origin. Standard errors clustered using two-way clustering by individual respondents and by county of origin. * p<0.1, ** p<0.05, *** p<0.01.

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

Table A.7: No Heterogeneous Effects w.r.t. Geographic Features at Ancestors' Origin

Dependent variable: Years of education

	(1)	(2)	(3)	(4)
Ancestor from Kresy	0.589***	0.480***	0.588***	0.558***
	(0.100)	(0.117)	(0.103)	(0.096)
Land suitability for wheat at origin (std)	-0.041			
	(0.082)			
Land suit. for wheat (std) × Kresy	0.014			
	(0.096)			
Annual temperature at origin (std)		0.030		
		(0.089)		
Annual temperature (std) × Kresy		-0.169		
		(0.115)		
Precip.-evatranspiration ration at origin (std)			-0.015	
			(0.064)	
Precip.-evatranspiration ration (std) × Kresy			-0.052	
			(0.099)	
Ruggedness at origin (std)				0.034
				(0.048)
Ruggedness (std) × Kresy				-0.088
				(0.082)
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	8,793	8,793	8,793	8,793

Notes: The table uses data from our Ancestry Survey. Regressions are run at the ancestor level. The table shows that the coefficient on Kresy ancestry does not vary systematically with geographic characteristics at the place of origin. Standard errors clustered using two-way clustering by individual respondents and by county of origin. * p<0.1, ** p<0.05, *** p<0.01.

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

Kresy migrants with destination in the Western Territories vs. Central Poland

Table A.8 restricts the Diagnoza sample to respondents with Kresy ancestors. It compares their education in the Western Territories and in Central Poland. Odd columns in Table A.8 show the raw differences (after controlling for individual characteristics). Note that we cannot control for local fixed effects in these specifications because the table compares individuals with Kresy ancestors *across* regions. Thus, differences in local labor markets affect the results. To account for at least some of this variation, even columns include an indicator for individuals who live in the counties of Warsaw or Krakow – the main university centers in Poland. The results imply that controlling for these educational centers is important, as it reduces the difference between WT and CP. We find that – after accounting for Warsaw and Krakow – respondents with Kresy ancestors who live in the Western Territories have, on average, 0.44 fewer years of education and are 5.0 and 6.0 percentage points less likely to complete secondary and higher education, respectively, as

compared to respondents with Kresy ancestors who live in Central Poland.[6] Thus, our Ancestry Survey results in the Western Territories – which show a significant education advantage of people with Kresy ancestors – are, if anything, underestimating the effect for Poland overall.

Table A.8: Education of Kresy Migrants in the Western Territories and Central Poland

Dependent variable: as indicated in column header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	Years of education		Secondary education		Higher education	
Dummy for Western Territories	-0.652***	-0.436***	-0.065***	-0.050**	-0.083***	-0.060***
	(0.135)	(0.136)	(0.019)	(0.020)	(0.019)	(0.019)
Warsaw or Krakow		2.225***		0.153***		0.236***
		(0.349)		(0.033)		(0.044)
Controls‡	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	12.58	12.58	0.61	0.61	0.29	0.29
R-squared	0.33	0.34	0.21	0.21	0.19	0.20
Observations	3,298	3,298	3,294	3,294	3,294	3,294

Notes: Regressions are run at the respondent level, restricting the sample to individuals with ancestors from Kresy in the Diagnoza Survey. Standard errors are clustered at the household level. * p<0.1, ** p<0.05, *** p<0.01. 'Warsaw or Krakow' is an indicator that takes on value one for the counties of Warsaw and Krakow.
‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties.

Congestion

Figure A.11 illustrates that the county-level share of autochthons in the 1950 Polish Census is highly correlated with the share of Polish speakers in the German Census of 1900. The 1900 German Empire Census was the last census in the German Empire that collected information on language spoken at home. Autochthons in the 1950 Polish Census are the people who had lived in the territories that Germany lost to Poland as a result of WWII and were not expelled, as they declared themselves to be Polish. Figure A.11 illustrates that autochthons are indeed largely people with ethnic Polish ancestry. They 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.

Out-migration

Figure A.12 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 Polish Census. The latter data are available at the regional level. The high correlation shown in the figure suggests that intention to emigrate measures something meaningful, as in previous years the same regions indeed saw larger realized emigration.[7] It supports

[6] Note that the counties Warsaw and Krakow are geographically smaller than commuting zones. When we account for larger areas – by using indicators for the Voivodeships of Mazowieckie and Lesser Poland (Małopolska), i.e., the areas around Warsaw and Krakow – the coefficients on Western Territories become even smaller.

[7] A linear regression yields a coefficient of 0.65 with a standard error of 0.18 and an R^2 of 0.53.

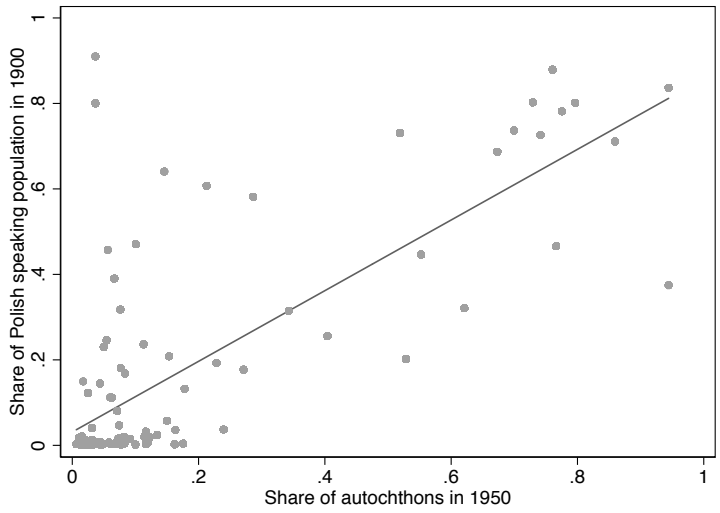


Figure A.11: Two Alternative Measures of the Share of Autochthons across WT Counties

Note: The figure plots the share of Polish speakers in the German Empire Census in 1900 against the share of autochthons in the 1950 Polish Census. The line shows a linear regression with coefficient of 0.83 and a standard error of 0.07; the R^2 is 0.57.

the validity of the evidence presented in Table 11 in the paper, which shows that the intention to emigrate does not differ for those with Kresy ancestors.

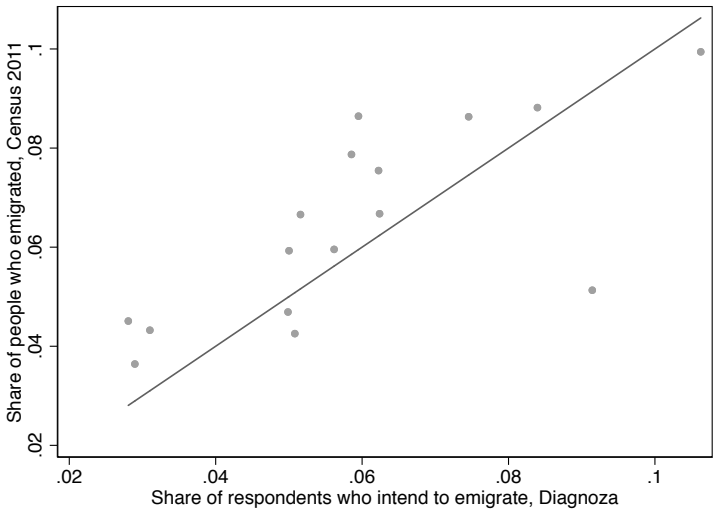


Figure A.12: Stated Intent to Emigrate vs. Emigration Rates

Note: The figure plots the share of respondents who intend to emigrate (Diagnoza 2015) against the share of people who emigrated (from the 2011 Polish Census) at the regional level. The figure also displays a 45-degree line.

Moving as communities

Table A.9 investigates whether migrants from Kresy tended to move more (or less) together with people from their origin location, as compared to migrants from Central Poland. We compute, for each municipality in WT, the number of ancestors in our Ancestry Survey who are from the same county of origin. We refer to this measure as the 'size of the local ancestor community.' This is likely to be a noisy measure, as it is based on a count within our survey alone. Note also that this number will mechanically tend to be larger in municipalities for which we have a higher number of ancestors in our sample. We thus control – for each municipality – for the total ancestors in the sample.

Table A.9 checks whether the size of local ancestor community is related to the Kresy origin of migrants, and whether our results are robust to controlling for this measure. Column 1 shows that there is no relationship between Kresy origin and the size of local ancestor communities. In other words, Kresy migrants are not more (or less) likely to live in municipalities with many migrants from the same origin. In column 2, we show that our main result from specification (2) also holds in the subsample for which we can construct the size of the local ancestor community.[8] In column 3, we use the size of the local ancestor community as a control, showing that the relationship

[8]The smaller sample is explained by two factors: First, to construct the size of the local ancestor community, we can only use data from our representative sample in the Ancestry Survey (see Section 3.2 and in particular footnote 23 in the paper). We need to exclude the oversample of people with Kresy ancestors to avoid that the community size from Kresy is overestimated. Second, we only compute the size of the local ancestor community for migrants from Kresy and Central Poland. We exclude ancestors from WT because these are autochthons, while the focus here is on *migrant* communities. In addition, we exclude ancestors from abroad because the community variable is undefined for them.

between Kresy origin and educational attainment is essentially unchanged. Finally, columns 4 and 5 show that our results for secondary and higher education are also robust to controlling for the size of the local ancestor community. Overall, Table A.9 suggests that our results are unlikely to be driven by variation in the size of the local community of people with common origin.

Table A.9: Size of Ancestor Communities in each Municipality: Ancestor-Level Data

Dependent variable: as indicated in column header

Dep. Var.:	(1) Size of local ancestor community[#]	(2) Years of education	(3) Years of education	(4) Secondary education	(5) Higher education
Ancestor from Kresy	-0.020	0.432***	0.433***	0.055***	0.032*
	(0.260)	(0.116)	(0.115)	(0.020)	(0.017)
Size of ancestor community[#]			-0.037*	-0.007***	-0.003
			(0.021)	(0.003)	(0.004)
Total ancestors in sample	0.010***		-0.002*	0.000	-0.000
	(0.001)		(0.001)	(0.000)	(0.000)
Controls[‡]	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓
Mean Dep. Var.	2.52	13.12	13.12	0.57	0.28
R^2	0.24	0.30	0.30	0.22	0.24
Observations	7,093	7,093	7,093	7,093	7,093

Notes: The table uses data from our Ancestry Survey. Regressions are run at the ancestor level; robust standard errors clustered at the municipality level in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.
[‡] Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties.
[#] This variable is constructed for each municipality in our Ancestry Survey sample. It measures the total number of ancestors who came from the same county of origin.

Recall bias: Missing information about ancestor origin locations

Table A.10 examines the role of missing information about ancestors in our 2016 Ancestry Survey in the Western Territories. We compute the share of ancestors with missing information as follows for each respondent: Let $N_a(i)$ be the number of ancestors for whom respondent i reported the location of origin. Remember that our Ancestry Survey asked for information about the generation of ancestors who were the youngest adults in the respondent's family in 1939. For this generation, let $N_{max}(i)$ denote the maximum possible number of ancestors (e.g., $N_{max}(i) = 4$ for the grand-parent generation). Then, the share of i's ancestors for whom information is missing is given by $1 - N_a(i)/N_{max}(i)$.

Column 1 in Table A.10 shows that missing information on ancestors is unrelated to Kresy origin in our baseline Ancestry Survey regression (which is run at the respondent level – see column 2, Panel A, in Table 5 in the paper). More specifically, the excluded category in this regression is the share of ancestors from Central Poland. Thus, the zero coefficient on the share of Kresy ancestors means that respondents with ancestors from Kresy are just as likely as those with ancestors from Central Poland to remember their ancestors. This makes it unlikely that any of our results are

confounded by missing information on ancestors. Note also that the mean of the dependent variable in column 1 is 0.12. That is, the share of ancestors with missing information is only 12% in our Ancestry Survey. Finally, the coefficient on the share of ancestors from WT in column 1 is negative and significant, meaning that respondents are more likely to remember the location of their ancestors in the Western Territories. This is not surprising, given that our survey was conducted in WT.

In the remaining columns in Table A.10, we use our education measures as outcome variables. Column 2 shows that there is a significantly negative relationship between years of education and the share of missing ancestor information. This is what one would expect: More educated respondents tend to be better informed about their ancestors. Columns 3-5 replicate the specification from columns 2, 5, and 6 in Panel A of Table 5 in the paper, adding the share of missing ancestor information as an additional control. The coefficients on the share of Kresy ancestors are literally unchanged. Thus, missing information about ancestor origin locations does not confound our results.

Table A.10: Accounting for Missing Ancestor Information in the Ancestry Survey

Dependent variable: as indicated in column header

	(1)	(2)	(3)	(4)	(5)
Dep. Var.:	Share missing ancestor info†	Years of education		Secondary education	Higher education
Share of Ancestors, Kresy	0.000		0.812***	0.110***	0.070***
	(0.010)		(0.139)	(0.021)	(0.018)
Share of Ancestors, WT	-0.059***		-1.079***	-0.166***	-0.138***
	(0.016)		(0.195)	(0.032)	(0.024)
Share of Ancestors, abroad	-0.126***		-1.267	-0.032	-0.015
	(0.038)		(0.867)	(0.113)	(0.100)
Share of Ancestors, rural	0.001		-0.458***	-0.063**	-0.034*
	(0.011)		(0.163)	(0.024)	(0.020)
Share missing ancestor info†		-0.671**	-0.865***	-0.125**	-0.122***
		(0.287)	(0.285)	(0.050)	(0.040)
Controls‡	✓	✓	✓	✓	✓
County FE	✓	✓	✓	✓	✓
Mean Dep. Var.	0.12	12.44	12.44	0.47	0.22
R^2	0.21	0.27	0.29	0.21	0.21
Observations	3,581	3,581	3,581	3,581	3,581

Notes: The table examines the role of missing information about ancestors in our 2016 Ancestry Survey in the Western Territories. Columns 3-5 replicate the specification from columns 2, 5, and 6 in Panel A of Table 5 in the paper, adding the share of missing ancestor information as an additional control. Regressions are run at the respondent level; robust standard errors in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

† For each respondent, the share of ancestors with missing information is computed specific to the generation of ancestors who were the youngest adults in the respondent's family in 1939. For example, if those were the grandparents, and the historical location for three out of four grandparent is known, then the share missing is 0.25.

‡ Controls include respondents' gender, age, age^2, dummies for six age groups, as well as indicators for rural places and urban counties, an indicator for the generation of the ancestor, and rural location of the ancestor.

Quick Guide to Mechanisms

The following is a brief guide to potential mechanisms behind the main result in the paper, namely, that Poles from the former Eastern Borderlands of Poland (Kresy) whose ancestors were forced to migrate after WWII invest significantly more in education than other Poles. We first present a table that summarizes historical and empirical evidence for the most likely mechanism behind our finding. We then present a second table that discusses alternative mechanisms, together with historical and empirical evidence that renders these alternative mechanisms unlikely. Note that the following tables focus only on possible mechanisms and do not discuss the empirical evidence for the main finding of the paper (namely, the educational advantage of Poles with Kresy origin).

Most likely mechanism: *Our empirical findings suggest that our main result is driven by a shift in preferences from investing in physical possessions towards investment in human capital, as a consequence of the loss of physical belongings during the expulsion.*	
Type of evidence: **H**istorical/**E**mpirical	Description of evidence
H: Section 1 – Introduction	Memoirs written by Kresy migrants in Western Territories in the 1950s suggest a *change* in preferences towards education in the aftermath of forced migration, for example: *"In Western Territories, there was a specific situation. People did not attach great importance to material wealth. ... In a new life situation, the cult of new values emerged, i.e., values that are indestructible, that cannot be lost, and that die with the man – the cult of knowledge, of skills, which can resist cataclysms."* This is also supported by interviews with descendants of forced migrants, e.g., with the former president Komorowski who stated: "*At home, nobody attached any importance to the material side, because everything that was valuable had been lost*."
H/E: Section 2.3	Historical evidence by sociologist Irena Turnau suggests an immediate shift towards higher school enrolment among children of Kresy migrants after the expulsion. Turnau assembled data on schooling in Wrocław (the former German Breslau) in 1948. She found that children of Kresy migrants were over-represented among secondary school students, and even more so among students in higher education.
E: Table 3 and Figure 3	Cohort-specific empirical evidence shows that this immediate shift is also true for educational attainment: The education effect is not present for forced migrants who had completed schooling before they were forced to migrate; while it is present for children of forced migrants who had the chance to complete education after migration.
E: Table 7	Evidence from the large-scale Diagnoza Survey shows that descendants of forced migrants value material goods less, while having a stronger aspiration for education of their children. They also possess fewer physical assets, relative to the number of physical assets they can afford. These results hold even when controlling for the level of education of the individual respondents, suggesting that different *preferences* among Kresy descendants drive the results (as opposed to Kresy descendants' higher own education explaining their aspiration for their children's education).
Literature: Section 1	Our preferred interpretation of the results is consistent with a robust body of existing evidence that describes how individual preferences change in response to exposure to violence, natural disasters, or economic shocks. Recent evidence suggests that these effects persist in future generations. We cite over a dozen related publications in the Introduction.

Alternative mechanism 1. Pre-existing skills / preferences *People from Kresy may have had higher education or different preferences for education already before WWII.*	
Type of evidence: **H**istorical/**E**mpirical	Description of evidence
H: Section 2.1 under heading "Poles in Kresy and Central Poland before WWII"	Same access to education for Poles in Kresy and CP before WWII (it was the same country). Also, no discrimination against Poles in Kresy.
E: Figure 3 and column 1 in Table 3	No difference in education for Kresy migrants among the pre-1930 cohort (that had finished schooling by the time of expulsions).
E: Figure 4 and column 7 of Table 2	RDD along the Kresy border (note that this border was arbitrarily drawn – see Section 2.1 under heading "Arbitrariness of the Kresy border of 1945"). Diagnoza Survey: i) No difference in pre-WWII education along Kresy border. ii) Kresy descendants are more educated than descendants of inhabitants ("*stayers")* to the West of the Kresy border. The remaining possibility is that "stayers" were negatively selected. This is addressed by Figure 5 (see next point).
E: Figure 5 and Table A.4	RDD along the Kresy border. Ancestry Survey: Comparing individuals within municipalities in WT. Kresy descendants are more educated than descendants of *movers* from the area to the West of the Kresy border. In combination, Figures 4 and 5 make it unlikely that selection drove our results.*

* Note: This point holds unless one reverts to the following (unlikely) explanation – a mix between a story of pre-existing skills and selection: Outmigration from the area in CP to the West of the Kresy border could have been such that i) unskilled migrants moved to WT; ii) skilled migrants moved to other places in CP. Point i) would explain Figure 5. Also, if flow ii) was large, the stayers to the West of the Kresy border would be less educated, explaining Figure 4. Note that (in addition to the purely speculative presumption about skill-biased migration, which cannot be examined in the data and for which there is no historical evidence), this would require a larger outflow from the area to the West of the Kresy border to CP than to WT (only this would yield relatively less educated stayers). To check this, we use the 1950 Census and examine outmigration from Polish regions (*wojewodztwa*) next to the Kresy border (to its West) to other regions in CP and in WT. We find that the overall flow from the area to the West of the Kresy border to CP was between 5% and 8%, while the outflow to WT was between 14% and 25%. Thus, the overall flow from the area to the West of the Kresy border to CP was much *smaller* (less than one-third) than the flow to WT. Consequently, the alternative interpretation outlined at the beginning of this note is not compatible with the data.

Alternative mechanism 2. Selection hypothesis *People from either Kresy or from other parts of the country differentially selected into specific locations or occupations.*	
Type of evidence: **H**istorical/**E**mpirical	Description of evidence
H: Section 2 **E**: Table 8	*Selection into migration from Kresy?* The historical narrative clearly speaks against selection out of Kresy: The vast majority of ethnic Poles in Kresy had no choice but to leave Kresy. This is particularly true for urban areas and for Ukraine. In Table 8, we confirm that our results hold equally in urban vs. rural areas and in the subset of the Ukrainian part of Kresy.
E: Tables 9 and 10	*Selection of voluntary migrants from CP to WT?* First, note that this type of selection would not affect our results for Poland overall (Table 2). We present two analyses, showing that both *regional* and *individual* selection of voluntary migrants is unlikely to affect our results within WT (i.e., from our Ancestry Survey). On regional selection, see Table 9, and on individual selection, see Table 10 and the corresponding description in the text. For both, we find that if anything, the evidence points to *positive selection* of voluntary migrants from CP, which would imply *smaller* effects of Kresy origin.
E: Tables 1, 2 & A.8	*Selection of Kresy migrants into WT vs. CP?* Three quarters of Kresy migrants moved to WT and one quarter to CP (Table 1). Did the most able Kresy migrants move to WT, explaining why Kresy migrants in WT are more educated? No: Table 2 (columns 5 and 6) show that the coefficients on Kresy ancestry are, if anything, larger in CP than in WT. Table A.8 performs an additional check, showing that respondents with Kresy origin are somewhat *less* educated in the WT than in CP. This confirms that, if anything, our results for the Western Territories are a lower bound on the effect of Kresy origin.

Alternative mechanism 3. Differential access at destinations *After moving, people from Kresy were either given preferential or restricted access to schooling or employment opportunities.*	
Type of evidence: **H**istorical/**E**mpirical	Description of evidence
H: Section 2.2 under heading "The arrival of migrants to the WT"	The historical narrative is clear: in Western Territories (ex-German areas), there was equal access to education, land, houses, and productive assets for Poles from Kresy and CP. There was neither affirmative action for Kresy people nor discrimination against them.
E: Table 11	*Differential congestion*: Locations in Western Territories with a higher share of autochthons might generate congestion that limits access to assets. However, Table 11 suggests that there is no differential effect of such potential congestion on education of Kresy migrants. Underlying this finding is the fact that Kresy migrants were not systematically resettled to areas in WT with more/fewer autochthons.
E: Figure A.4	*Preferential access for CP migrants*: Voluntary migrants from Central Poland were closer to Western Territories and might have grabbed the best opportunities before Kresy migrants arrived. This would be a story of congestion for Kresy migrants because of fast-moving CP migrants. However, Figure A.4 suggests that CP and Kresy migrants arrived into WT in parallel throughout.
E: various tables: location fixed effects	*Differential assignment to locations*: We routinely use county fixed effects or even municipality fixed effects, i.e., we compare survey respondents within the same location. If different groups of migrants were assigned differently to different locations, our within-location comparison removes such worry.
E: Tables 2,5,8,10	*Differential assignment within locations:* Voluntary migrants may have been attracted by the promise that they would receive land, potentially making it more likely that they were given land and thus worked in agriculture within destination locations. This is very unlikely, given that our results hold within the subsample of urban locations. In urban municipalities, the share of farmers among all occupations is smaller than 1%.