Uprootedness, Human Capital, and Skill Transferability*

Stelios Michalopoulos† Elie Murard‡ Elias Papaioannou§ Seyhun Orcan Sakalli¶

March 11, 2025

Abstract

More than a century has passed since the abrupt exodus of 1.2 million Greek Orthodox from Anatolia and their resettlement in Greece, a transformative event for the country's social and demographic landscape. Today, more than one in three Greeks reports a refugee background. While its historical significance is well-documented, its short-, medium-, and long-term impact on human capital accumulation remains unexplored. How did forced displacement shape the educational trajectories of the uprooted and their offspring? Did refugees invest in portable skills to respond to uncertainty, or did they struggle to catch up with the autochthonous? To address these questions, we trace the educational investments of refugees and their descendants over the last 100 years, leveraging granular census data and a comprehensive mapping of both their origins in Anatolia and their settlements in Greece. The analysis provides compelling support for the *uprootedness* hypothesis. Though initially lagging, refugees settling in the Greek countryside eventually outperformed nearby natives in educational attainment. Their university choices also diverged, with refugees' lineages favoring degrees transferable beyond the Greek labor market, such as engineering and medicine, and natives specializing in law and other fields with a strong home bias. Exploring additional mechanisms reveals the critical role of linguistic barriers and local economic conditions in shaping these outcomes, rather than the refugees' pre-migration economic background. The widespread educational gains of refugees and their descendants over four generations offer some hope that the ongoing surge of forced displacement, despite its tragedy, if properly addressed by the international community, can be a backbone of economic resilience for the affected communities.

Keywords: Forced Displacement, Human Capital, Transferable Skills, Uprootedness

JEL Codes: J24, N34, N44, O15

*We are deeply thankful to Athanasios Stavrakoudis and the team of young scholars at the Applied Informatics and Computational Economics Lab, University of Ioannina for data processing, cleaning, and digitizing. We thank Ekaterina Zhuravskaya, Hans-Joachim Voth, David Yanagizawa-Drott, Christos Genakos, and Sascha Becker for their insightful comments. We received constructive feedback from seminar and conference participants at Bocconi, Cattolica-Milan, the Athens University of Economic and Business, CRETE 2023 and 2024, the National Kappodistrian University of Athens, Brown, Università della Svizzera italiana, Zurich, UCLouvain, KCL, World Bank, Georgetown, and LBS. Special thanks to Nikos Benos (University of Ioannina) and Stelios Karagiannis (EU Commission), with whom we started this research. We also wish to thank the Center for Asia Minor Studies for granting us access to the archive of oral histories, and the Hellenic Statistical Authority (ELSTAT), especially Efi Chantsouli, Athanasios Thanopoulos, and Apostolos Kasapis. The National Science Foundation (NSF Grant 1729908) and the LBS Wheeler Institute for Development provided generous financial support. Additional Web Material (not for publication) can be found at http://www.anatolia-imprints.org. We thank the Hellenic Bank Association for supporting the portal's development. All errors are our responsibility.

†Brown University, Department of Economics, 64 Waterman St., Robinson Hall, Providence RI, 02912, United States. Email. smichalo@brown.edu, Web: https://sites.google.com/site/steliosecon/

‡University of Trento, DEM, Italy. Email. elie.murard@unitn.it, Web: https://sites.google.com/site/eliemurardinfo

§London Business School, Regent's Park, London, NW 1 4SA, United Kingdom. Email: eliasp@london.edu, Web: https://sites.google.com/site/papaioannouelias/home

¶King's College London, Strand, London, WC2R 2LS, United Kingdom. Email: seyhun.sakalli@kcl.ac.uk, Web: https://sites.google.com/site/sosakalli/

1 Introduction

The farming villages of Kato Komi and Roditis in Kozani, Macedonia, Greece, have many similarities. Just 10 kilometers apart, they share the same agricultural landscape. In 1928, both had 285 residents, while by 2021, their populations had declined slightly to 237 and 224, respectively. Yet, their educational paths starkly diverged. Before 1920, Kato Komi residents were five percentage points (pps) ahead in primary school completion. This educational gap had reversed by the end of the century, as high school graduation in Roditis was 10 pps higher. Moreover, university graduates from Roditis have pursued degrees with skills that are portable beyond Greece, like science and medicine, while those from Kato Komi have specialized in fields that are less transferable (but equally valuable). The key difference? Unlike the natives of Kato Komi, Roditis' inhabitants in 1928 were refugees from Pontus—part of the 1.2 million Greek Orthodox uprooted from Anatolia in the 1920s, settling in thousands of villages (and cities) across Greece. This paper examines how forced displacement shaped the educational investments of refugees and their descendants by looking at this defining episode for modern Greece, *the Asia Minor (or Great) Catastrophe*.

The "uprootedness" hypothesis suggests that the experience of forced displacement compels individuals and their descendants to invest in portable and adaptable capital, such as education and transferable skills, to navigate uncertainty and change. Stigler and Becker (1977) summarize the idea: "*Since Jews have been persecuted so often and forced to abandon their ancestral homes, they have not invested in immobile land, but in mobile human capital—business skills, education, etc.—that would automatically go with them.*"[1] Besides, education may be particularly desirable for the displaced and their offspring because it facilitates integration into the new communities. In her best-selling book, *We Are Displaced*, Malala Yousafzai alludes to this mechanism, writing that *"education is the key that unlocks the doors of opportunity for displaced individuals, empowering them to build a better future for themselves and their families."*

Despite its widespread popularity, there is a modicum of empirical work linking forced displacement to human capital and the few studies (discussed below) reach inconclusive results; for example, innovative research by Becker et al. (2020) looking at Polish relocations after WWII documents strong uprootedness effects, while studies on the forced movements of Germanic communities after WWII do not detect economically meaningful links between their experience and human capital (Bauer, Braun and Kvasnicka, 2013, Ciccone and Nimczik, 2023). It is even challenging to validate the uprootedness hypothesis in the Jewish history context, as "*the transition from farming to urban skilled occupations occurred in Mesopotamia and Persia of the eighth and ninth centuries, where the Jewish minority enjoyed significant security of life*" (Botticini and Eckstein, 2012). The challenge of empirically establishing this phenomenon may stem from relatively small samples, data reflecting a particular point in time, and numerous identification challenges related to selection, omitted variables, and heterogeneity

[1] Likewise, Botticini and Eckstein (2012) write: "*Scholars such as Werner Sombart, Reuven Brenner, Nicholas Kiefer, and Yuri Slezkine have maintained that the Jews, like members of other persecuted religious or ethnic minorities, preferred to invest in education rather than land because human capital is portable and cannot therefore be expropriated.*"

of origin and destination characteristics.

In this study, we examine the uprootedness hypothesis by analyzing the short-, medium- and long-term educational investments of Greek Orthodox displaced to Greece from Anatolia (Asia Minor, Pontus, Cappadocia, East Thrace, Caucasus and Bulgaria). These individuals were compelled to leave their ancestral homelands after the Great War (1914–18) and the subsequent Greco-Turkish War (1919–22), resulting in the loss of territories that had been home to Greek communities for millennia. Amidst chaos, refugees fled abruptly, arriving in Greece impoverished and destitute. They joined a nation struggling with internal divisions after a humiliating military defeat, increasing the country's population by approximately 20% in a few months. The *Convention Concerning the Exchange of Greek and Turkish Population*, signed shortly after the end of the fighting in January 1923, mandated (for the first time in the history of international relations) the compulsory relocation of Greek-Orthodox and Muslim populations between Greece and the nascent Turkish state. Using individual-level census data that span a century, we examine the dynamics of human capital over four generations of refugees whose families settled in the Greek countryside, mainly in Macedonia and West Thrace, where the departing Muslims lived.

The refugees' eagerness to acquire education was salient enough to be noted by the key figures involved in the resettlement process. Henry Morgenthau, the chairman of the Refugee Settlement Commission (RSC), the international body that undertook the resettlement, provides a first-hand account. *"They [refugees in Pella, Macedonia] were so poor that they were unable to build even a church, but for five years they had been holding their church services in a barn. The priest himself then explained what they wanted. It was not relief from taxes, nor an extension of time on their land payments, nor any of the selfish advantages one might have expected; it was not even a church they wanted. Said the priest: We are willing to go on worshiping in a stable until better times come, but we implore you to help us build a school, so that our children shall not grow up in ignorance"* (Morgenthau, 1929). His deputy, Sir John Hope Simpson, would similarly note: *"A very marked feature of the life of an agricultural (refugee) village is the intense desire for education. But even before they are comfortably settled, they commence to agitate for a school... The desire for education among the children is so great as to be incredible and is most astounding to an English observer... When one of the members of the Commission makes an inspection in a village where no school has yet been provided, it is the invariable rule that he is waylaid by the children, demanding its provision, and where there is a school in existence, you will find that the children are in attendance long before the hour fixed for the commencement of instruction, and that they leave it with the greatest reluctance"* (Simpson, 1929).

1.1 Results Preview

The starting point of our research is the compilation of a complete granular mapping of refugee settlements in the Greek countryside, distinguishing by their villages of origin in Anatolia (Section 3). We do so by processing, cleaning, and digitizing the 1928 General Population Census and the refugee

catalogs compiled by the Refugee Settlement Commission in the mid-1920s. As the Greek state did not inquire again about refugee background, we trace human capital across all rural settlements (hamlets, small and large villages, and towns), distinguishing between predominantly refugee and autochthonous ones in 1928, geo-referencing all censuses of the past century. We provide a comprehensive dataset of Greece's economic geography, which not only allows contrasting refugees' and natives' education at a granular level but is also of independent interest to the millions of Greeks of refugee origin.[2]

Our empirical analysis proceeds in three steps. First, in Section 4, using a difference-in-differences strategy, we compare education in predominantly refugee rural settlements and native ones in the same province (admin-3 level unit) before and after the compulsory population exchange (1923). At the time, primary school completion rates were *lower* in refugee settlements than in autochthonous ones. Likewise, the aggregate statistics of the 1928 General Population Census reveal that refugees in the countryside had a five percentage points *literacy deficit*. Eight decades later, junior high school completion rates were ten percentage points *higher* in refugee settlements than in native ones. The refugee-native difference in education, which applies to both men and women, is also present when we compare individuals from neighboring settlements (within a 25 km and even 10 km radius). Moreover, even among people residing in the same municipality (admin-4 level unit), those born in refugee settlements have about half a year of schooling more than Greeks hailing from native settlements. The refugee-native gap in education is similar among internal migrants (who moved away from their birthplace) and non-movers, reflecting refugees' deeper valuation of human capital, not directly related to employment opportunities or integration in urban areas. Consistent with a shift in educational preferences transmitted intergenerationally, the refugee-native gap in schooling persists and remains stable until the fourth generation of refugees, the last we observe. The displacement-schooling nexus also reveals interesting dynamics: First-generation refugees, born just before the displacement, and second-generation refugees, born in Greece in the 1920s and 1930s, catch up and eventually surpass natives in primary school completion; as primary school completion, made compulsory in the early 1930s, reaches 100% for all Greeks, the refugee-native educational difference for those born in the 1940s-1960s emerges in (junior) high school completion and, over time, in university education.

Second, in Section 5, we test the portability dimension of the "uprootedness hypothesis" using detailed census data on the major of undergraduate studies among those with tertiary degrees in 2011. Descendants of refugees pursue university degrees in fields that systematically differ from natives' offspring. College-educated Greeks from refugee villages and towns are more likely to obtain degrees with portable skills (e.g., science, engineering) than Greeks from native settlements. Even when we narrow our analysis to those completing law, medicine, or engineering in the most selective schools, third- and fourth-generation refugees are more likely to study engineering and medicine than law, which is less transferable across borders due to the legal system's idiosyncrasies and tight regulation.

[2]Today, roughly one-third of Greeks report being descendants of refugees from the maternal or paternal line and an additional 10–12% from both parents. Surveys suggest that most Greeks today consider the Asia Minor or Great (*Megali*) Catastrophe and the influx of 1.2–1.5 million Greek Orthodox from Anatolia one of the most traumatic events of the modern Greek state. As part of this project, we created www.Anatolia-Imprints.gr, where visitors can search for their ancestry and visualize the displacement trajectory by origin and destination.

The uncovered tendency of refugee lineages to invest in portable skills brings new evidence of the lasting impact of forced relocation on the preference for transferable skills among descendants of uprooted populations (Stigler and Becker, 1977, Kessel, 1958).

Third, in Section 6, we use the rich variation in origin, displacement trajectories, and refugee settlement characteristics to illuminate additional mechanisms. Increases in human capital are muted when looking at the non-Greek-speaking refugees from Cappadocia (about 10% of the inflow), highlighting the importance of linguistic barriers in integrating refugees. Apart from this, the educational advantage is widespread across dozens of sample splits. It applies to refugees and their descendants settling in newly established villages and those abandoned by Muslims; in places with low and high agricultural potential; to those from exclusively rural or more urban origins. The magnitudes are somewhat more pronounced for Greeks growing up in settlements with refugees from more diverse origins, hinting at the positive impacts of mixing. While the link between displacement and education applies to refugees from (almost) all backgrounds settling in all types of destinations, it is stronger for those settling in more densely populated, manufacturing, and commerce-oriented provinces, where the demand for an educated workforce was presumably higher. Although the refugee advantage in schooling years shifts in response to labor market opportunities, this is not the case for the type of university degree: irrespective of background and destination characteristics, refugees' offspring prioritize studies with portable skills, including regions where engineering degrees were not in demand.

1.2 Related Literature

Our work belongs to the broad research agenda that explores the economic effects of displacement episodes for the affected populations and the receiving communities; Becker (2022) and Becker and Ferrara (2019) provide detailed surveys. Of most relevance are studies examining the lasting effect of forced migration at the end of WWII.[3] Becker et al. (2020) explore the impact of the redrawing of Poland's borders and the movement of millions of Poles from the East (*Kresy*) to the West (former Germanic areas) on human capital. Surveys from 2015 and 2016 show that descendants of the displaced have, on average, one extra year of schooling and value education more than non-displaced. However, research on forced migration in Germany from the East after WWII does not find much of an effect on human capital. Bauer, Braun and Kvasnicka (2013) document a tiny increase in the education of German migrants from the East (Silesia, Prussia, Danzig, and Pomerania) compared to non-migrants in West Germany in 1971 despite a move towards manufacturing and services (Bauer and Kvasnicka, 2014). Peters (2022) and Ciccone and Nimczik (2023) do not find any education impact of German refugees despite strong agglomeration effects (see also Grosser (2006) and Schumann (2014)). Sarvimäki, Uusitalo and Jäntti (2022) explore the income effect of the relocation of about 11% of the Finnish population after the border realignment with the Soviet Union during WWII. They

[3]See Chiovelli et al. (2023) for (internal) displacement during civil conflict and Nakamura, Sigurdsson and Steinsson (2022) in the context of natural disasters. For the broader literature on the economic effects of immigration, see Card (2001), Sequeira, Nunn and Qian (2020), Abramitzky and Boustan (2017), and Dustmann, Schönberg and Stuhler (2017), among others.

find some positive albeit heterogeneous effects on education (i.e, non-existent for refugees from cities). Besides, their results do not directly speak to the uprootedness hypothesis as Finnish refugees' higher education is a consequence of their transition from agriculture to the modern sector and the resulting higher returns.

We advance the literature on the legacy of displacement in various domains. First, we trace the educational progress of Greeks from Anatolia—who became refugees in 1922–23—and autochthonous, covering a century leading up to the displacement and almost a century afterward. Second, using census data, we study displacement's short-, medium-, and long-term impacts on human capital. Third, we study additional mechanisms, thanks to the detailed mapping of refugees' origins and settlements, which carry lessons for policymaking: the importance of linguistic barriers in integrating refugees and that of destination rather than origin features. Fourth, we provide new evidence on the portability aspect of the uprootedness hypothesis by showing that refugees' descendants are considerably more likely to acquire *transferable* college-level skills (Stigler and Becker, 1977, Ayal and Chiswick, 1983, Brenner and Kiefer, 1981).

Our work also contributes to the extensive body of historical research assessing the role of the refugees from Anatolia in Greece, which has received renewed interest given the recent centenary celebrations and, more generally, to population movements in the Balkans and Eastern Europe during the Great War (e.g., Pentzopoulos (1962), Kontogiorgi (2006), Kritikos (2005), Mavrogordatos (1983), Mazower (1991), Hirschon (2003), Pavlowitch (1999)).[4] Our contribution is three-fold. First, we provide a plethora of new georeferenced data that allow us to quantify the behavior of refugee lineages over time, complementing the narratives. For example, the newly compiled data are inconsistent with a widespread view in Greece that refugees from Anatolia were more educated than the autochthonous upon arrival. Second, we move beyond case studies, identifying the impact across all rural settlements. Third, by uncovering the mechanisms, we draw more general lessons and connect this defining episode for Greece to other forced population movements.

2 Historical Background

2.1 Greco–Turkish Wars

Greece sprang from the Ottoman Empire after the War of Independence (1821–29). The new Kingdom, established in 1832, included only the Peloponnese and parts of *Roumeli* (Sterea Ellada). Even after the annexation of the Ionian Islands in 1864 (ceded by Great Britain) and Thessaly in 1881, most Greek communities resided outside the country's territory. Uniting Greek populations scattered mainly across the Ottoman Empire became the dominant ideology. The irredentist aspiration, known as the "*Megali Idea*" (Grand Idea), would define the nascent Greek state in its first century. The opportunity came in

[4] Arbatli and Gokmen (2023) investigate how the removal of Greek-Orthodox and Armenian communities shaped the development of Turkish districts. Hornung (2014) documents significantly positive productivity effects of the high-skilled Hugenot migration in Prussian textile manufacturing.

1912 during the Balkan Wars, in which Greece, fighting alongside Serbia, Montenegro, and Bulgaria, defeated the Ottoman Empire—and then Bulgaria. By the end of 1913, Greece's size and population almost doubled as it annexed Epirus, Macedonia, Crete, and the Aegean Sea islands (Chios, Lesvos, Samos, Icaria, Limnos). After tensions with the royalty ended in a civil war, Eleftherios Venizelos, the Greek prime minister, brought Greece into WWI on the *Entente* side.

In the Paris Peace Conference, Greece, on the winning side of WWI, secured West Thrace from Bulgaria with the *Neuilly Treaty* (1919), which also included provisions for the voluntary reciprocal emigration of Greeks and Bulgarians. With the *Treaty of Sèvres*, which dissolved the Ottoman Empire (August 1920), stripping it of its European territories, Syria, Lebanon, and Palestine, *Megali Idea* reached its apogee. Greece secured East Thrace, extending its borders a few kilometers from Constantinople (Istanbul). Greece also got the administration of Ionia, an area with significant Greek populations around Smyrna (Izmir), with a referendum on its status to be held ten years later. The Greek Army triumphantly landed in Smyrna in May 1919 to protect the Greek population, which faced attacks and deportations to labor camps in the interior. Under the leadership of Mustafa Kemal, the Turkish forces, which did not recognize the Treaty of Sevres, staged fierce opposition. After Greek troops failed in the interior of Asia Minor, the Turkish army staged a swift counter-attack, capturing Smyrna in September 1922.

The persecution of the Greek-Orthodox communities by the advancing Turkish Army and militias, the arson of the Christian, Greek and Armenian, quarters of Smyrna, and ethnic cleansing triggered a mass exodus. Hundreds of thousands fled, leaving their homes, farms, livestock, and properties behind. As the Turkish troops were about to cross into East Thrace, the Greek side signed the Armistice of Mudanya in October 1922. Hostilities ended, but Greece conceded East Thrace, which triggered the exodus of hundreds of thousands of Greek-Orthodox from East to West Thrace.

2.2 The Population Exchange

Following the consolidation of power by Kemal's government in Ankara and the temporary return of Eleftherios Venizelos to Greek politics, the two leaders negotiated the Peace Accord under the League of Nations. The *Convention Concerning the Exchange of Greek and Turkish Populations* of 30 January 1923 provided the *"compulsory exchange of Turkish nationals of the Greek Orthodox religion established in Turkish territory, and of Greek nationals of the Moslem religion established in Greek territory"* (article 1). Exemptions were made for the Greeks of (the city) Constantinople (but not the broader area, *vilayet*) and the two islands of Imvros and Tenedos and the Muslims of West Thrace (article 2). To accommodate for the population movements during the Balkan Wars (1912–3), World War I (1914–8), and the Greco–Turkish War (1919–22), the treaty's definition of *"emigrant applied to all persons who have been obliged to emigrate since 18th October 1912"* (article 3). Emigrants *"acquire the nationality of the country of their destination, upon their arrival"* or at the convention date for those already displaced (article 7).

The Treaty of Lausanne (July 1923) ratified the population exchange. It settled all issues of the

dissolving Ottoman Empire, now represented by the newly established Turkish State. Many refugees (of both creeds) did not identify themselves as Greek (or Turkish) nationals, as their ancestors had lived in Ionia, Pontus, Cappadocia, and East Thrace (Macedonia and Crete) for centuries. Some did not even speak Greek, for example, the communities in Cappadocia and Cessaria, or spoke dialects, which were hardly intelligible in Greece. The exchange was monitored and facilitated by the League of Nations Commissioner for Refugees, Fridtjof Nansen.[5] Despite its compulsory nature and the amnesty for war crimes, the convention would serve as a blueprint for the resolution of disputes exposing minorities to human rights violation, persecution, or genocide.[6]

According to the 1928 General Population Census, in a population of 6,204,684 there were 1,221,849 refugees (19.6%); about 151,892 arrived before 1922–3 and 1,069,957 after 1922–3.[7] About 50.2% settled in big cities (with a population exceeding 10,000), and 49.8% in the countryside.

2.3 Refugee Settlement in Greece

2.3.1 The Refugee Settlement Commission

Accommodating such a large-scale population influx was far beyond the capacity of the Greek State, which was almost bankrupt after a decade of warfare, with social and political divisions. Initial support came from primarily American charitable foundations, such as the American Red Cross and the Near East Relief Fund. The League of Nations stepped in founding the Refugee Settlement Commission (RSC) in September 1923. The RSC, fully independent, had complete authority over refugee settlements and the distribution of funds.[8] It consisted of an American Chairman representing the charitable organizations, another foreign member, and two members nominated by the Greek government. The RSC, which operated initially with a bridge loan from the Bank of England, would raise 10 million pounds sterling in December 1924 (net of charges and commissions) with a 7% interest and a second loan of 9 million pounds sterling in 1928 (with a 6% interest), a third of which was devoted to the completion of the refugee settlement.

The RSC operated in two broad domains: settling refugees in the countryside and supporting urban refugees who settled in shacks, slums, and the periphery of Athens, Piraeus, and Thessaloniki. Absolute priority was given to rural (agricultural) refugees since the RSC and the Greek government deemed the immediate increase in agricultural produce fundamental, while at the same time limiting

[5]Nansen, whose work on the Great War's displaced was recognized with the 1922 Nobel Peace Prize, led the negotiations with the Turkish side for the release of 100,000 Greek-Orthodox men captured during the hostilities and placed in infamous military battalions. See Kitromilidis (2008), Mavrogordatos (2017), and Clark (2006) for the population exchange and Psomiades (2009) for an overview of the Nansen (1923) Report.

[6]For example, Katz (1992) discusses the efforts of the British Royal Committee and the Zionist Committee to emulate the Greco-Turkish Population Exchange in Palestine in the mid-1930s. Mazower (2009) connects the Greco-Turkish Population Exchange to the discussions of population movement during the dissolution of Yugoslavia in the 1990s.

[7]The actual number was higher, given the high mortality refugees faced during the first years. The census misses refugees who emigrated to Europe and the Americas before 1928. Kitromilidis (2008) approximates that about 75,000 died between 1922 and 1928 and 66,000 emigrated. The refugee count includes children under five years.

[8]For RSC's work, see Kritikos (2000), Hill (1931), the detailed quarterly reports from the League's Archives, and RSC leadership eloquent memoirs (Morgenthau, 1929, Howland, 1926, Ladas, 1932, Eddy, 1931, Simpson, 1929).

the presence of refugees in the cities, and preventing their radicalization and swing to communism. About 9 millions of the initial loan were allocated to the *"colonisation"* of the countryside, and just one million for the refugees in cities, although roughly half settled in urban centers.[9]

2.3.2 Agricultural Refugees Settlement

The agricultural settlement started with the RSC representatives assembling refugees into groups and then moving them to their new settlements (Pentzopoulos, 1962). The RSC attempted to keep communities together. However, the haphazard exodus dispersed the original communities across many settlements (see Section 6 below). The typical farm allocated to a refugee family was 3.5 hectares (Simpson, 1929), reflecting the average landholding size at the time and today (as three-fourths of Greek farms are less than five hectares (European Union, 2022)).

Housing and Land Regarding housing, basic facilities, and land, the properties of roughly 400,000 Muslims offered immediate relief. But, as this was not enough, the RSC built new villages and expanded small ones.[10] RSC planners determined the new settlement locations mainly by assessing the availability of suitable land (Eddy, 1931). The resettlement process *"left little if any scope for choice, on the part of the refugees, as regards the place of their final resettlement"* (Kontogiorgi, 2006, p. 143). The RSC concentrated its operations in Northern Greece, where large tracts of arable land were left vacant by the expelled Muslims, who predominantly worked in agriculture. Drawing on earlier legislation (1915–7), which, however, had not been enforced, the authorities split the large estates of departing Muslim landlords (*chifliks*) into smallholdings and distributed them to refugees. Palis (1929) reports that 60% and 11% of land transferred to the RSC by the Greek state, in total $850,000$ hectares, came from expellees to Turkey and Bulgaria, respectively, while about 10% was acquired by expropriating Greek citizens. The RSC also confiscated some unused land of foreigners, and expanded the cultivable land area by undertaking drainage and reclamation works, mainly in Macedonia.

Other RSC assistance to agricultural refugees went well beyond providing shelter and land. First, the RSC provided technical assistance, nudging farmers into cereals (wheat, corn, maize), deemed chief for surviving, and tobacco due to the high yields even for small lots. Second, it gave much-needed livestock, plows, seeds, tools, and machinery (Kontogiorgi, 2006). Palis (1929) reports that

[9] Kritikos (2005) reports that by the end of 1928, policymakers had spent $9,117,362£$ for agricultural and $1,302,734£$ for urban settlement. In the words of Sir John Hope Simpson (1929), the RSC's Vice-Chair: *"From the start, the Commission considered the settlement of the agricultural population on the land primary importance, in order, first, that agricultural production might be increased as rapidly as possible so as to provide for needs of the new population, and secondly, to avoid the congregation in the towns of large masses of the population."* Likewise, Kritikos (2005) quotes the Greek Minister of Agriculture in 1924 saying: *"The rural establishment of the refugees had exclusively occupied our main attention, for the agricultural settlement of these people is easier than the rehabilitation of the urban masses."*

[10] Palis (1929) reports that by the end of 1928, there were $122,402$ agricultural refugee houses and apartments in $1,954$ rural refugee settlements; $62,770$ (51.3%) were houses of departing Muslims, $55,264$ were new constructions, $4,356$ were in former state buildings.

the RSC had provided by the end of 1928 about 60,000 plows alongside 150,000 animals (e.g., cows, horses, mules, sheep), with the state giving an extra 135,000 animals. Third, the RSC helped with infrastructure, built some minor roads, cleared trails, and *"erected and equipped fifty-nine hospitals and dispensaries, and maintains for each of them at least one medical man and one dispenser"* (Simpson, 1929). The RSC did not build schools, as education was outside its duties.

Assessment At its dissolution at the end of 1930, the RSC had completed its main tasks, providing immediate relief, refuge, land, seeds, and production inputs for more than half a million rural refugees. Tensions and unrest were limited. Despite the small plots, production boomed as cultivated areas in Macedonia and West Thrace more than doubled, boosting cereal, corn, tobacco, and wine production and bringing much-needed foreign income from exports.

3 Data and Preliminaries

This section first gives an overview of the primary data and then discusses some preliminary patterns. Appendix Section A gives details, additional visualizations and results, and summary statistics.

3.1 Main Data

In 1928, Greece was divided into 35 prefectures (admin-2 units, *nomos*) and 141 provinces (admin-3 units, *eparchies*), excluding the autonomous region of Mount Athos. Provinces are local economic units with a similar history, dating back to the Byzantine era and the administrative structure of the Greek-Orthodox church. The mean (median) population of provinces in 1928 was 43,970 (30,992) while the 90th and 10th percentiles were 130,201 and 3,855, respectively. Since the islands of the Dodecanese joined Greece in 1948, we exclude them from the analysis.

3.1.1 The 1928 Population Census

The starting point is the processing, cleaning, and digitization of the 1928 General Population Census (conducted on May 15–16), the only census that inquired about refugee status. While the Greek Statistical Agency (ELSTAT) has not kept the micro-data, we managed to access the information across the universe of inhabited places: 11,003 settlements (cities, towns, villages, and even hamlets). Figure 1, Panel *A*, plots the population share of refugees in 1928 across 5,605 administrative communities (we use the latter polygons to visualize the variation).[11] Most refugees settled in Thessaloniki, in rural Macedonia (about 540,000), as most Muslims departed from there, and in West Thrace. 45% of Macedonia's population and 35% of West Thrace's in 1928 consisted of refugees. Athens, Piraeus, and the larger area of Sterea Ellada (Central Greece) received nearly 300,000, 20% of its population.

[11] ELSTAT makes publicly available the refugee counts at the *settlement level* only for those arriving after the Asia Minor Catastrophe (1922–3), roughly 87.5% of the total. Appendix Figure A2 gives the share of all refugees arriving pre and post-1922 across provinces. The cross-province correlation between the two exceeds 0.90; see Appendix Figure A3.

About 85,000 refugees settled in Crete (9% of its population) and the Aegean Sea Islands (18.4%). Few refugees settled in Southern and Western Greece (Peloponnese, Epirus, and the Ionian Sea Islands); their share was below 3%. As shown in Panels B and C that zoom into Macedonia and West Thrace and the area surrounding Giannitsa province in central Macedonia, there is a considerable spatial variation within admin-1 level (region – "*peripheries*"), admin-2 level (prefecture – "*nomoi*"), and admin-3 level (provinces – "*eparchies*") units. Admin-1 level constants explain 24% of the variability in the refugee share; admin-2 level fixed-effects 46%, and admin-3 level fixed-effects 52%.[12]

We also extract the number of departing Muslims between 1920 and 1928 across provinces, in total 415,942 (Appendix Figure A4). Since this information is unavailable at a more disaggregated level, we combine the province-level numbers with changes in settlements' population and an ethnographic map of Macedonia in 1915 (Güvenç, 2010) to proxy the share of departing Muslims in the latter. Appendix Figure A5 gives the mapping of Muslims across settlements. Using this information, we differentiate refugee settlements between former Muslim and newly established.

3.1.2 Rural and Urban Refugee Catalogs

We processed, cleaned, and digitized 100% of the rural refugee catalog compiled by the RSC, as it allows us to detect the land grants to refugees and trace both origin and destination at a granular level; see Appendix Figure A9 for examples of the catalog's origin-destination structure. The rural catalog lists about 245,000 entries. After deleting duplicates and cleaning up, we ended up with 211,056 unique household heads, along with information on the village of origin and the settlement of the land grant. The grants span 1,427 settlements, mapped in Figure 1.

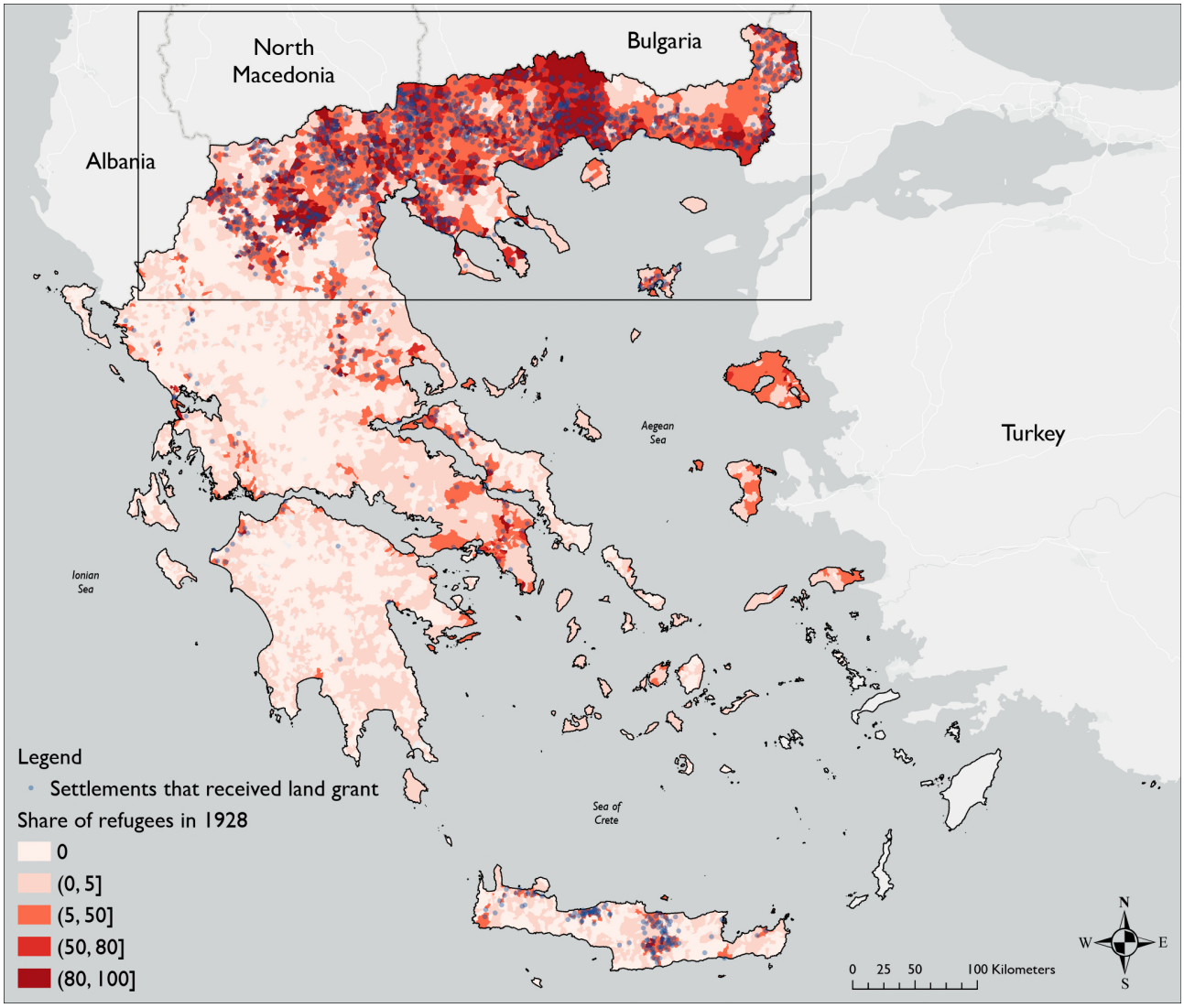
The urban catalog has about 380,000 hand-written entries of household heads' names alongside the origin settlement. However, information on the city (and neighborhood) of residence are missing. These families resided in the mid-1920s in Thessaloniki, Piraeus, Athens, and 70 other big cities. We transcribed 164,617 entries (alphabetically), approximately 43%.

Figure 2 maps for the first time all 1,850 Greek-Orthodox settlements in Anatolia, tabulating the origin of all household heads in the rural and about half of the urban refugee catalogs. The refugees came from places with diverse cultural, professional, and linguistic backgrounds. Moreover, their experience during the preceding decade-long fighting, displacement trajectories, and settlements in Greece differed considerably. In Section 6, we give details, drawing on the digitization and decoding of 1,500 interviews with first-generation refugees from the *Center of Asia Minor Studies* archive. The bilateral nature of the refugee catalog data yields a battery of heterogeneity checks, which shed light on additional mechanisms.

[12]The 1928 Census records refugees' origin across provinces, distinguishing coarsely across regions of provenance. Refugees came from Asia Minor (626,954), Pontus (182,169), East Thrace (256,635), Bulgaria (49,027), Caucasus (47,091), Russia (11,435), and about 11,000 from Serbia, Albania, Romania, Cyprus, Egypt, and the Dodecanese.

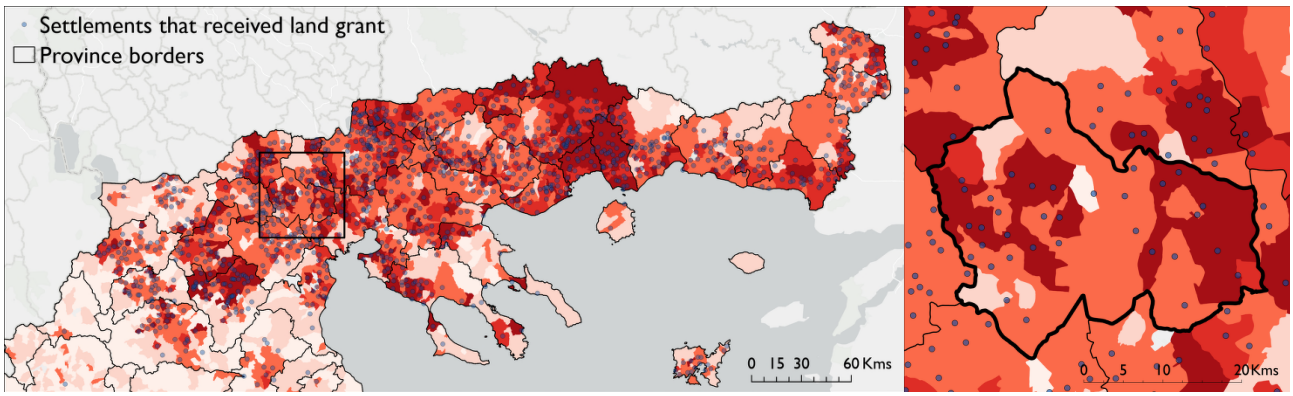
Figure 1: Refugees in 1928 and Land Grants across Settlements

(a) Greece



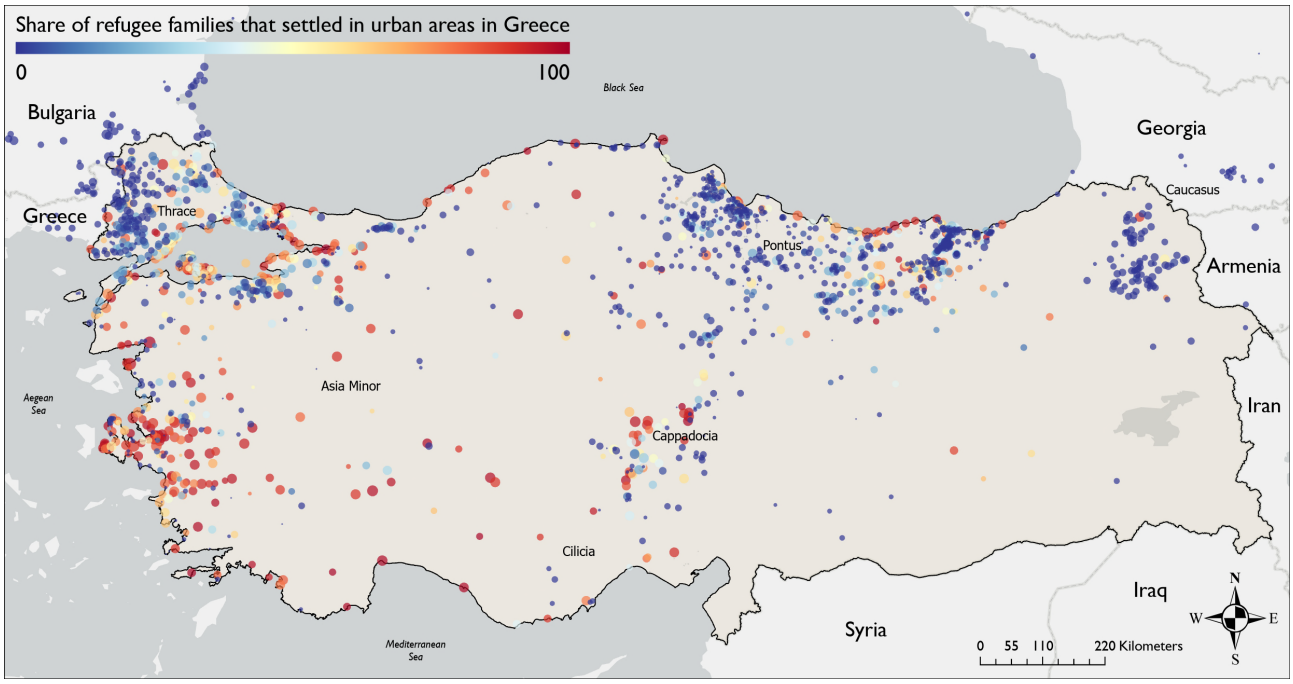
(b) Macedonia and West Trace

(c) Giannitsa province



Notes. The figures plot the spatial distribution of refugees arriving after the Asia Minor Catastrophe (1922–23) across 5,605 communities in the 1928 General Population Census. Community (*koinotita*) is an administrative unit above the settlement level, which we use in our analysis. Panel *A* plots the share of all refugees (rural and urban) across Greece. Panel *B* zooms into Macedonia and West Thrace. Panel *C* zooms into Giannitsa province in the prefecture of Pella, Central Macedonia. Darker (lighter) color indicates a higher (lower) share of refugees. The figures also indicate with blue dots 1,427 settlements with land grants given to agricultural refugees from the RSC's Rural Refugees Catalog.

Figure 2: Settlements of Greek-Orthodox Refugees in Turkey before the Population Exchange



Notes. The figure plots the origin of all Greek refugees in the rural and urban refugee catalogs compiled by the Refugee Settlement Committee in the mid-1920s. The information is based on 211,056 household entries in the rural catalog (1,850 settlements) and 164,617 household entries in the urban catalog (923 settlements). 1,735 origin settlements, out of the 1,850, are in contemporary Turkey. Blue (red) dots indicate the origin settlements of refugees in Greece's rural (urban) areas in the mid-1920s. The dot size presents the (log) number of refugee households from that location. To compute the share of refugees who settled in urban areas, we scale up the number of families in the urban catalog by dividing it by 0.43.

3.1.3 Greek Censuses

We digitized and georeferenced all (12) population censuses from 1907 to 2011 to map pre-population exchange conditions and trace the impact of the refugee influx. Our empirical analysis uses micro-samples from the 1971 (10%), 1981 (10%), 1991 (25%), 2001 (25%), and 2011 (25%) General Population Censuses, kindly shared by the Hellenic Statistical Authority (ELSTAT). We extract information on individuals' education and demographics, while from the 2011 Census, we also extract information on migration and birthplace.

3.1.4 Other

We also collected additional data. First, we georeferenced the rail network and stations in the mid-1920s. Second, since some land from monasteries was transferred to rural refugees, we identify 389 villages in the 1920 General Population Census whose name indicates the seat or property of a monastery, nunnery, or *metochion*. Third, we processed information on swamps and malaria in the 1920s, as some refugee settlements were near reclaimed/drained swamps (Kardamatis, 1924). Fourth, we compiled settlement-specific geographic and locational measures: soil suitability for agriculture (Zabel, Putzen-

lechner and Mauser, 2014), temperature and precipitation (Fick and Hijmans, 2017), altitude (Japan Aerospace Exploration Agency, 2015), and distance to the coast and the provincial capital.

3.2 Samples

Stable Geographical Units across Censuses To conduct the analysis at the settlement level, the most disaggregated unit, we had to create a crosswalk of censuses. In 1928, there were 11,003 settlements, whereas, in 2011, 13,537. Tracing settlements over a century was far from straightforward, as there have been myriad name changes, annexations, creations, abolition, and detachments. Using information from the official *Gazette of the Government* detailed in the Hellenic Society for Local Development and Self-Government website, we constructed a correspondence for each settlement in the 1928 Census over 100 years, resulting into 8,031 stable geographical units.

Baseline Sample After dropping 146 cities and district capitals in the 1928 census and 116 settlements within 25 km of Athens, Piraeus, and Thessaloniki, we are left with the baseline sample of 7,769 rural settlements.[13] The mean (median) share of the refugee population across the 7,769 rural settlements is 10% (0%); for the 3,445 settlements with at least one refugee, the mean (median) is 22.5% (3.2%); see Appendix Table A1. On the one hand, there are 698 settlements with a refugee majority population. Out of these, in 410, mainly in Northern Greece, the share of refugees (arriving after the summer of 1922) exceeds 80%. Among the latter, 405 received at least one land grant; we consider them as our "treated units," labeling them "predominantly refugee villages."[14] On the other hand, there are 6,228 villages where the refugees' share is below or equal to 5%. Among the latter, 6,091 had no land grant; we consider them as "native or autochthonous settlements," which will serve as the control units. Our baseline sample analysis leverages comparisons across these 6,496 settlements; Appendix Figure A12 gives the mapping.

Local Sample To further account for time-varying unobservable features, we restrict estimation to nearby predominantly refugee and autochthonous settlements, splitting provinces into Voronoi polygons so that each native settlement is matched to the nearest refugee settlement within a 25 km (or smaller) radius. Appendix Figure A13–Panel *C* illustrates our procedure zooming into the province of Giannitsa in Central Macedonia. The province is split into five polygons so that (at least) one predominantly refugee settlement is in the same sub-province unit as each native one. The local analysis sample compares 262 predominantly refugee settlements to 1,042 native settlements (Appendix Figure A13–Panel *A*). The mean (median) distance of native and refugee villages in the same sub-province (Voronoi polygons) is 11.5 (12) km, and the 90th-10th percentile range is 21.6 and

[13] Urban centers have transformed over the past century due to the dramatic rise in urbanization, making sharp comparisons between native and refugee areas difficult. Nevertheless, focusing on the urban sample delivers similar patterns (available upon request).

[14] There are 21 settlements with more than 80% refugees without land grants. However, as there is a nearby settlement with a land grant within 5 km of them, we treat them as predominantly refugee settlements. There are five settlements with more than 80% refugees where no land grant was received.

4.2 km. In comparison, in the baseline sample, the mean (median) distance between native and refugee villages in the same province is 18.3 (20) km; the 90th-10th percentiles are 37.7 and 6.7 km.

Settlement Population The comparisons of predominantly refugee settlements to native ones involve small and medium-sized rural villages and towns with an average of 450 inhabitants in 1928 (median 300, 90th-10th percentiles from 100 to 1,000); the median difference was just 24 individuals. The population has evolved similarly in the two sets of settlements. Despite Greece's population doubling, these settlements have remained rural. Their population increased until 1961 but gradually fell back to the 1928 levels by 2011 (Appendix Figure A15). This trend mimics the population dynamics of the Greek countryside as urbanization rates increased from roughly 30% in 1928 to 80% today. Hence, our results do not reflect agglomeration forces, often stressed in studies of (forced) migration, allowing us, therefore, to more cleanly identify uprootedness effects.

3.3 Refugee Settlements Features

We explore the correlates of refugee presence in the Greek countryside both to shed light on differences between native and refugee settlements that may affect our estimates (due to selection and omitted variables) and to provide quantification to the historical research. We do so by associating the share of the refugee population in 1928 with geographic, ecological, and locational features, as well as with "balance tests" that examine mean differences in these features between the two sets of settlements. The tabulations, given in Appendix Section A.2, reveal that, by far, the most significant correlate is the share of Muslims in 1920, as it explains more than half of the overall variation in refugee presence in 1928. A ten-percentage points increase in the share of the Muslim population is associated with a 7.0–8.2 percentage points increase in the share of refugees, even when exploiting within-province variation. In line with the historical narrative, the refugee share is higher in areas closer to rail stations and reclaimed marshlands. Refugee settlements are mainly in the lowlands; the mean altitude difference is 59 meters. However, these differences shrink and turn insignificant when comparing settlements in the same province or conditioning on altitude. When we zoom into proximate settlements, geographical, locational, and ecological differences are small; only the share of Muslims and altitude retain significance. Besides, refugee settlements had a smaller population before the exchange, as the RSC established many new villages.

4 Forced Displacement and Human Capital

This section reports the results linking the forced displacement shock to subsequent human capital investments, the core of the *uprootedness hypothesis*. We first present some simple before-after tabulations. Second, we present difference-in-difference specifications comparing the evolution of education in predominantly refugee and native settlements. Third, we present alternative specifications that account for residence and distinguish between (descendants of) refugees living in their birthplace and

those residing elsewhere. Fourth, we report some additional results and sensitivity checks, reported in Appendix Section B.

4.1 Preliminary Patterns

Before proceeding with the formal analysis, it is instructive to review the "raw" data patterns in human capital before and after the refugees' arrival in Greece.

Pre-Trends As the 1928 General Population Census reports age-specific literacy rates for natives and refugees by province, we can explore trends until the forced population exchange. In Figure 3, we trace the literacy dynamics between the two groups in the Greek countryside, going back to natives and refugees born shortly after Greek Independence in 1830. We also complement the tabulations of the 1928 Census, overlaying (with dashed lines) the literacy rates of those residing in rural predominantly refugee and native settlements using the 1971 and 1981 General Population Censuses. Although the samples and sources differ, the patterns are reassuringly similar. Two patterns are apparent. First, literacy increases steadily *both* for refugees, obtaining primary education in the Ottoman Empire (Asia Minor, East Thrace, Cappadocia, and Pontus), and the autochthonous population attending schools in the Peloponnese, Roumeli, and Thessaly. The upward literacy trend is similar for the two groups; various tests fail to reject the null hypothesis of no differential trend between natives and refugees across Greece or when we zoom into Macedonia and West Thrace, which were also part of the Ottoman Empire until the 1910s (Appendix Figure B2 and Appendix Table B1).

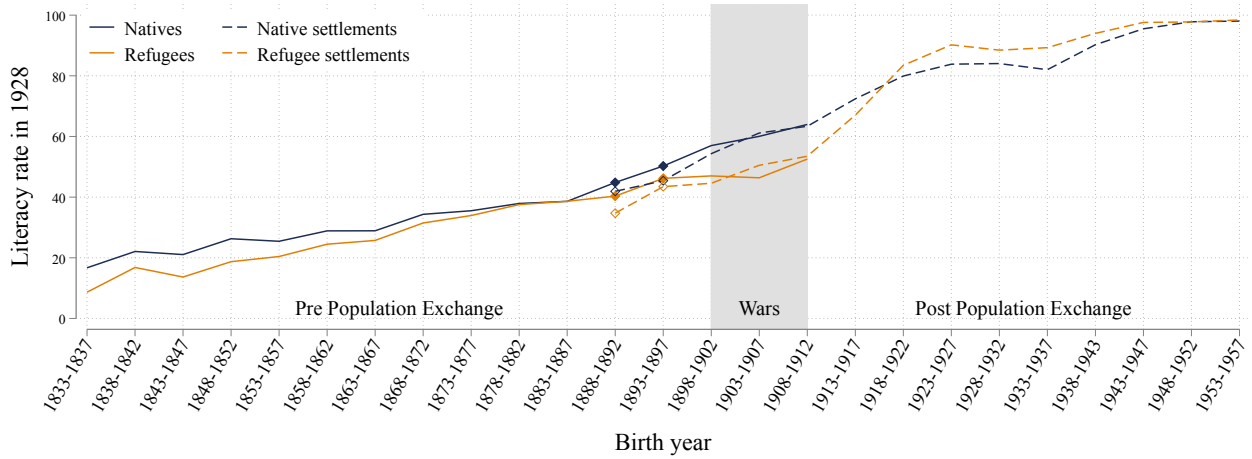
The similar trajectories in literacy during the 19th century reflect the significant autonomy the Ottoman Empire had granted the Greek-Orthodox (and other monotheistic non-Muslim) communities—grouped in the Rum *millet*—in administrating and funding their schools (Masters, 2009, Kazamias, 1991, Zervas, 2017, Hatzopoulos, 1991). Second, Figure 3 reveals (to the best of our knowledge for the first time) the adverse impact of the violence and repression that Greeks living in the Ottoman Empire faced in the late 1900s and 1910s from the Young Turks and the intensification of conflict. Literacy gains halt for those born in Anatolia at the turn of the century, whose primary school enrollment age coincides with the rise of ethnic tensions; see also Appendix Figure A14 for the mortality effects.

Conditions upon Arrival Turning now to the time of the refugees' settlement in Greece, Figure 4–Panel *A* shows that in the 1928 Census rural refugees' literacy rates (43%) were lower than natives in the countryside (49.3%).[15] The right panel tabulates the statistics for Macedonia and West Thrace, where 80% of rural refugees settled. The literacy rate in these regions, which became part of Greece from the Ottoman Empire shortly before the population exchange (1913 and 1920, respectively), was 5 pps lower than the rest of Greece. The literacy rate among refugees was 40.3%, three and a half percentage points lower than natives'.

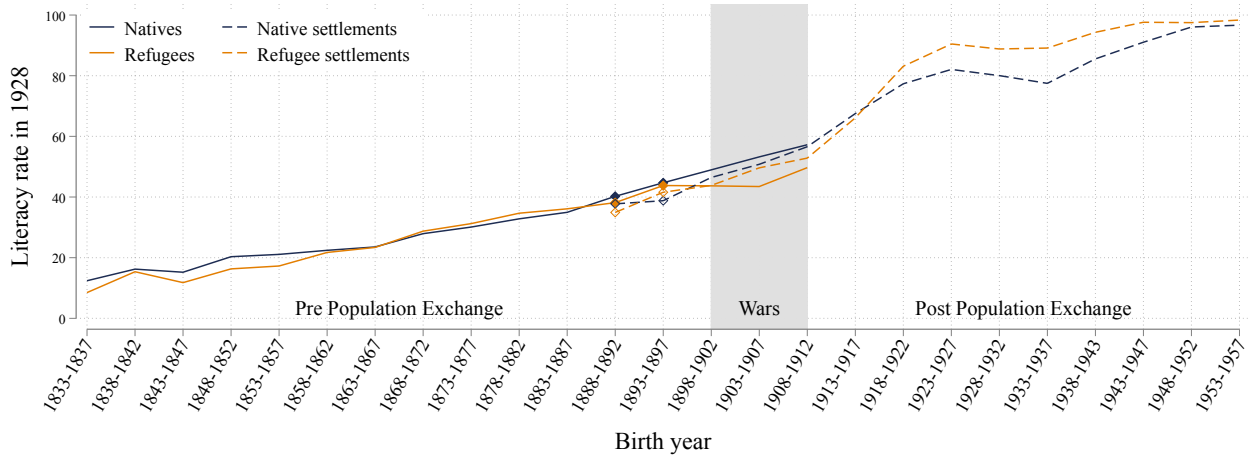
[15]Likewise, refugees in urban centers were less educated than natives, their literacy rate was 65.2% compared to natives' 71%. Overall, refugees' literacy was marginally lower than natives' (52.8% vs. 53.7%), as relatively more refugees than natives lived in cities in 1928; 50.2% of refugees and 28.1% of natives lived in urban areas.

Figure 3: Literacy across Rural Refugees and Natives by Birth Cohort

(a) All Greece



(b) Macedonia and West Thrace

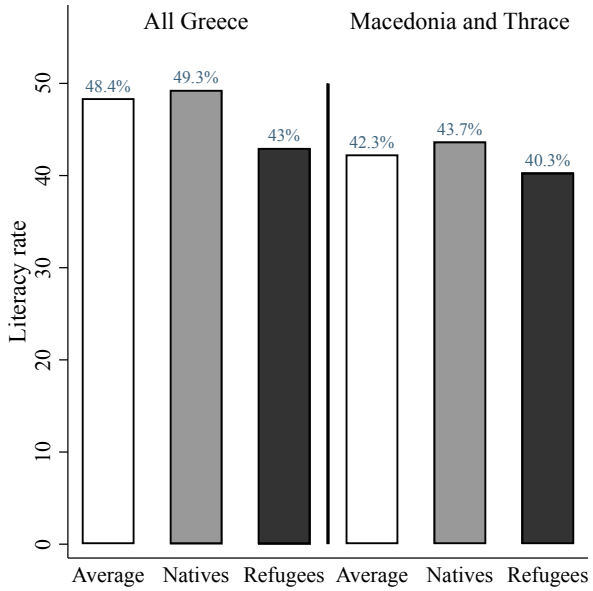


Notes. Panel *A* plots literacy rates for rural refugees and non-refugees across five-year birth-cohorts in the Greek countryside. Panel *B* plots literacy rates for the two groups in Macedonia and West Thrace, as 79.8% of rural refugees settled in Macedonia and Thrace. Literacy rates for cohorts born between 1833 and 1912 (in solid lines) are retrieved from the 1928 General Population Census. Literacy rates for cohorts born between 1888 and 1957 (in dashed lines) are from the 1971 and 1981 General Population Censuses. The cohorts born between 1888 and 1897, with diamond markers, serve as the excluded cohort in the difference-in-differences analysis. Gray indicates the cohorts affected by the Balkan Wars, the Great War, and the Greco-Turkish War.

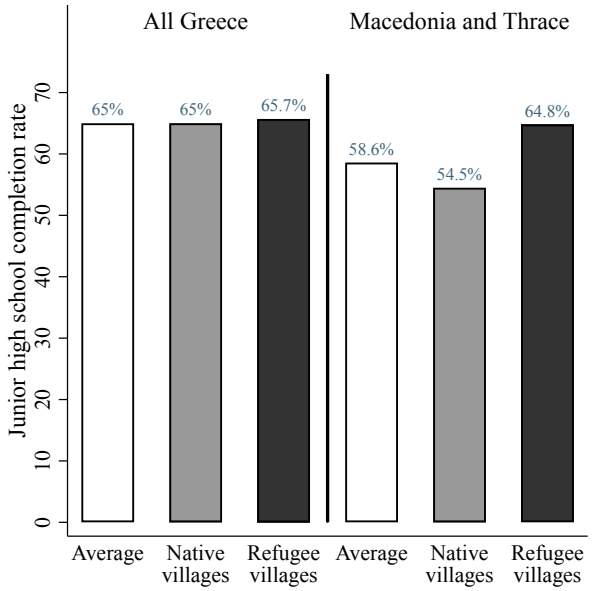
The First Cohorts After Displacement Figure 3 further shows that literacy dynamics changed for refugees who attended primary school in the 1920s upon settling in the Greek countryside. The first cohorts of refugees who attended school in Greece (born in 1918–22) swiftly caught up with natives, reaching a literacy rate of roughly 80%. In addition, those born in refugee settlements in Greece during 1923–7 had five percentage points higher literacy rates than those born in native ones. Literacy rates remained higher in refugee settlements for at least a decade (the late 1920s and 1930s).

Figure 4: Education in 1928 and 2011 across Rural Refugees and Natives

(a) Literacy in 1928



(b) Junior high school completion in 2011



Notes. Panel *A* plots literacy rates for all Greeks (white bars), aged 15–64, living in the countryside (rural) in 1928, natives (gray bars), and refugees (black bars). Literacy rates are retrieved from the 1928 General Population Census. Panel *B* plots the share of Greeks, aged 15–64, who had completed junior high school (9 years of schooling) in 2011 using data from the 2011 General Population Census. White bars give the mean values for all those born in rural settlements, gray bars across those born in rural native settlements, and black bars across those born in rural, predominantly refugee settlements. Both panels report means across Greece (left), and Macedonia and West Thrace (right).

The difference disappeared only when literacy became universal, following a significant educational reform in 1928–30 that made primary schooling (6 years) compulsory.

Contemporary Patterns Figure 4–Panel *B* plots the share of 15–64-year-olds who completed junior high school (9 years of schooling) for the two sets of birth settlements—across the country (left) and in Macedonia and West Thrace (right)—using the data from the 2011 General Population Census. Junior-high-school completion rates in refugee birth settlements were a staggering ten percentage points *higher* than in native birth settlements in Macedonia and West Thrace (64.8% vs 54.5%).

Examples Two sets of examples illustrate these patterns. Take Repanidion, a native settlement, and Lera, a refugee settlement 10 km apart, on Limnos Island. In 1928, their populations were comparable, 632 and 556, respectively. Among those born before 1917, the primary school completion was two-thirds (67.4%) for the exclusively native population of Repanidion and less than a third in Lera (32.5%). Descendants of refugees in Lera who attended school in Greece leapfrogged the native population of Repanidion. Among the residents of Lera born after 1917, the high school completion rate was 20.4% while in Repanidion just 12.7%. Likewise, take Promakhi, an autochthonous town with a population of 1,160 in 1928, and Polikarpi, a predominantly refugee one, 10 km away, with a similar

population in 1928—1,070 inhabitants, in Almopia province, Pella, Central Macedonia. Primary school completion rates for those born before 1917 were about 19.6% in Promakhi and 16.3% for the refugee population of Polikarpi. But residents of Polikarpi made considerable educational progress, and high school completion rates were double in the refugee town (15.5% vs 8%) for those born after 1917.

4.2 Before-After Analysis

To formally explore how displacement shaped educational choices among refugees and their lineage, we first estimate difference-in-difference specifications that compare education's (dynamic) evolution (across cohorts) before and after the population exchange between residents of predominantly refugee and nearby native settlements.

4.2.1 Empirical Specification

The difference-in-difference specification takes the following form:

$$Y_{ispct} = \underbrace{\lambda_{(c)} REF_s^{post}}_{\substack{\textbf{Uprootedness} \\ \textbf{Post Displacement}}} + \underbrace{\phi_{(c)} REF_s^{1907-1916}}_{\substack{\textbf{Repression} \\ \textbf{Pre-Displacement}}} + \alpha_s + \beta_{p,c} + \gamma_{t,c} + X'_{s,c}\Psi_{s,c} + Z'_{i,c}\Gamma_{i,c} + \epsilon_{ispct} \qquad (1)$$

Y_{ispct} denotes years of schooling, as well as indicators for completed primary, secondary, and tertiary education of individual i, residing in settlement s of province p, born in a 10-year birth-cohort c, and recorded in census year t. REF_s is an indicator that takes the value of one for individuals residing in a predominantly refugee settlement. For the baseline estimates, we classify as refugee settlements those where refugees' share in 1928 exceeded 80%, while "native" settlements are those where the corresponding share was below 5%; we explore the stability of the results to alternative cutoffs below. $\lambda_{(c)}$ captures the differences in the outcome between Greeks residing in refugee and non-refugee settlements compared to the reference cohort, individuals born between 1887 and 1896, whose schooling decisions took place before the displacement and before the Ottoman restrictions on schools and the intensification of the attacks on the Greek population (*Uprootedness*). To capture those affected by the turmoil, ethnic tension, and warfare of the 1910s and early 1920s (*Repression*), we interact the refugee indicator with the likely impacted cohorts born between 1907 and 1916. Besides estimating average before-after refugee-native settlement differences, we also explore their dynamics by interacting REF_s with cohort indicators.

Settlement-of-residence constants, α_s, account for local time-invariant factors, which may affect the settlement process and education dynamics. Birth-cohort fixed effects interacted with province constants, $\beta_{p,c}$ allow for differential educational attainment trends, which rises considerably within relatively homogeneous economic zones ("provinces") during the last century. We also exploit more granular variation, allowing for differential trends across sub-province units (Voronoi polygons). $\gamma_{t,c}$ net out cohort-specific differences across the five censuses. The specification also accounts for settlement features, $X'_{s,c}$, allowed to differ across cohorts. Vector X_s comprises location features: log

distance to (i) the coast, (ii) the provincial capital, (iii) a railway stop, and (iv) reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; log population in 1920, and an indicator for settlements with a monastery in 1920. The empirical model also includes gender and age, $Z'_{i,c}$, allowed to differ across birth cohorts. The error term, ϵ_{ispct}, is clustered at the settlement-of-residence level.[16]

4.2.2 Difference-in-Difference Estimates

Table 1 reports the difference-in-difference estimates looking at years of schooling (columns (1)–(2)), and the indicator variables for completing primary (columns (3)–(4)), junior high (columns (5)–(6)), high school (columns (7)–(8)), and tertiary education (columns (9)–(10)) among Greek citizens aged 25–90 residing in rural, predominantly refugee and native settlements. Odd-numbered columns give within-province estimates. Even-numbered columns restrict the estimation to refugee and non-refugee settlements no more than 25 km apart in the sub-province unit. Despite narrowing the comparisons, the local sample is considerable, comprising 90, 609 individuals residing in 237 predominantly refugee settlements and 266, 559 residents of 1, 001 native settlements in 123 provinces.

Static Estimates Compared to those born between 1887 and 1896, residents of predominantly refugee settlements who started schooling after the population exchange experience an increase of about 0.5 years of education over their counterparts in native settlements in the same province (column (1)). The magnitude is substantial, as residents of native villages born in the Interwar period had just 3.2 years of schooling, and those born between 1937–76 had 7 years. The specifications also quantify the impact of the repression that Greek-Orthodox communities faced post-1908 in the Ottoman Empire; the coefficient for the cohorts affected by warfare is significant, implying a drop in schooling years of about 0.2 compared to the earlier cohorts born in the 19th century. Columns (3)–(10) show that residents of predominantly refugee settlements have significantly higher propensities of completing primary, secondary, and tertiary education than Greeks born in native settlements in the same province or even sub-province units. On average, living in a predominantly refugee settlement increases the post-population-exchange completion rates of junior high school by 3.9 percentage points (pps) (15.5% of the mean increase for natives), high school completion rates by 3 pps (15.6%), and tertiary education rates by 1 pps (18.5%). Across all outcomes, the coefficients on the refugee indicator are smaller in the local sample. Years of schooling increase by 0.33 years, primary school by 3.6 pps, junior high school completion rates by 1.9 pps (7.8% of the average increase for natives), high school completion rates by 1.5 pps (8.4%), and tertiary education by 0.4 pps (7%).

[16] We also calculated Conley standard errors with various distance cutoffs to account for spatial correlation, finding, again, similar results (Conley (1999), Bester, Conley and Hansen (2011), and Colella et al. (2019)). See Appendix Figure B11.

Table 1: Refugees and Human Capital Dynamics (1971–2011)

Dependent Variable:	Years of schooling		Completed education level: Primary school		Jun. high school		High school		Tertiary education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Refugee settlement × Post	0.486***	0.328***	3.998***	3.580**	3.910***	1.840**	2.955***	1.440**	1.012***	0.378
	(0.110)	(0.121)	(1.497)	(1.633)	(0.749)	(0.805)	(0.658)	(0.629)	(0.340)	(0.339)
Refugee settlement × War cohort	-0.177*	-0.285***	-2.950**	-4.403***	0.338	-0.105	0.013	-0.002	-0.259	-0.444
	(0.097)	(0.110)	(1.408)	(1.612)	(0.540)	(0.648)	(0.403)	(0.497)	(0.220)	(0.312)
R-squared	0.461	0.488	0.356	0.382	0.398	0.438	0.300	0.327	0.090	0.104
Observations	1,476,538	357,168	1,476,538	357,168	1,476,538	357,168	1,476,538	357,168	1,476,538	357,168
Average outcome in native settlements: 1887-1906	2.0	1.5	31.0	23.7	3.9	2.3	1.2	0.9	0.4	0.3
Average outcome in native settlements: 1907-1916	2.7	2.1	40.7	33.5	5.6	2.8	2.0	1.1	0.5	0.4
Average outcome in native settlements: 1917-1986	6.4	6.1	77.6	75.6	29.0	25.8	20.1	18.0	5.8	5.7
Settlement FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Census year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Settlement controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Province FE	✓		✓		✓		✓		✓	
Cohort × Sub-province FE		✓		✓		✓		✓		✓
# Refugee observations	128,305	90,609	128,305	90,609	128,305	90,609	128,305	90,609	128,305	90,609
# Refugee settlements	360	237	360	237	360	237	360	237	360	237
# Native observations	1,348,233	266,559	1,348,233	266,559	1,348,233	266,559	1,348,233	266,559	1,348,233	266,559
# Native settlements	5,922	1,001	5,922	1,001	5,922	1,001	5,922	1,001	5,922	1,001
# Admin units for FE	142	123	142	123	142	123	142	123	142	123

Notes. All specifications report OLS estimates associating various educational outcomes to a binary refugee settlement indicator, *Refugee settlement*, that identifies those residing in predominantly refugee villages where the share of refugees in 1928 exceeded 80%. It equals zero for Greeks residing in native settlements, where the refugees' share in 1928 was below 5%. *Post* indicates cohorts born after 1917 who started school in Greece after the Population Exchange. *War cohort* is an indicator for cohorts born between 1907 and 1916 who were of school age during the Balkan Wars, the Great War, and the Greco-Turkish War. The sample is restricted to Greek citizens aged between 25 and 90. The dependent variable is years of schooling in columns (1)–(2); an indicator for completing primary school in columns (3)–(4); an indicator for completing junior high school in columns (5)–(6); an indicator for completing high school in columns (7)–(8); and an indicator for completing tertiary education in columns (9)–(10). Odd-numbered columns focus on the sample of all rural refugee and native settlements, conditioning on birth-cohort specific province (*eparchia*) fixed effects. Even-numbered columns focus on the local sample, conditioning on birth-cohort specific sub-province (Voronoi polygon) fixed effects. In each Voronoi polygon, refugee and native settlements are less than 25 km away. All specifications include settlement fixed effects, interactions of census-year and birth-cohort fixed effects, interactions of individual characteristics (age and gender) and birth-cohort fixed effects, and interactions of various settlement characteristics with birth-cohort fixed effects. Settlement characteristics include log distances to the coast, to the provincial capital, to railway stops, and to reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1887–1906 birth cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Examples A few examples help illustrate the difference-in-difference estimates. Take Mega Doukato, a native settlement, and Adriani, a refugee settlement, in Komotini, West Thrace. The two villages, only 5.7 km apart, had similar populations in 1928, 465 and 411. Years of schooling for those born before 1917 was 0.5 in the native settlement and close to 0 in the refugee one. But when focusing on those born after 1917, the difference is reversed, as now schooling years are 3.9 in Mega Doukato and 4.5 in Adriani. The comparison of Gouledіana (196 inhabitants), a native settlement, and Xerokhorion (176 inhabitants), a refugee settlement 7.6 km away, offers an illustration from Rethymnon, Crete. Years of schooling for those born before 1917 was much higher in the native village (3.0 vs 0), but for those attending school in Greece, the gap is reversed, with higher average schooling in the predominantly refugee village (6.8 vs 6.3 years). The difference is striking when looking at high-school completion rate, which in the refugee settlements is three times higher (30% vs 8.3%).

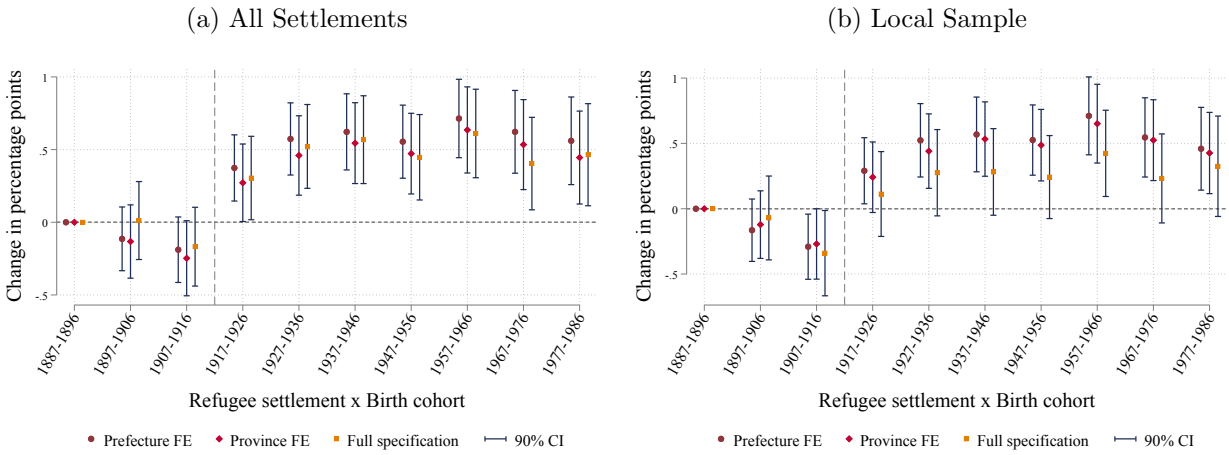
Dynamics Figure 5 illustrates the evolution in years of education across 10-year birth cohorts between residents of refugee and native villages across all rural settlements (Panel A) and in the local sample (Panel B). The patterns are similar when exploring within-prefecture or within-province variation without any controls and weaken modestly when conditioning for settlement and individual characteristics (*Full specification*). Figure 6 shows how the uncovered stability of the educational gain of refugee vis-à-vis native communities reflects a continuous intergenerational upgrading along the educational ladder throughout the 20th century.[17] Refugees born in Turkey (and also Bulgaria) just before displacement (1917–23) or shortly after settling in Greece (1924–37) have a five to seven percentage points higher likelihood of completing primary education than natives in nearby settlements. Over time, as almost all Greeks completed elementary school, those born in refugee settlements after the mid-1950s enjoyed a five to nine percentage points higher likelihood of completing junior high (9 school years) and high school (12 years). As more Greeks born after WWII completed high school, differences between refugee and native settlements appeared in college education. Generation-X Greeks residing in refugee villages have about 2 percentage points higher likelihood of completing university. (See Appendix Figure B4 for the local sample estimates.)

4.3 Leveraging Birthplace Variation and Accounting for Residence

By comparing residents in refugee and native settlements, the analysis misses those migrating to the big cities (Athens, Piraeus, and Thessaloniki) or elsewhere. Fortunately, the 2011 General Population Census recorded not only an individual's current residence but also their mother's residence at birth at a granular level (the stable geographic unit); we refer to the latter as birth settlement hereafter. Hence, our second estimation design exploits variation across place of birth (instead of residence), accounting for the considerable internal migration in the 1950s–1980s.

[17] We also calculated measures of absolute intergenerational mobility in education to study dynamics. The analysis suggests significantly higher upward educational mobility in refugee settlements when looking at the offspring of parents without completed primary or without (junior) high school (results available upon request).

Figure 5: Years of Schooling across Refugee and Native Settlements (1971–2011)



Notes. Both panels present various event-study OLS estimates that associate years of schooling to an indicator variable that takes the value of one for individuals residing in predominantly refugee settlements where the share of refugees in 1928 exceeded 80%, and zero for individuals in native settlements, where the share of refugees in 1928 was below 5%; regression equation (1). Panel *A* gives estimates across all rural refugee and native settlements (baseline sample). Panel *B* focuses on the local sample, where refugee settlements are matched to native ones no more than 25 km away in the same sub-province unit (Voronoi polygon). Maroon circles and red diamonds plot the estimates from parsimonious specifications exploiting within-prefecture (*nomos*) and within-province (*eparchia*) variation, respectively; these specifications condition on settlement fixed effects, interactions of census and birth-cohort fixed effects, and interactions of the prefecture or the province fixed effects and birth-cohort fixed effects. Orange squares plot the estimates from a rich (full) specification that further controls for interactions between birth-cohort constants with individual and settlement characteristics, and also for interactions of sub-province fixed effects and birth-cohort constants in Panel *B*. Individual characteristics are age and gender. Settlement characteristics are: log distances to the coast, to the provincial capital, to railway stops, and to reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1887–96 birth cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level.

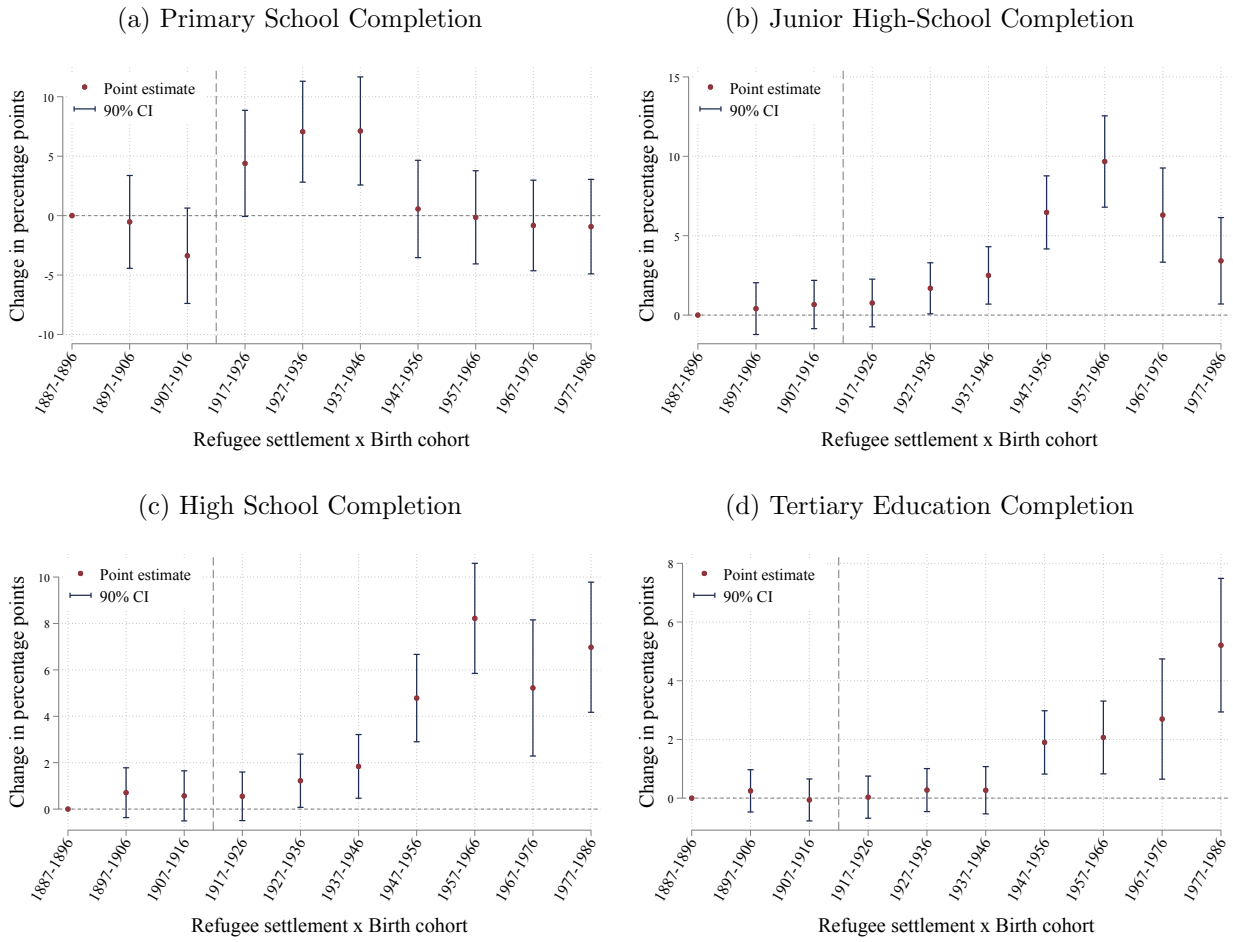
4.3.1 Accounting for Residence among Stayers and Internal Migrants

Our second specification reads:

$$Y_{ispcr} = \underbrace{\lambda_{(c)} REF_s}_{\substack{\textbf{Uprootedness} \\ \textbf{Post Displacement}}} + \underbrace{\mu_r}_{\textbf{Residence}} + \beta_{p,c} + X'_{s,c}\Psi_{s,c} + Z'_{i,c}\Gamma_{i,c} + \xi_{ispcr} \tag{2}$$

Y_{ispcr} denotes the education of individual i, born in settlement s of province p and in a 10-year birth-cohort c, and residing in, potentially different, municipality r in 2011. The specification utilizes variation not only among Greeks residing in their birthplace (either a predominantly refugee settlement or not) but also from respondents living outside their settlement of birth, "movers." The specification thus accommodates municipality-of-residence fixed effects, μ_r, that allow comparing individuals born in nearby refugee and native settlements living in the same municipality (typically city in the metropolitan areas of Thessaloniki and Athens-Piraeus). $\beta_{p,c}$ are province-specific birth cohort constants. The coefficient on the $REF_{(s)}$ indicator, $\lambda_{(c)}$, captures (cohort-specific) differences in education between those born in refugee and non-refugee villages residing in 2011 in the same

Figure 6: Education Dynamics across Refugee and Native Settlements (1971–2011)



Notes. The panels plot event-study OLS (linear probability model) estimates obtained from associating indicator variables for completing primary school (in Panel *A*), for competing junior high school (in Panel *B*), for competing high school (in Panel *C*), and for completing tertiary education (in Panel *D*) on a birth-cohort varying binary variable indicating whether an individual resides in a predominantly refugee settlement, where the share of refugees in 1928 exceeded 80%. The indicator equals zero for those born in native settlements, where the share of refugee in 1928 was below 5%. All specifications include settlement fixed effects, interactions of census-year and birth-cohort fixed effects, interactions of province (*eparchia*) fixed effects and birth-cohort fixed effects, interactions of individual characteristics (age and gender) and birth-cohort fixed effects, and interactions of various settlement characteristics and birth-cohort fixed effects. The settlement characteristics are: log distances to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1887–96 birth cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level.

municipality. As few respondents in the 2011 Census were born before 1920, specification (2) does not exploit before-after differences. But as pre-displacement differences in education between refugees and natives (and between residents of refugee and native settlements) are small—with natives having marginally higher education—the estimates still identify the impact of forced displacement on human capital. Given the specification's "bilateral" nature, we adjust standard errors for clustering at the birth-settlement and the municipality-of-residence levels.

Table 2 presents the estimates of specification (2) with schooling years (of 25-to-90-year-olds) as the outcome. (Appendix Table B2 gives results with the educational attainment.) The full sample estimates in columns (1)–(2) show that individuals from predominantly refugee settlements have more than a half year (0.58) of schooling than Greeks born in native settlements in the same province, and residing in the *same* municipality (column (2)); for comparison, mean years of schooling among those born in native settlements were eight years. Reassuringly, the educational advantage estimated based on birthplace is close to the estimate in Table 1–column (1) that used only variation in residence.

In columns (3)–(4) and (5)–(8), we estimate specification (2) separately for Greeks residing in their birthplace ("non-movers") as well as in settlements other than their birth one ("movers") to shed light on the underlying mechanisms. If human capital investments are essential for integrating minority groups, especially for those moving to more competitive urban labor markets, then one should expect stronger effects for migrants. Likewise, if refugees' educational investments reflect low attachment to the rural settlement in Greece and a desire to work in skill-intensive services and manufacturing, the associations should be more salient among movers. However, if the displacement-education nexus reflects a deeper valuation of human capital, we should still observe this educational advantage even among those who continue residing in the countryside and working in the same industry. Columns (3)–(6) show that descendants of refugees living both in and outside their birthplaces have about half a year of extra schooling (0.5–0.56). Focusing on internal migrants in the local sample yields only slightly smaller estimates (columns (7)–(8)).

In addition, we estimate specification (2) augmented with cohort-industry of employment fixed effects among 25-to-55-year-olds, as employment drops significantly after that age. Since the employment sector is an outcome, these specifications are prone to the "bad controls" critique. However, they shed light on the underlying motives for educational investments. Two key results emerge from the analysis in Appendix Table B3. First, the coefficient drops by a third (0.53 vs 0.32), suggesting that the increased human capital among refugees allows them to join industries with higher educational requirements. In line with this, in Michalopoulos et al. (2025), we show that refugees are more likely than Greeks from neighboring autochthonous rural settlements to migrate to urban hubs and abroad. Second, even when comparing Greeks within the same municipality and industry, those from predominantly refugee settlements are systematically more educated than individuals from autochthonous villages and towns.

The displacement-education link uncovered among non-movers, internal migrants, and those employed in the same industry suggests that the educational investments of refugee descendants stem from a heightened valuation of education—one that extends beyond the aspirations of second-, third- and fourth-generation refugees to migrate to cities or enter human-capital-intensive occupations, where returns to education are often higher. This emphasis on human capital can be understood within its historical and cultural context. In the 1920s, many refugees harbored hopes of returning to their ancestral homes in Anatolia. Although the 1930 and 1933 pacts between Greece and Turkey put to rest these aspirations by affirming the Treaty of Lausanne as the final settlement of their borders the nos-

talgia endured. The collective trauma of displacement, the longing of the so-called *Lost Motherlands* (Hamenes Patrides) remained deeply ingrained, persisting through memory, literature, oral histories, and songs that revolve around the wound of uprootedness, emphasizing the importance of learning to survive in a new environment, and perseverance amidst adversity and uncertainty (Logothetis, Michalopoulos and Papaioannou, 2025).

Dynamics The dynamics of educational attainment echo the patterns identified with the event-study estimates. Primary school completion rates increase significantly for the second generation of refugees, while the positive impacts on secondary and tertiary education materialized for third- and fourth-generation refugees. The estimates illustrate a spike in primary school completion for refugee children (1917–26) and second-generation refugees (1927–46), followed by a higher likelihood of completing (junior) high school and even tertiary education for third- and fourth-generation refugees. (See Appendix Figures B5 and B6.)

Table 2: Displacement and Years of Schooling in 2011, cond. on Residence

Sample:	All sample		Non-migrants		Internal migrants			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Born in a refugee settlement	0.708***	0.576***	0.569***	0.565***	0.674***	0.495***	0.447***	0.339***
	(0.131)	(0.127)	(0.132)	(0.169)	(0.156)	(0.147)	(0.149)	(0.129)
R-squared	0.356	0.402	0.441	0.457	0.331	0.369	0.363	0.403
Observations	549,480	549,474	243,038	243,005	289,213	289,208	62,495	62,421
Average outcome in native settlements	8.0	8.0	7.2	7.2	8.6	8.6	8.3	8.3
Cohort × Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth settlement controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth province FE	✓	✓	✓	✓	✓	✓		
Cohort × Birth sub-province FE							✓	✓
Residence municipality FE		✓		✓		✓		✓
# Refugee observations	44,115	44,114	18,259	18,255	22,674	22,674	14,604	14,592
# Refugee birth settlements	357	357	351	351	355	355	235	235
# Native observations	505,365	505,360	224,779	224,750	266,539	266,534	47,891	47,829
# Native birth settlements	5,788	5,788	5,688	5,688	5,704	5,704	971	971
# Birth provinces for FE	141	141	140	140	140	140	41	41
# Residence-municipality FE		1,021		943		1,020		857

Notes. The table reports OLS estimates associating years of schooling and a binary variable that takes the value of one for those born in a refugee settlement (where the share of refugees exceeded 80% in 1928) and zero for those born in a native settlement (where the share of refugees in 1928 was below 5%). Columns (1)–(2) consider all individuals born in all rural predominantly refugee and native settlements (baseline sample). Columns (3)–(4) restrict the estimation to individuals who reside in their birth settlement and who did not migrate abroad. Columns (5)–(6) restrict the estimation to those who reside in a settlement other than their birth settlement. Columns (7)–(8) further restricts the estimation to the local sample, where refugee and native birth settlements are within a 25 km radius. All specifications include interactions of individual characteristics (age and gender) and birth-cohort fixed effects, interactions of birth-settlement characteristics and birth-cohort fixed effects. In columns (1)–(6) we add interactions of birth-province fixed effects and birth-cohort fixed effects whereas in columns (7)–(8) include interactions of the birth-subprovince (Voronoi polygon) fixed effects and birth-cohort fixed effects. Columns (2), (4), (6), and (8) control for municipality-of-residence fixed effects. Birth settlement characteristics are: log distances to the coast, to the provincial capital, to railway stops, and to reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. Standard errors are (double) clustered at the birth-settlement and the municipality-of-residence levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Comparison The half-year (0.5–0.6) schooling increase for second-, third-, and fourth-generation refugees compared to Greeks from nearby native villages (with a mean of 8 years) is comparable to

the survey-based estimates of Becker et al. (2020), which estimates that migrants from the Eastern province of Kresy, ceded by Poland to the Soviet Union, have 0.8–0.9 extra years of schooling in 2015–6; mean of control 11.9.[18] Our estimates are quantitatively larger than the ones documented by Sarvimäki, Uusitalo and Jäntti (2022), who show that Finnish migrants from the Eastern areas ceded to the Soviet Union eventually experienced modest educational gains as a result of moving out of agriculture. The strong effects we identify go against the tiny and, in general, insignificant effects on education from the forced relocation of German communities from Eastern Europe to West Germany after WWII (e.g., Bauer, Braun and Kvasnicka (2013), Bauer and Kvasnicka (2014), Peters (2022), Ciccone and Nimczik (2023)).

4.4 Further Evidence

We conducted various checks to further shed light on the dispalcement-education nexus and explore the sensitivity of our estimates.

Spatial Spillovers First, we explored spatial externalities and whether the SUTVA assumption, requiring no spillovers from the "treatment" to "control" settlements, holds. Appendix Table B4 gives difference-in-difference estimates testing whether the presence of refugees nearby impacts schooling in native settlements. For each native settlement, we compute the share of refugees in 1928 within 5 km, considering all settlements. The analysis does not reveal evidence of spatial spillovers both in the baseline and the local sample.

Gender Second, we explore gender differences. The estimates are significant for men and women and are slightly larger for women. Moreover, consistent with the historical narratives that males were disproportionately affected by the ethnic strife in the decades leading to displacement, the cohort-specific estimates reveal that males (unlike females) suffered a substantial decline in primary schooling in the years leading to the exodus; see Appendix Figures B7 and B8, and Appendix Table B5.

Cutoffs Third, we explored the sensitivity of the estimates to using more or less restrictive cutoffs to distinguish between predominantly refugee and native settlements. We repeated our estimations classifying a settlement as predominantly refugee if the share of refugees in 1928 exceeds 50%, 60%, 70%, 80% (baseline), and even 90%. Likewise, we classify native settlements as those where the share of refugees is less than 10%, 5% (baseline), and even 1%. Two results stand out from these 15 combinations (Appendix Figure B9). First, all coefficients on the refugee indicator are more than two standard errors larger than zero, suggesting that no particular cutoff drives the results. Second, when we compare settlements where the share of refugees exceeds 90% to those without any refugees

[18] Looking at peaceful migration after a volcanic eruption in cities, Nakamura, Sigurdsson and Steinsson (2022) report a considerable increase of 3.5 years of schooling for displaced population younger than 25 children, with a control mean of 13.4, in a sample of 2,262 respondents.

(less than 1%), the refugee-native gap in years of schooling is larger than in the baseline classification, suggesting that if anything, the baseline estimates are attenuated.

Very Local Comparisons Fourth, we re-estimated the model, imposing that refugee and native settlements within the same province to be no farther than 20 km, 15 km, 10 km, and even 5 km. The estimates' statistical and economic significance is stable despite the drop in observations (Appendix Figure B10); the coefficient on the refugee indicator for schooling years is about 0.3 to 0.5, even restricting estimation to predominantly refugee and native settlements no more than 5 kilometers apart.

All Rural Settlements Fifth, we explore the displacement-human capital nexus across all rural settlements, replacing the binary refugee settlement indicator with the share of refugees in 1928 as the variable of interest (Appendix Tables B6–B7). Doing so makes the comparisons less sharp, as the sample blends predominantly refugee and native settlements with mixed ones, which are significantly more populated. However, it allows studying the link between education and refugee share in the entire sample, using information from all 431,895 refugees settling in the countryside across 7,507 rural settlements in all (35) prefectures and (141) provinces. All permutations reveal significantly higher schooling for Greeks residing in or born in places with a larger share of refugees in 1928. The estimates suggest that moving from a rural settlement without any refugees to an exclusively refugee settlement in the same province is associated with an increase of 0.4 years of schooling, similar to the baseline estimates with the indicator variable reported in Tables 1 and 2.

5 Uprootedness and Portable Skills

5.1 Hypothesis and Approach

A core aspect of the uprootedness hypothesis linking forced displacement (and discrimination) to human capital regards its *portability* aspect. However, it has been empirically challenging to evaluate this dimension and distinguish its importance relative to (minority) group-specific differences in access, taste, and returns to education. We make progress on this front by recognizing that not all forms of human capital are equally transferable; the literature has, for example, pinpointed that medicine or engineering skills are more "portable" than law, which by its country-specific nature and tightly regulated labor market offers valuable, albeit not easily transferable skills (Ayal and Chiswick (1983), Brenner and Kiefer (1981), Stigler and Becker (1977), Kessel (1958)).[19]

We empirically examine this conjecture by exploiting the variation in the transferability of skills acquired in tertiary education.[20] We utilize information on the tertiary degree pursued as recorded

[19] Beine, Fromentin and Bachiller (2024) show that labor market conditions abroad impact individuals' choices of university degrees when emigrating.

[20] In Greece, tertiary education degrees are granted by (i) universities and polytechnics (AEI), (ii) Technological Educational Institutes (TEI), which became AEI in 2018, (iii) military, police, firefighter and theological academies, and

in the 2011 General Population Census. The latter provides *the study area and the type of degree-granting institution*; for example, accounting, finance, dentistry, geology, philosophy, Greek literature, medicine, law, police academy, electrical engineering, computer science, naval engineering, pedagogical studies, public administration, and economics. We classify the 266 fields of study into relatively portable degrees and those that are less so, and explore whether college-educated Greeks from refugee settlements are more likely to complete programs teaching (more) portable skills.

Classifying tertiary education degrees into those with (more) portable skills entails subjectivity. We rely on GPT-4 for the initial classification and then manually check the plausibility of each assessment. Appendix Table C1 reports the classification for the 50 most-studied degrees nationwide and across the rural settlements in the baseline sample. The top-50 degrees account for more than 80% of the tertiary degree graduates in Greece. 206 fields of study were classified as transferable and 60 as non-transferable. Examples of degrees with transferable skills include economics, medicine, accounting, engineering, IT, computer science, and statistics. Examples of degrees with limited transferability include all Greek military schools (navy, air force, infantry), police academies, theology (as the curriculum is tilted towards Greek-Orthodox Studies), Greek literature, public administration, and the numerous pedagogical schools set up to train teachers on the centralized Greek schooling system. We then estimate the regression equation (2), with the portability of tertiary degrees indicator as the outcome, focusing on Greek citizens, aged 25 to 90, who completed tertiary education.

5.2 Results

Table 3 gives the linear probability model estimates. The specifications (1)–(2) look across all individuals born in rural settlements with completed tertiary education in 2011; the analysis spans 61,985 individuals with 7% coming from predominantly refugee settlements. College graduates from refugee settlements have a 6.1 percentage points (pps) higher likelihood of pursuing a tertiary degree with transferable skills than college-educated Greeks from native settlements in the same province, with similar observable characteristics, living in the same municipality (column (2)). The magnitude is not small, as about 66% of Greeks from non-refugee rural settlements complete a college degree with portable skills. We estimate a 6.7-pps gap in the much smaller local sample, comprising 13,639 college graduates, 22.5% from refugee settlements (in columns (3)–(4)).[21] The comparisons of a few neighboring pairs of settlements are telling. More than three-fourths (78.3%) of the college graduates from the refugee settlement of Polikarpi in Pella, Macedonia, pursued degrees with portable skills; less than half (48.3%) did in the nearby native settlement of Promakhi. In Central Greece (Thessaly), in Tsiampaslar (Vounena), a refugee settlement, all college graduates complete portable degrees, while in nearby Kourtiki (Metamorphosi), with a similar population in 1928 (366 vs 394) the corresponding share is 71%.

(iv) professional studies programs.

[21] The results are similar when comparing those attending tertiary education from predominantly refugee and native settlements within 20 km, 15 km, and even 10 km radii. Besides, the results are not sensitive to the refugee population share in 1928 to distinguish between predominantly refugee and native settlements. See Appendix Figure C2.

Given the uncertainty in classifying university degrees' portability, in columns (5)–(6), we only look at those with degrees in law, medicine, and all engineering degrees granted by an AEI/university (mechanical, electrical, computer science, naval, aerospace, chemical, architecture, material) but omit civil, as its curriculum is embedded in the Greek legal framework regulating construction and zoning laws. The uprootedness hypothesis would suggest a higher propensity to study medicine and engineering than law, which is hardly transferable across countries. As the Law School, the School of Medicine, and the Polytechnic Schools for Engineering have been the most prestigious and selective schools for aspiring Greeks and their families, by focusing on them, we examine the legacy of forced displacement from Anatolia among the most academically selected.[22] University graduates with refugee lineage have a significantly higher propensity to study medicine and engineering rather than law compared to natives from the same cohort, born in the same province, and residing in the same municipality in 2011. The coefficient suggests an 8.5 to 15.8 pps increase in likelihood—a sizable effect given the 56.8% average for those born in native settlements. (When we include civil engineering in the transferable skills, the estimates imply an increase of about 7 to 13 pps.) In columns (7)–(8), we omit medicine, as there are some licensing barriers for doctors to move across borders, and only compare engineering to law. Despite the sample drop, there is a 12 to 22 pps increase in the probability of refugee descendants studying engineering compared to law.

Taking Stock While earlier scholars have long argued that the uprootedness hypothesis implies investments in portable skills, it has been challenging to detect it. Therefore, these results, exploiting variation among college graduates—and even within the highly selected group of those studying law, engineering, and medicine—support a core aspect of the uprootedness hypothesis, stressing the portability of human capital investments.

6 Additional Mechanisms

In this section, we explore the rich heterogeneity of refugees' backgrounds, trajectories, and settlement characteristics to shed further light on the underlying economic forces; besides, we explore the sensitivity of the results to various (restrictive) tests. Appendix Section D gives descriptives, summary statistics, and additional results.

6.1 Approach

A helpful framework for understanding the educational choices of refugees and natives is the Roy assignment model, suggesting that individuals' education is based on (i) the relative (perceived) re-

[22] Annual nationwide standardized exams determine admission to each school of the exclusively public Greek university system. Students indicate their ranked preferences to the Ministry of Education, and the cutoff (base) for each school and field of study (e.g., Law at the National and Kapodistrian University of Athens or Architecture at the University of Macedonia) is determined after the exams, based on the submitted preferences and the grades. Historically, the base for the highly competitive nationwide exams entering Law, Engineering, or Medical Schools exceeds 18/20.

Table 3: Displacement and College Skill Portability across Rural Native and Refugee Birth Settlements

Sample:	All sample		Local sample		All sample Law., doc., eng.		All sample Lawyers & engineers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Born in a refugee settlement	6.368***	6.074***	6.513***	6.670***	8.471**	15.843***	12.070**	22.037***
	(1.246)	(1.347)	(1.836)	(2.433)	(4.227)	(5.449)	(5.030)	(5.865)
R-squared	0.053	0.062	0.058	0.070	0.116	0.135	0.201	0.231
Observations	62,049	62,035	13,643	13,499	7,076	6,930	4,752	4,557
Average outcome in native settlements	66.0	66.0	65.2	65.2	56.8	56.8	36.5	36.5
Cohort × Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth settlement controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth province FE	✓	✓			✓	✓	✓	✓
Cohort × Birth sub-province FE			✓	✓				
Residence municipality FE		✓		✓		✓		✓
# Refugee observations	4,335	4,333	3,074	3,046	466	435	300	270
# Refugee birth settlements	332	332	217	216	185	172	153	138
# Native observations	57,714	57,702	10,569	10,453	6,610	6,495	4,452	4,287
# Native birth settlements	5,081	5,079	870	866	2,673	2,625	2,194	2,109
# Admin units for FE	140	140	122	122	134	132	131	126
# Residence-municipality FE		984		531		700		585

Notes. The table reports OLS (linear probability model) estimates, associating indicator variables on the portability of tertiary-degree skills with a binary variable that takes the value of one for individuals born in a refugee settlement, where the share of refugees exceeded 80% in 1928, and zero for individuals born in native settlements, where the share of refugees in 1928 was below 5%. The outcome variable in columns (1)–(4) takes the value one for fields of study classified as transferable by ChatGPT-4, and zero otherwise. The outcome variable in (5)–(8) equals one for individuals with a medical or an engineering degree from universities and polytechnics and zero for individuals with a law degree. Columns (1)–(2) consider all individuals with a university/college degree across all rural refugee and native settlements, while columns (3)–(4) restrict the estimation to the local sample, which focuses on individuals born in refugee and native settlements in the same province, no more than 25 km away. Columns (5)–(6) consider individuals with a degree in law, medicine, and engineering (excluding civil engineering) from universities and polytechnics across all rural refugee and native settlements. Columns (7)–(8) consider individuals with a degree in law and engineering from universities and polytechnics across all rural refugee and native settlements. All specifications include the interactions of individual characteristics (age and gender) and birth-cohort fixed effects and the interactions of birth-settlement characteristics and birth-cohort fixed effects. In addition, columns (1)–(2) and (5)–(8) include interactions of the birth-province fixed effects and birth-cohort fixed effects. Columns (3)–(4) control for the interactions of the birth-sub-province (Voronoi polygon) fixed effects and birth-cohort fixed effects. Columns (2), (4), (6), and (8) control for municipality-of-residence fixed effects. The set of birth settlement characteristics comprises log distance to the coast, to the provincial capital, to railway stops, and to reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. Standard errors are clustered at the birth settlement and the municipality of residence levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

turns to education, (ii) the costs of acquiring human capital, and (iii) transitioning frictions to the nonagricultural labor market.[23] The results in Section 4 reveal that refugees arriving in the Greek countryside with lower levels of education accumulate human capital more rapidly than natives, eventually surpassing them. But how generalizable is the comparative advantage of refugees in education? If uprooted individuals place greater value on human capital due to its portability or non-pecuniary benefits and/or face lower (psychological) costs of leaving agriculture, their educational advantage should be similar across different settings. However, some origin or destination features might amplify or attenuate the propensity of refugees and their offspring to invest in human capital.

Recognizing that expellees came from diverse backgrounds, faced different displacement trajectories, and settled in destinations with non-negligible variation in agricultural productivity, proximity to markets, and manufacturing presence, we explore the heterogeneity of effects to uncover additional

[23]The literature employing Roy assignment models to study migration and asymmetric information is considerable. See Lagakos, Mobarak and Waugh (2023), Bryan and Morten (2019), Young (2013), Nakamura, Sigurdsson and Steinsson (2022), Sarvimäki, Uusitalo and Jäntti (2022), among many others; Lagakos (2020) provides an overview.

mechanisms and assess the stability of the benchmark results across these dimensions. Specifically, we ask: What background features influenced the behavior of refugees and their descendants after the population exchange? Did specific destination conditions at the settlement or broader labor market level enhance or weaken refugees' educational gains? We rerun the baseline regression equations (1) and (2) interacting the REF_s variable identifying those either living or born in predominantly refugee settlements (where the share of refugees in 1928 exceeded 80%) with various origin and destination features. Appendix Table D1 gives summary statistics of the variables we consider in the heterogeneity analysis across all predominantly refugee settlements.

6.2 Results

Figure 7 summarizes the heterogeneity analysis of the forced-displacement shock on schooling years, reporting the difference-in-difference estimates (Panel A) and the conditional on residence estimates in 2011 (Panel B). All panels report OLS estimates from rich specifications, which exploit differences between those residing (or born) in predominantly refugee and native settlements in the same province, conditioning also on interactions of birth-cohort fixed effects with individual attributes and settlement geographic and locational characteristics. Appendix Figures D1–D4 give corresponding results looking at primary, (lower- and upper-) secondary, and tertiary education completion rates.

6.2.1 Origin. Language, Occupational Background, and Assets

As discussed above (in Section 2), refugees from Anatolia were far from a homogeneous group, coming from a diverse set of origins. Appendix Figure D5 shows the share of refugees by origin as recorded in the 1928 General Population Census and the Rural Refugee Catalog.

Decoding Narratives Besides broad historical narratives and a few case studies, there is no comprehensive record of the characteristics of Greek-Orthodox settlements in Anatolia at a granular level. To map conditions at the origin, we visited in the summer of 2019 the *Center of Asia Minor Studies* in Athens, read, and decoded 1,500 oral histories from their archive of approximately 5,000 interviews with refugees. The interviews provide a unique window into the characteristics of refugees' lives in Anatolia. We complemented these interviews with information from *Xenophanis*, a periodical publication of the Association of Asia Minor Greeks from 1896–1910, and a monograph on the Greek-Orthodox communities of Asia Minor by Anagnostopoulou (1997). We collected information on the main occupation, the primary agricultural products, and the language spoken. We recovered information for 250 out of the 843 villages in the Pontus-Caucasus region, 75 from the 103 in Cappadocia, 333 out of the 518 in Asia Minor, 19 out of the 38 in Constantinople, and 8 out of the 284 from East Thrace. Appendix Table D2 gives the tabulations, while Appendix Figure D6 maps the refugee settlements in Greece, distinguishing by origin.

Patterns and Differences Primary sector occupations (agriculture, fishing and herding) constitute the main source of income in more than three-fourths of villages. In Pontus, more than 80% of the villages relied on herding and agriculture for their livelihood, while the share in Coastal Asia Minor was 65%. The importance of intermediation and commerce among the Greek Orthodox population ranged from 10% in Pontus to 30% in Constantinople. Linguistically, Cappadocia stands out. In more than half of the villages, the population spoke only Turkish, and an additional 25% mainly Turkish. The remaining villages used a mixture of Cappadocian Greek, Pontic, and Turkish. In Thrace, Asia Minor, and Constantinople, most Greek Orthodox communities spoke exclusively Greek, while in Pontus, the linguistic landscape was a blend of Greek, Pontic, and Turkish. The statistics from the decoding of the interviews accord with historical evidence on the handicap of Cappadocian refugees in Greece due to their unique language. Refugees' displacement trajectories also varied. The Greek-Orthodox from Smyrna (Izmir), Ionia, and the coastal cities of Asia Minor had to flee in chaotic conditions after the swift advance of Turkish troops in September 1922. The Eastern Thrace Greeks abandoned their homes in a relatively organized manner, carrying movable property and animals, crossing the Evros River to Western Thrace. As part of the voluntary population exchange, the Greek-Orthodox communities from Bulgaria, unaffected by the Greco-Turkish wars, moved relatively peacefully south with their belongings. Pontic Greeks faced labor camp internments and large-scale massacres before moving to Greece in 1923.

Results We explore heterogeneity with respect to the origin, which we observe for families receiving land grants in 384 of the 405 predominantly refugee settlements.[24] In 147 refugee settlements, most of the families come from Pontus and Caucasus; in 94 from Eastern Thrace; in 75 from coastal Asia Minor; in 27 from Cappadocia, in one from Constantinople, and 7 from Bulgaria. In 33, there is no majority in the allocation of land grants (mixed settlements). The estimates show that, despite the heterogeneity of background and displacement trajectories, refugees from all areas have higher schooling than Greeks from native settlements in the same province. But there is one exception: Refugees from Cappadocia and Cilicia have a smaller—and not always statistically significant educational gain vis-à-vis those born in native settlements. As Cappadocians did not speak Greek but a unique Turkish dialect, these results reveal that even in a context where refugees and natives share religion and cultural heritage and where the refugees received legal rights upon settlement, their integration might be challenging if they do not speak the native language.[25] The cohort analysis (Appendix Figure D8) shows that the linguistic impediment for Cappadocian refugees is substantial for those born in the 1920s, 1930s, and 1940s and persists even for those born after WWII.

[24]Exploring heterogeneity across primary origin is of independent interest as most refugee offspring in contemporary Greece self-identify as Pontians, Cappadocians, Ioanians, Smyrniotes, and Constantinoupolite.

[25]The limited educational gains for settlements with Cappadocian and Cilician refugees do not stem from differences in their characteristics compared to other refugee settlements; see Appendix Figure D7. The Cappadocians faced discrimination from the natives, making their integration into the labor market challenging. This linguistic bias led to the abandonment of the language and the widespread belief that the language had disappeared by the 1960s. However, recent ethnographic studies reveal that the dialect persisted in hidden forms (Janse and Kavcic, 2020).

Example Three settlements in Giannitsa province, all within 25 km, illustrate the broader pattern. Stavrodromi, a native settlement with 299 inhabitants; Damianon, a 320-people refugee settlement inhabited by Cappadocians and Cilicians; and Arxontiko, a mixed refugee settlement of 307. In 1928, the share of the population born before 1917 with completed primary school was higher in the native settlement (17.4%) than in the refugee villages of Damianon (14.3%) and Arxontiko (12.5%). But, educational progress in the mixed refugee settlement outpaced that of the native settlement while gains in the Cappadocian settlement were modest. Among those born in 1917 and later, the high school completion rate in Arxontiko reached 26.1%, surpassing Stavrodomi, the autochthonous village (11.9%), and Damianon, the Cappadocian refugee settlement (10%).

6.2.2 Urban-Rural Background

Since the RSC catalog records refugees' exact origin (village, town, city, hamlet) in Anatolia, we can infer the share of the refugees in a settlement coming from an urban location. We classify as "urban" background refugees those from the 47 provincial capitals in Anatolia. The average (median) share of refugees from province capitals across the 384 predominantly refugee settlements was 4.4% (2.1%). In 116 settlements, none of the refugees came from a provincial capital, while in six settlements, more than half of the refugees came from province capitals (e.g., Drosia in Edessa, Sandali in Giannitsa, and Kouvouklia in Kozani).[26] Strong uprootedness effects are evident regardless of whether refugees were born in settlements with smaller or larger proportions of individuals from provincial capitals, highlighting the widespread desire of refugees to pursue education.[27]

6.2.3 Diversity

We then explore the role of refugee-settlement diversity, which is of independent interest and policy relevance. This investigation, therefore, connects to the literature on the effects of ethnic mixing, often originating from forced population movements and resettlement programs (e.g., Toews and Vezina (2025) and Chen and Zhou (2020)). Rather than looking at the mixing of natives and immigrants, we examine the impact of diversity within the refugee population, similarly to Bazzi et al. (2016). Using the rural refugee catalog, we construct standard fractionalization indexes for each refugee village in Greece (d), reflecting the likelihood that two randomly selected refugees will have a different origin (o): $Diversity_d = 1 - \sum_{o=1} a_o^2$, where a_o is the fraction of the settlement population (d) from origin o, defined at a granular level ($1,850$ settlements) and across 98 provinces. In the median refugee settlement, the probability of finding two individuals from the same village and province of origin was 30% and 50%, respectively. For example, among settlements in the Kilkis province, in Central

[26] If anything, the literacy rates and primary school completion in these six settlements for those born before 1902 were lower than those in the other refugee settlements, most likely because they engaged in peri-urban agriculture, common in late 19th and early 20th century Anatolia.

[27] Becker et al. (2020) find slightly stronger effects on education of displacement from urban areas in Poland. Sarvimäki, Uusitalo and Jäntti (2022) show that the income and education dividends of displacement are present only among Finns who are engaged in agriculture before resettlement.

Macedonia, Aspros and Kalindia have low diversity (less than 0.13) as the RSC largely managed to reconstruct their origin village; in contrast, nearby Kiladi, Koronouda, and Kritika have high fragmentation (0.88 and 0.63) reflecting the mosaic of origins.[28] Two results emerge from the various permutations. First, the uprootedness effects, with higher education for refugee-originating Greeks compared to autochthonous, are present when looking at both relatively homogeneous and diverse settlements. Second, the coefficients capturing the differences between refugee-native settlements in education are somewhat higher for more diverse refugee settlements, yielding some support to the idea that diversity of backgrounds, skills, and occupations may spur human capital investments.

6.2.4 New and Old Refugee Settlements

We distinguished between (the offspring of) refugees settling in new villages, built almost from scratch from the RSC, and former Muslim settlements (see Section 2). Whether one type of settlement is superior to another is unclear. New refugee settlements are closer to the coast, marshlands, and at lower altitudes than former Muslim villages.[29] Nevertheless, the latter had some infrastructure, and the farmland was of good quality and already cultivated.[30] Of the 405 predominantly refugee settlements, 291 (71.9%) had a considerable Muslim presence before the population exchange, while 114 (28.1%) were new settlements; see Appendix Figure A8. The estimates reveal significant uprootedness effects for those born after the population exchange in both settlement types.[31]

6.2.5 1928 Settlement-Level Population (Agglomeration)

Earlier works have identified significant scale effects from refugees (Peters (2022) and Ciccone and Nimczik (2023)). By increasing local market size, large influxes of refugees may initiate a chain of agglomeration dynamics, boosting development at the destination. Although these forces may be at play in the major urban centers, our comparisons involve small-to-medium neighboring villages and small towns of natives and refugees, where agglomeration effects are likely muted. However, it is still important to explore the educational gap when comparing refugee to native villages above and below the median population (285 residents). There is no evidence of heterogeneity, as refugees and their descendants have over-performed natives to a similar extent both in hamlets and small villages and in larger villages and small towns.

[28] Literacy and primary school completion for those born before 1902 were somewhat higher in refugee settlements with low diversity than in more diverse settlements. In the 1920s, education in native settlements was considerably higher than both in diverse and more homogeneous refugee settlements. In 2011, the high school completion rates in Macedonia and West Thrace were higher for both types of refugee settlements.

[29] These differences become less pronounced and appear statistically insignificant conditional on altitude. In addition, new settlements were somewhat more populated in and after 1928; see Appendix Figures D9-D10.

[30] As Morgenthau (1929) notes *"fortunately for the refugees, the Turkish landlords had been extremely inefficient. The result was large areas of virgin soil which, though it would require much heavy work to break it up for the first crop, was nevertheless available for new shelters."*

[31] Refugees' literacy and primary school completion rates in new settlements were somewhat higher than in former Muslim villages, but compared to native settlements, education was still much lower. Compared to natives, junior high completion rates in Macedonia and West Thrace in 2011 were higher in both types of refugee settlements.

6.2.6 Agroclimatic Aspects

Given the agricultural nature of refugees' origins and settlements in Greece's counryside, we explore in detail how agricultural conditions might have shaped their post-resettlement educational investments, using three metrics that conceptually speak to different underlying forces.

Farming Suitability across Destinations First, we use a 0-100 range measure that reflects the average crop suitability under rainfed conditions for 16 crops to approximate the farming attractiveness and potential productivity at the destination (barley, cassava, groundnut, maize, millet, oil palm, potato, rapeseed, paddy rice, rye, sorghum, soy, sugarcane, sunflower, summer wheat, and winter wheat). The 1928-population-weighted mean (median) value across the 405 predominantly refugee and the 6,091 native settlements are 27 (26.4) and 37.7 (37.4), respectively. The considerable autochthonous-refugee differences reflect broader regional differences and disappear when exploiting within-province variation (Appendix Figure A17). Allowing the coefficients to differ above and below the refugee settlement median yields no significant heterogeneity. In addition, the "direct effect" of high suitability for agriculture is statistically insignificant.

Geographic Similarity Between Origin and Destination Villages Second, motivated by the work of Michalopoulos (2012) and Bazzi et al. (2016) showing that the transferability of skills between regions facilitates mobility and shapes integration to the host communities, we explored heterogeneity with respect to the *geographic similarity* between destination and origin settlements in (i) agricultural suitability for crop cultivation (rainfed agriculture), (ii) elevation and ruggedness, (iii) precipitation and temperature, and (iv) distance to the coast. We calculated each characteristic's Gower destination-origin similarity index and took the average. The 0-1 range Gower index equals one minus the absolute difference between the characteristics of the origin and destination, normalized by the range. $Sim_d = 1 - \sum_{o=1} \frac{abs(x_d - x_o)}{x_{max} - x_{min}} a_o$, where d represents the destination, o the origin, and a_o is the fraction of the settlement population that came from the origin o. The 1928-population-weighted mean (and median) is high, 86.1, as origins and destinations had similar geographies. Looking across the broad macro-origins from Anatolia, those from Eastern Thrace who mainly settled in West Thrace and East Macedonia experienced the smallest changes in geographic conditions (Gower similarity of 0.91). Refugees from Cappadocia saw the largest changes, with a Gower similarity of 0.76, as the origin settlements were in the high-altitude Central Anatolian Plateau and far from the coast. The Gower similarity index for Pontians was 0.84, as the origin settlements had a mean altitude of 815 meters compared to 435 in the Pontic-majority villages in Greece, mainly in Macedonia.[32] The educational comparative advantage of refugees and their descendants is equally strong for settlements with relatively high and low geographic similarity with the origin.

[32]For example, the geography in Agora in Kilkis and Gramvousa in Nestos was very similar to refugees' origin settlements. In contrast, the refugee settlements of Aniphoria in Kilkis and Stegnon in Nestos had rather dissimilar elevation, temperature, and ruggedness to the settlers' origin villages.

Gains-Losses in Agricultural Quality Third, we distinguish between refugee settlements where the majority faced better or worse conditions than at the origin. On average, refugees settled in locations less suitable for rain-fed agriculture. The 1928 population-weighted mean (median) value of suitability gain is -7.4 (-8.8).[33] For example, Parameron in Drama and Agios Markos in Kilkis are among the settlements with the largest agriculture suitability gains, while Sterna in Drama and Aniphoria in Kilkis are among those with the biggest loss. Refugees from Asia Minor faced the most significant decline in soil suitability for agriculture, from an average of 43 to 29, while refugees from Pontus faced farming lands of similar quality. However, these differences do not appear to mediate or amplify the impact of forced displacement on education. Whether settling in relatively more or less fertile areas than their origins, refugees matched natives' education in the short run and surpassed them in the medium and long run.

6.2.7 Proximity to Province Capitals

We also distinguish between rural settlements close and far from provincial capitals to examine the role of proximity to high schools and other infrastructure. During 1929–32, more than $3,000$ primary schools were constructed, enabling for the first time access to education for the rural population. (At the time, there were fewer than 1,000 primary schools). While elementary schools were built in many medium-sized villages, attending junior high and high school required children to travel to—or stay for the weekdays in—provincial capitals and other towns. The 1928 population-weighted median distance to province capitals across the 405 predominantly refugee settlements and the 6,089 native settlements was 14.1 and 12.3 kilometers. The displacement-schooling nexus is substantial when comparing both proximate and more distant to province capitals (refugee and autochthonous) settlements; the effects are only slightly larger when looking at those born or residing in settlements relatively closer to provincial capitals .When we distinguish between primary, secondary, and tertiary educational completion (Appendix Figures D1–D4), some additional interesting patterns emerge. First, there is no evidence of heterogeneity for primary education, consistent with the fact that by the early to mid-1930s, most rural settlements had an elementary school. Second, heterogeneity emerges when looking at junior high and high school completion: Compared to natives, refugees were more likely to seize educational and other opportunities in the nearby provincial capitals.

6.2.8 Broader Province-Level Characteristics

A core dimension of Roy assignment models concerns the returns of education, which are higher in the "modern" sector (manufacturing and services) vis-à-vis agriculture. In our sample villages, hamlets, and small towns, non-agriculture employment was concentrated in the province capitals. As provinces proxy local labor markets, we explore whether the forced displacement-education nexus is stronger

[33] There is no evidence of spatial sorting concerning education. In Macedonia and West Trace, first-generation refugees born before 1902, settling in villages and towns with better-than-origin farming conditions, had similar literacy and primary school completion rates to refugees settling in areas with worse conditions.

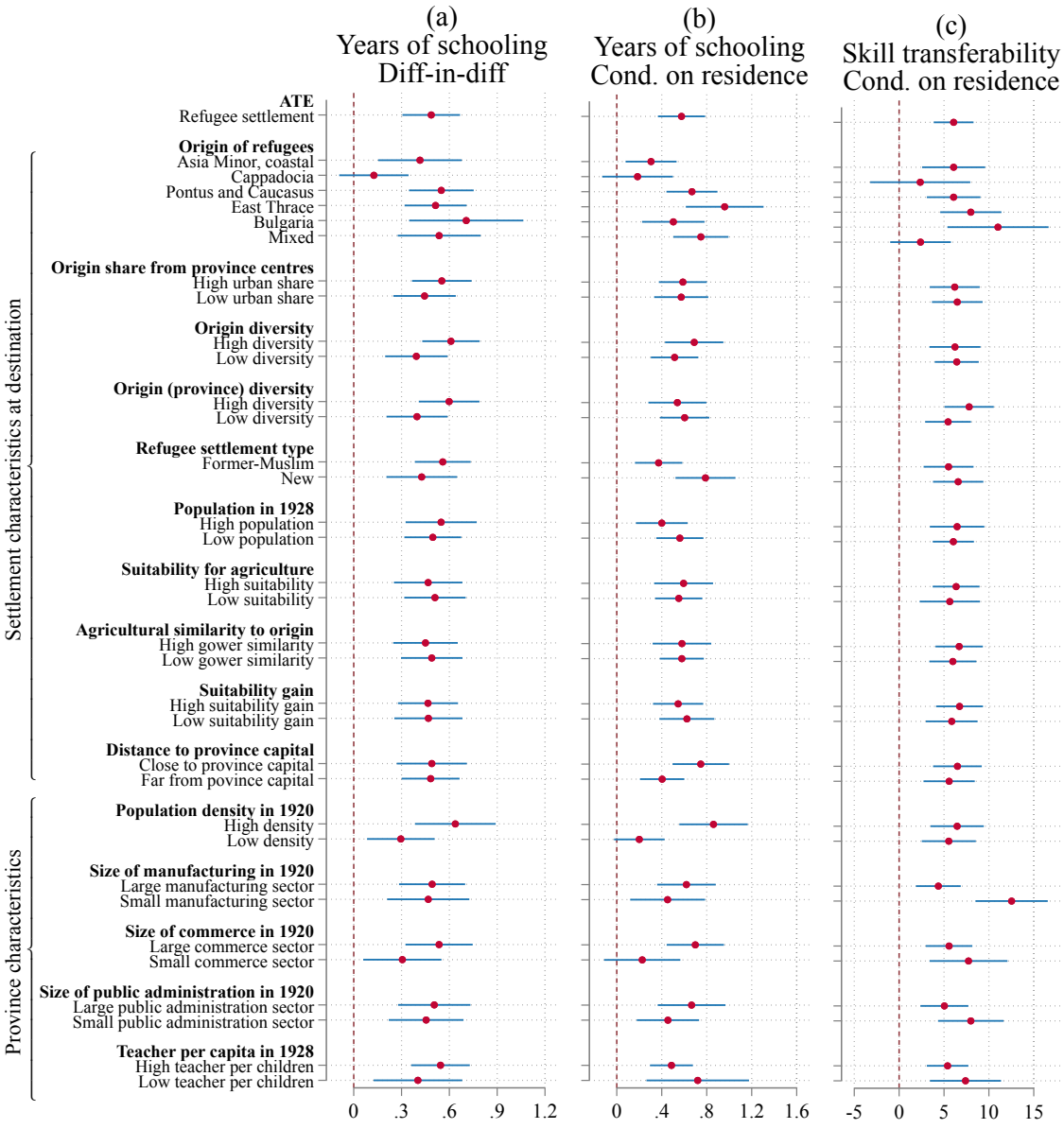
for refugees settling in more densely populated and less agriculture-oriented provinces. For example, even when looking across the 35 provinces of Macedonia and West Thrace, some refugee villages are in provinces where more than 25% of the labor force worked in manufacturing and commerce in 1920 (e.g., Komotini, Konitsa, Thassos, and Anasselitsa), while in other provinces the corresponding share was below 7% (e.g., Nestos and Kilkis). As the demand for an educated workforce would differ between these labor markets, we explore the role of province (local labor market) features related to the pre-population exchange level of development.

Pre-Displacement Labor Markets First, we interact the indicator for those reside in or coming from predominantly refugee settlements (REF_s) with three alternative measures of development in 1920: binary variables indicating whether provinces are either above the median population density, above the median employment share in manufacturing, or the median share of commerce in 1920 ("modern sectors"). The analysis reveals a significant displacement-education nexus in all sample splits. At the same time, the refugee-autochthonous gap is more pronounced in the above-median population density, manufacturing, and commerce provinces, as well as in 2011, when educational attainment is the highest.

Examining the completion of primary, secondary, and tertiary education sheds light on the underlying forces (Appendix Figures D1–D4). Refugees were more likely than natives to complete primary schooling, even in provinces with minimal demand for human capital. However, like natives, they were less inclined to advance beyond this educational milestone. In contrast, in provinces with significant economic activity outside of agriculture in 1920—and thus a higher demand for skilled labor—refugees climbed the educational ladder more rapidly, surpassing natives in secondary and tertiary education.

State Presence and School Supply Next, we examine whether the state presence and the school availability amplify the observed patterns. As a proxy for state presence, we take the share of the labor force employed in public administration in each province in 1920. State employment is minimal, with the median province at just 1.4%. However, there is substantial variation across provinces with refugee settlements, with the 90th-to-10th-percentile ratio being 5.4. For school availability, using the 1928 Census, we calculate the number of teachers per 6-to-14-year-old children across provinces. The variance in teachers per school-aged children across all provinces and those receiving refugees is limited, with a 90th-to-10th-percentile ratio of 1.5. The regressions reveal little heterogeneity in the displacement-education link in provinces with below or above-median public administration or teachers per pupil. The same pattern emerges with completing primary, junior high, and secondary school, suggesting that teacher and school availability differences in the initial years after the refugees' arrival do not account for their distinct educational paths.

Figure 7: Effects of refugees on years of schooling by settlement characteristics (1971–2011)



Notes. The panels present the heterogeneous treatment effects by settlement characteristics at the destination and settlement province characteristics. Panels A and B consider years of schooling as the outcome variable, while Panel C considers skill portability of tertiary degrees. Panel A presents the heterogeneity of difference-in-difference estimates presented in Table 1, column (1); Panels B and C presents the heterogeneity of estimates conditional on residence presented in Table 2, column (2), and in Table 3 column (2), respectively. For each characteristic we consider, we split the sample into two by the population-weighted median value of the characteristic at the settlement level in 1928; see Panel B of Appendix Table D1 to see the cutoff values. The confidence intervals in Panel A are based on standard errors clustered at the settlement-of-residence level, while in Panels B and C are based on standard errors double clustered at the birthplace-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

6.3 Heterogeneity in Skill Portability

Despite the limited sample size, we also explored how the link between refugee background and investment in *transferable* university degrees differs along the dimensions mentioned above. Figure 7–Panel C gives the estimates conditional on current municipality of residence; regression equation (2) estimated across college graduates in 2011. First, all estimates across the various subsets are significantly positive, uncovering how widely applicable the portability aspect of uprootedness is. Second, although province-level economic conditions are significant mediators for the native-refugee gap for the *quantity* of schooling, this is not the case for the *type* of university degree. What is perhaps most striking is that in the provinces with below-median manufacturing intensity in 1920, i.e., in the regions where engineering degrees were not in high demand, refugees and their descendants invested about 13% percentage points more often in portable degrees than natives. This gap is smaller but still sizeable in the above-median manufacturing provinces.

6.4 Taking Stock

Besides demonstrating the pervasiveness of the uprootedness effects across various sample splits, the heterogeneity analysis uncovers additional insights, which may apply in other contexts. First, linguistic barriers can impede the integration of refugees, even with religious and ethnic backgrounds similar to natives. Second, while the strong link between initial displacements and subsequent human capital investments applies to refugees from all backgrounds and types of destination settlement, it is muted at higher levels of education for refugees and their offspring who settle in rural areas where the broader labor market opportunities and demand for human capital are negligible. Third, the absence of heterogeneity in college degrees hints at a deeper valuation of human capital's portability, as it is unrelated to settlement, province, or origin characteristics.

7 Discussion

About a century ago, the international community, represented by the newly-established League of Nations, experienced one of its first humanitarian crises as millions of predominantly Christian Armenians, Greeks, and Assyrians, facing persecution and ethnic cleansing, fled the disintegrating Ottoman Empire. Following a unique-at-the-time Convention for a Forced Population Exchange, about 1.2-1.5 million Greek Orthodox abandoned their ancestral homelands in Anatolia and settled in Greece, raising its population by 20%. At the same time, more than 400,000 Muslims followed the reverse route from Macedonia and Crete to Turkey. Refugees who had lost children, family, and land arrived destitute, sick, and hopeless in Greece under the chaotic conditions that followed the burning of Smyrna and the atrocities in Pontus (Black Sea). The so-called *Asia Minor Catastrophe* has profoundly shaped modern Greece, leaving a lasting imprint on the society, literature, poetry, art, music, and sports.

Against the plethora of narratives, case studies, monographs, literary output, songs, and oral histories, which have preserved memory, we quantify the impact of this event, compiling a complete and

granular mapping of refugees' settlements in Greece alongside their origins. Inspired by the uprootedness hypothesis, which posits that the forcibly displaced will invest in portable human capital to integrate in the new country and be ready to move, we explore the displacement's dynamic impact on human capital, linking all Greek censuses of the past century. We focus on the universe of refugees settling in the Greek countryside, about half of the total, who received land grants alongside home construction inputs, technical assistance, tools, and animals, and compare them to Greeks living in nearby autochthonous settlements. The dynamic analysis, which spans over 100 years, reveals three results. First, refugees, particularly those growing up in the violent years leading to the population exchange, had lower literacy and primary school completion rates upon arrival than natives. Nevertheless, they caught up quickly, and their offspring surpassed the autochthonous' education. In 2011, people from refugee settlements had, on average, half to one year more schooling than those from nearby native settlements. This difference, present for women and men, applies to those living in the initial settlements and elsewhere, mainly in the big cities. Second, we document the portability aspect of the uprootedness hypothesis, which, despite its popularity, has been hard to quantify. Focusing on individuals with a university education, we show that refugees' descendants are more likely to pursue degrees with transferable skills, such as medicine, sciences, and engineering, while Greeks from rural native villages and towns obtain college degrees with substantial home bias, such as law. Third, we explore additional mechanisms using the richness of the newly-constructed data. Uprootedness effects are present in dozens of sample splits. Nonetheless, the heterogeneity analysis uncovers the importance of linguistic barriers, even for refugees with ethnic and religious backgrounds similar to the host population. Besides, the forced-displacement education nexus is stronger for refugees settling in more densely populated, manufacturing- and professional-services-oriented provinces, especially for secondary and tertiary education investments.

So, where do these results leave us? And what are the policy lessons? When the refugees arrived in Greece and the international community read and saw the pictures of their plight, few would have expected that they would catalyze the country's reconstruction and that these unfortunate individuals would overtake the native population in terms of education. Hope is, therefore, a message coming from our analysis, especially since the uprootedness hypothesis and a couple of other works on displacement yielding similar results allow for some generalization. However, hope is conditional. A hundred years ago, the League of Nations played an instrumental role by setting up the Refugee Settlement Commission and allowing it to raise funds for the refugees' support. Refugees got political rights and citizenship upon arrival. Besides, foreign (philanthropic) aid was chief, as the Near East Relief Fund, the very first large-scale humanitarian project, would support refugees in the initial months when the Greek state had no capacity. Lastly, despite considerable discrimination in the 1920s, the Greek governments genuinely tried to integrate refugees. So, as the Eastern Mediterranean and the Near/Middle East currently face large refugee outflows, our study's lesson from history is that the forcibly displaced, eager to learn and invest in education, can serve as an engine of growth. Yet, this depends on giving them a fair chance to do so.

References

Abramitzky, R., and L. Boustan (2017): "Immigration in American Economic History," *Journal of Economic Literature*, 55(4), 1311–1345.

Anagnostopoulou, S. (1997): *Greek Orthodox Communities in Asia Minor - 19th century - 1919*. Ellenika Grammata, Athens.

Arbatli, C. E., and G. Gokmen (2023): "Human Capital Transfers and Sub-national Development: Armenian and Greek Legacy in Post-expulsion Turkey," *Journal of Economic Growth*, 28(1), 1–43.

Ayal, E. B., and B. Chiswick (1983): "The Economics of the Diaspora, Revisited," *Economic Development and Cultural Change*, 31, 861–875.

Bauer, T. K., S. Braun, and M. Kvasnicka (2013): "The Economic Integration of Forced Migrants: Evidence for Post-War Germany," *Economic Journal*, 123(571), 998–1024.

Bauer, T. K., and M. Kvasnicka (2014): "Immigration and Structural Change: Evidence from Post-War Germany," *Journal of International Economics*, 93(2), 998–1024.

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

Becker, S. O. (2022): "Forced Displacement in History: Some Recent Research," *Australian Economic History Review*, 62, 2–25.

Becker, S. O., and A. Ferrara (2019): "Consequences of Forced Migration: A Survey of Recent Findings," *Labour Economics*, 59, 1–16.

Becker, S. O., I. Grosfeld, P. Grosjean, N. Voigtländer, and E. Zhuravskaya (2020): "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review*, 110(5), 1430–63.

Beine, M., V. Fromentin, and J. S. Bachiller (2024): "Emigration Prospects and Educational Choices: Evidence from the Lorraine-Luxembourg Corridor," Cesifo working paper no 11158.

Bester, C. A., T. G. Conley, and C. B. Hansen (2011): "Inference with Dependent Data Using Cluster Covariance Estimators," *Journal of econometrics*, 165(2), 137–151.

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

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

Bryan, G., and M. Morten (2019): "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 127(5), 2229–2268.

Card, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 19(1), 22–64.

Chen, Yi, Z. F. X. G., and L.-A. Zhou (2020): "Arrival of Young Talent: The Send-Down Movement and Rural Education in China," *American Economic Review*, 110(11), 3393–3430.

Chiovelli, G., S. Michalopoulos, E. Papaioannou, and S. Sequeira (2023): "Civil Conflict, Forced Displacement Trajectories, and Human Capital," *NBER Working Paper 29589*.

Ciccone, A., and J. Nimczik (2023): "The Long-Run Effects of Immigration: Evidence Across a Barrier to Refugee Settlement," *University of Mannheim Working Paper*.

Clark, B. (2006): *Twice a Stranger. The Mass Expulsions that Forged Modern Greece and Turkey*. Harvard University Press, Cambridge, MA.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019): "Inference with Arbitrary Clustering," *IZA DP No. 12584*.

Conley, T. G. (1999): "GMM estimation with cross-sectional dependence," *Journal of Econometrics*, 92(1), 1–45.

Dustmann, C., U. Schönberg, and J. Stuhler (2017): "Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment," *The Quarterly Journal of Economics*, 132(1), 435–483.

Eddy, C. B. (1931): *Greece and the Greek Refugees*. G. Allen and Unwin, New York.

European Union (2022): "Farms and Farmland in the European Union - Statistics," Discussion paper, Eurostat.

Fick, S. E., and R. J. Hijmans (2017): "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas," *International Journal of Climatology*, 37(12), 4302–4315, The WorldClim data are available here: https://www.worldclim.org/data/worldclim21.html (accessed May 25, 2020).

Grosser, T. (2006): *Die Integration der Heimatvertriebenen in Württemberg-Baden (1945 - 1961)*, Publications of the Commission for Historical Regional Studies in Baden-Württemberg. W. Kohlhammer Verlag, Stuttgart.

Güvenç, S. (2010): *Atlas of Old and New Toponyms of Northern Greece: Before and After the Population Exchange [Mübadele Öncesi ve Sonrası Eski ve Yeni Adları ile Kuzey Yunanistan Yer Adlari Atlası]*. Golden Medya Matbaacılık, Istanbul.

Hatzopoulos, K. (1991): "Ελληνικά σχολεία στην περίοδο της οθωμανικής κυριαρχίας 1453-1821 [The Greek School during Ottoman Rule 1453-1821]," *Thessaloniki: Vanias Publishers*.

Hill, M. (1931): "The League of Nations and the Work of Refugee Settlement and Financial Reconstruction in Greece, 1922—1930," *Weltwirtschaftliches Archiv*, 34, 265–283.

Hirschon, R. (2003): *Greece and the Inter-War Economic Crisis*. Berghan, New York and Oxford.

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

Howland, C. P. (1926): "Greece and Her Refugees," *Foreign Affairs*, 4.

Janse, M., and J. Kavcic (2020): "Cappadocian Greek: A Sociocultural and Linguistic History," *Keria: Studia Latina Et Graeca*, 22, 13–33.

Japan Aerospace Exploration Agency (2015): "ALOS World 3D - 30m resolution," The elevation data (AW3D30) are available here: https://www.eorc.jaxa.jp/ALOS/en/aw3d30/data/index.htm (accessed June 20, 2020).

Kardamatis, I. (1924): *Statistical Tables of Marsh Regions and Frequency of Malaria in Greece*. National Printing House, Athens, Greece.

Katz, Y. (1992): "Transfer of Population as a Solution to International Disputes: Population Exchanges between Greece and Turkey as a Model for Plans to Solve the Jewish-Arab Dispute in Palestine during the 1930s," *Political Geography*, 11(1), 55–72.

Kazamias, A. (1991): "The Education of the Greeks in the Ottoman Empire, 1856-1923: A Case Study of Controlled Toleration," *Schooling, Educational Policy and Ethnic Identity*, 1, 343–367.

Kessel, R. A. (1958): "Price Discrimination in Medicine," *The Journal of Law and Economics*, 1, 20–53.

Kitromilidis, P. (2008): "The Greek-Turksih Population Exchange," in *Turkey in the Twentieth Century, Edited by Erik-Jan Zurcher*. Klaus Schwarz Verlag, Berline.

Kontogiorgi, E. (2006): *Population exchange in Greek Macedonia the rural settlement of refugees 1922-1930*. Oxford University Press, Oxford.

Kritikos, G. (2000): "State policy and urban employment of refugees: The Greek case (1923–30)," *European Review of History: Revue européenne d'histoire*, 7(2), 189–206.

Kritikos, G. (2005): "The agricultural settlement of refugees: A source of productive work and stability in Greece, 1923-1930," *Agricultural History*, pp. 321–346.

Ladas, S. P. (1932): *The Exchange of Minorities: Bulgaria, Greece and Turkey*. Macmillan, New York.

Lagakos, D. (2020): "Urban-Rural Gaps in the Developing World: Does Internal Migration Offer Opportunities?," *Journal of Economic Perspectives*, 34(3), 174–192.

Lagakos, D., A. M. Mobarak, and M. E. Waugh (2023): "The Welfare Effects of Encouraging Rural-Urban Migration," *Econometrica*, 91, 803–837.

Logothetis, V., S. Michalopoulos, and E. Papaioannou (2025): "Songs of Refugees," *working paper, LBS, Brown, and University of Ioannina*.

Masters, B. (2009): "Millet," in *Encyclopedia of the Ottoman Empire*, ed. by G. Ágoston, and B. Masters. Facts On File, New York.

Mavrogordatos, G. T. (1983): *Stillborn Republic: Social Coalitions and Party Strategies in Greece, 1922-1936*. University of California Press, Berkeley, Los Angeles, London.

——— (2017): *The Prolongation of the Schism*. Patakis, Athens.

Mazower, M. (1991): *Greece and the Inter-War Economic Crisis*. Clarendon Press, Oxford.

——— (2009): "Minorities and the League of Nations in Interwar Europe," *Dedalus*, 126(2), 47–63.

Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4), 1508–1539.

Michalopoulos, S., E. Murard, E. Papaioannou, and S. O. Sakalli (2025): "Once a Refugee, Always a Refugee," Discussion paper, London Business School, mimeo.

Morgenthau, H. S. (1929): *I Was Sent to Athens (in collaboration with French Strother)*. Doubleday, Doran, and Co, Garden City, New York.

Nakamura, E., J. Sigurdsson, and J. Steinsson (2022): "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *The Review of Economic Studies*, 89(3), 1557–1792.

Nansen, F. (1923): "Report of the High Commission for Refugees," Discussion paper, League of Nations, Geneva.

Palis, A. A. (1929): "The Greek Census of 1928," *The Geographical Journal*, 73(6), 543–548.

Pavlowitch, S. K. (1999): *A History of the Balkans 1804-1945*. Routledge, London.

Pentzopoulos, D. (1962): *The Balkan Exchange of Minorities and Its Impact Upon Greece*. Mouton & Co, Paris.

Peters, M. (2022): "Market Size and Spatial Growth - Evidence from Germany's Post-War Population Expulsions," *Econometrica*, 90(5), 2357–2396.

Psomiades, H. (2009): "Fridtjof Nansen and the Greek refugee problem (September-November 1922)," *Center for Asia Minor Studies Bulletin*, 16, 287–346.

Sarvimäki, M., R. Uusitalo, and M. Jäntti (2022): "Habit Formation and the Misallocation of Labor: Evidence From Forced Migrations," *Journal of the European Economic Association*, 20(6), 2497–2539.

Schumann, A. (2014): "Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II," *American Economic Journal: Applied Economics*, 6(3), 189–205.

Sequeira, S., N. Nunn, and N. Qian (2020): "Immigrants and the Making of America," *Review of Economic Studies*, 87(2), 382–419.

Simpson, J. H. (1929): "The Work of the Greek Refugee Settlement Commission," *Journal of the Royal Institute of International Affairs*, 8(6), 583–604.

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

Toews, G., and P.-L. Vezina (2025): "Enemies of the People," *American Economic Journal-Macroeconomics*, 17(1), 310–342.

Young, A. (2013): "Inequality, the Urban-Rural Gap, and Migration," *Quarterly Journal of Economics*, 128(4), 1727–1785.

Zabel, F., B. Putzenlechner, and W. Mauser (2014): "Global Agricultural Land Resources – A High Resolution Suitability Evaluation and Its Perspectives until 2100 under Climate Change Conditions," *PLOS ONE*, 9(9), 1–12, The data are available here: https://zenodo.org/records/3748350 (accessed June 8, 2020).

Zervas, T. G. (2017): "From Ottoman Colonial Rule to Nation Statehood: Schooling and National Identity in the Early Greek School," *Espacio, Tiempo y Educación*, 4(1).

Online Appendix[34]

Uprootedness, Human Capital, and Skill Transferability

Stelios Michalopoulos
Elie Murard
Elias Papaioannou
Seyhun Orcan Sakalli

Abstract

The Online Appendix gives details on the data, descriptive and summary statistics, sensitivity checks, and further evidence. Section A gives mappings of the refugees' origin in Anatolia and settlement in Greece, summary statistics, and correlations on their settlement patterns. Section B reports additional descriptive and regression results associating forced displacement with refugees and their offspring's subsequent human capital investments. Section C gives additional results on the link between displacement and skill transferability of college degrees. Section D provides further descriptive and regression results, complementing the heterogeneity analysis and exploring additional displacement-education nexus mechanisms.

[34]Additional descriptives, examples, and details on the forced population exchange, the origin, and settlement of the Greek refugees from Anatolia can be found in https://anatolia-imprints.gr/. We thank Thanassis Stavrakoudis and his team of young researchers at the University of Ioannina for their help with the data. All errors are our responsibility.

A Descriptives and Summary Statistics

This Appendix complements the analysis in Section 3 of the paper. First, we provide a plethora of mappings and descriptions illustrating the settlement of refugees in Greece. Second, we explore differences between predominantly refugee and native rural settlements.

A.1 Descriptives and Mappings

Summary Statistics Appendix Table A1 gives summary statistics (mean; standard deviation; min, 10th percentile, median, 90th percentile, and max values) of the main variables (share of refugees, number of land grants, population in 1920 and 1928, share of Muslims in 1920) across 141 provinces (admin-3 level units) and 7,769 rural settlements.

Partial Refugee Census 1923 Appendix Figure A1 plots the share of refugees across 5,605 communities as recorded in the specialized (partial) Census in April 1923. Community (*koinotita* in Greek) is an administrative unit above the settlement level, which we use in our analysis. The specialized 1923 Census records 778,690 refugees, which is lower than the 1.2 million refugees recorded in the 1928 General Population Census, as the 1923 Refugee Census missed Greek-Orthodox refugees who arrived after the Population Exchange Convention and the Treaty of Lausanne.

Refugee Share across Provinces 1928 Appendix Figure A2 plots the share of refugees across 141 provinces, as recorded in the 1928 General Population Census. *Eparchies* (provinces), alongside *nomoi* (prefectures), have been the critical administrative units, coarse economic zones, from the establishment of the modern Greek state in the late 1820s till the 2000s. The boundaries of *eparchies* and the council's governing duties largely followed the *de facto* split during the Ottoman times, which in turn followed the church's diocese boundaries. Appendix Figure A3 plots the OLS association between the log number of refugees who arrived before and after 1922, as recorded in the 1928 General Population Census, across 141 provinces. The elasticity is 0.93, and the R^2 is 0.85.

Muslim Presence Appendix Figure A4 plots the share of Muslims (Bulgarians and Turks) who were forced to leave Greece after the 1923 Population Exchange Convention across 141 provinces, as recorded in the 1928 General Population Census. The Census suggests that 415,942 Turks and Bulgarians left. Appendix Figure A5 maps Muslim settlements and the share of Muslim populations across settlements. To impute the share of Muslims at the settlement level, we begin by computing, for each settlement, the change in its non-refugee population between 1920 and 1928. A fall in the population implies the out-migration of Muslims. Then, we scale up (or down) the settlement-level out-migration levels such that, once aggregated at the province (*eparhia*) level, it matches the province-level number of Muslims and Bulgarians left after 1920, as recorded in the 1928 census. Panel *A* plots the share of Muslims across Greece, where darker blue indicates a higher share. Panel

B zooms into Macedonia. In addition, the figure maps predominantly Muslim, Mixed, and Non-Muslim settlements, using information from Güvenç (2010) – the share of Muslims in 1915 is greater than 80% in "Predominantly Muslim" settlements; it is between 5% and 80% in "Mixed" settlements; and less than 5% in "Non-Muslim" settlements.

Ethnographic Map of Macedonia 1915 Appendix Figure A6 reproduces the digitized ethnographic map of Macedonia in 1915 (Güvenç, 2010). The map reports the main settlements alongside estimates of the Turkish, Bulgarian, and Greek-Orthodox populations.

Muslim and Refugee Settlements Appendix Figure A7 plots the density of the imputed share of Muslims departing after 1920 in Panel A and of the share of refugees in 1928 in Panel B across "Predominantly Muslim" villages (share of Muslims in 1915 equals or exceeds 80%), "Mixed" villages (share of Muslims in 1915 between 5% and 80%), and "Non-Muslim" villages (share of Muslims in 1915 equals or is less than 5%) according to the ethnographic map of Macedonia in 1915 (Güvenç, 2010). Panel A shows that the imputed share of departing Muslims proxies well the observed distribution of Muslims in 1915 across settlements in Macedonia—the imputed share of departing Muslims is skewed to the right among "Predominantly Muslim" settlements in 1915 and skewed to the left among "Non-Muslim" settlements in 1915. The correlation is high: 218 (80.7%) of 270 predominantly Muslim villages in 1915 are also predominantly Muslim in 1920 with our procedure. We split predominantly refugee settlements, where the share of refugees in 1928 exceeded 80%, into former Muslim settlements and new ones. Appendix Figure A8 plots the two types of refugee settlements across Greece (Panel A) and in Macedonia and West Thrace (Panel B), as four out of five rural refugees settled there.

Rural Refugee Catalog Examples Appendix Figure A9 illustrates the origin-destination structure of the rural refugee catalog, compiled by the Refugee Settlement Committee (RSC) in the mid-1920s. Panel A gives the settlements of origin in Anatolia of Greek-Orthodox refugees in Agios Constantinos, in Aetolia-Acarnania in West Greece. Most refugees come from Pontus, on the southern shore of the Black Sea, and from the Interior of Ionia on the West Coast of Asia Minor. Panel B maps the settlements of refugees from Panormos (Bandırma) in the Marmara Sea. Most Greek-Orthodox who left Panormos settled in West Thrace and East Macedonia.

Origin Settlements Appendix Figure A10–Panels A and B portray the Greek-Orthodox settlements in contemporary Turkey from where refugee families come from. We identify $1,850$ origin settlements in the rural refugee catalog, which we have processed in its entirety (Panel A). There are 923 settlements in the urban refugee catalog, which we have partly processed (about 43%). 923 settlements appear in both catalogs, 927 only in the rural catalog, and one only in the urban catalog. $1,735$ origin settlements are in modern Turkey; 115 are outside of Turkey (in contemporary Bulgaria, Ukraine, Russia, and Georgia).

Census Validation Appendix Figure A11 illustrates the correlation between the (log) number of refugees in the 1928 Population Census and the (log) number of families that received a land grant, as recorded in the Rural Refugee Catalog, compiled by the Refugee Settlement Committee in the mid-1920s. We focus on rural settlements, i.e., exclude settlements with greater than a population of 10,000 in 1928, province capitals, and settlements within 25 km of Athens, Piraeus, and Thessaloniki. Panel *A* gives the correlation across 1,284 rural settlements with at least one land grant. Panel *B* plots the correlation across 384 settlements with at least one land grant where the share of refugees in 1928 exceeds 80%. The correlation is strong. The number of families settled by the RSC with a land grant explain 88% of the variation in the presence of refugees in a settlement in 1928.

Baseline Sample Appendix Figure A12 maps the 405 predominantly-refugee villages that received at least one land grant, and where the refugees' population share exceeds 80% (in red), and the 6,091 native villages (in blue), where refugees' share is below 5%, and there was no recipient of a land grant. The figure also maps mixed refugee-native settlements (in low-tone gray). Panel *A* maps the country; Panel *B* maps Macedonia and West Thrace; Panel *C* zooms into the Giannitsa province in Central Macedonia, respectively.

Local Sample Appendix Figure A13 maps the local sample of predominantly-refugee and native settlements, where each native settlement is matched to the nearest predominantly-refugee settlement in the same province within 25 km. To perform this matching, first, we identify all pairs of refugee and native settlements within 25 km of each other in the same province (excluding provinces with no pairs). Second, for each native settlement, we identify the closest predominantly refugee settlement. Third, we split each province into Voronoi polygons based on the location of these refugee settlements. Fourth, we exclude refugee settlements further than 25 km from a native settlement within the same Voronoi polygon. Panel *A* maps the whole country, while Panel *B* zooms into Macedonia and West Thrace. Panel *C* presents the sub-province Voronoi polygons constructed for Giannitsa province in Central Macedonia, along with the names of native and predominantly refugee settlements nearest to them in the same province.

Age Distribution in 1928 Appendix Figure A14 presents the age distribution by gender and refugee status as provided by the 1928 General Population Census. The share of males aged 20 to 44 is largely affected by the Balkan Wars, the Great War, and the Greco-Turkish War. However, it is much more pronounced among refugees, particularly in the 20-24 age range, as the Greek population was subject to persecution in the Ottoman Empire during this long decade of conflict. In particular, young Greek men served in the Labor Battalions (*Amele Taburları*), where mortality rates were very high during WWI (Minasidis, 2024). Notably, the lower representation of children below the age of 9 among refugees in 1928, mainly born in Greece after the exodus, reflects the adverse economic conditions, poor health, and refugees' high mortality in their first years in Greece.

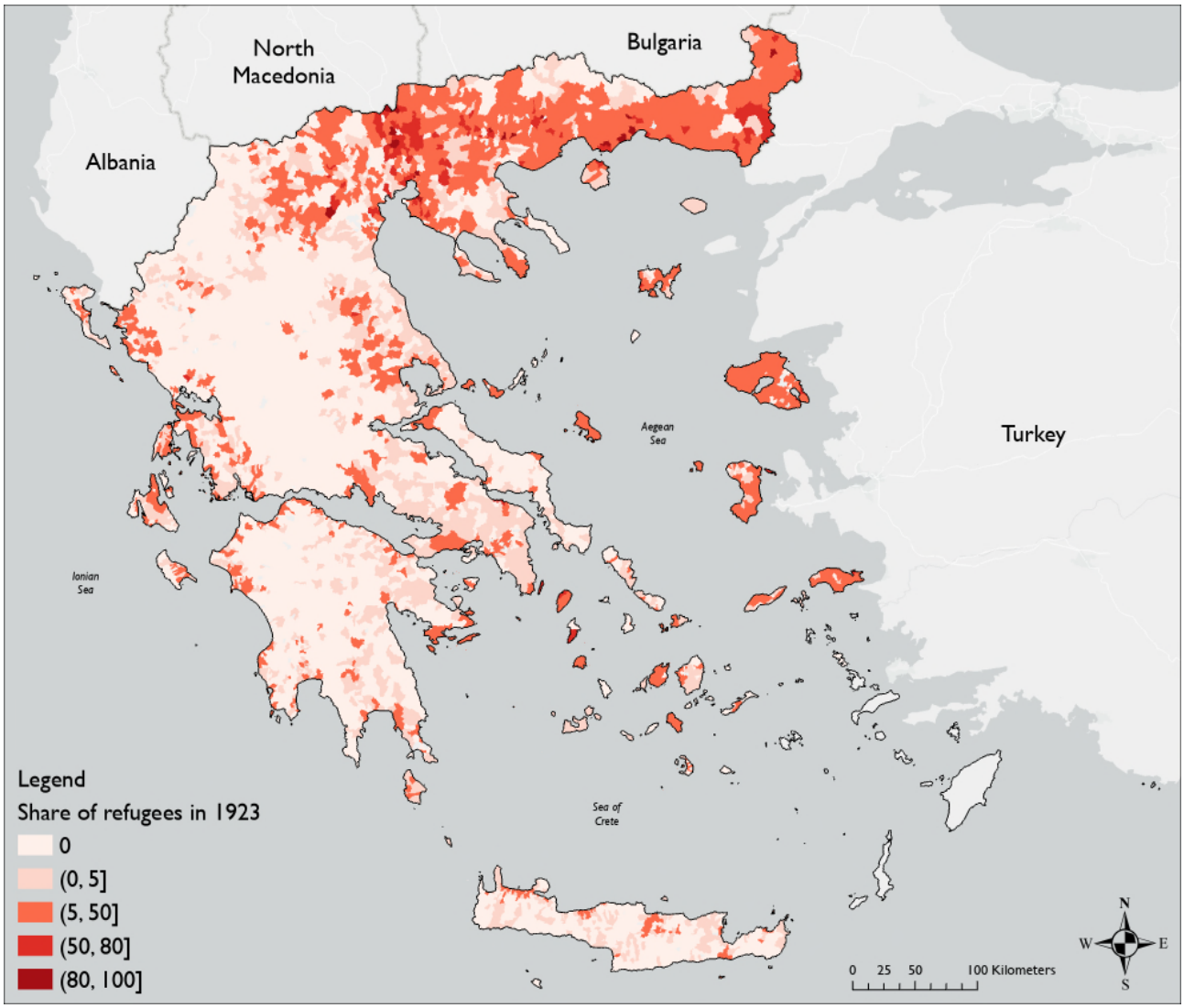
Population Dynamics Appendix Figure A15 plots the evolution of the population of predominantly refugee and native settlements from 1907 until 2011 in the baseline (Panel *A*) and the local sample (Panel *B*), using the information from the respective General Population Censuses. Three main takeaways emerge. First, our comparisons mainly entail small and mid-sized villages. Second, population differences between the two sets of settlements were minor at the time of the population exchange (as recorded in the 1928 General Population Census). Third, the population evolves similarly in both types of settlements. Despite the overall increase in Greece's population, the population in rural refugee and native settlements does not increase much.

Table A1: Summary Statistics across Provinces and Settlements

	Mean	SD	Min	10p	Median	90p	Max	Obs.
Panel A:				**Province characteristics**				
Population in 1928	43,970.4	78,119.6	3,037	10,342	30,992.0	73,403	866,924	141
Population in 1920	35,544.3	47,561.4	3,190	9,709	25,302.0	60,569	501,615	141
Share of rural population in 1928	86.3	21.3	14	57	100.0	100	100	141
Surface area	914.0	630.6	29	268	742.7	1,743	3,067	141
Share of refugees in 1928	14.0	20.2	0	0	3.5	46	90	141
Share of refugees in rural areas in 1928	13.4	21.3	0	0	2.4	46	95	141
Share of Turks and Bulgarians in 1920	8.4	18.0	0	0	0.1	33	91	141
Number of settlements with land grant	10.4	17.7	0	0	1.0	35	110	141
Number of refugee families settled with land grant	1,318.4	2,587.1	0	0	63.0	4,651	14,298	141
Share of refugees from Constantinople in 1928	0.2	0.4	0	0	0.1	0	3	141
Share of refugees from Asia Minor in 1928	5.6	7.4	0	0	2.0	16	35	141
Share of refugees from Pontus in 1928	2.7	6.6	0	0	0.1	9	49	141
Share of refugees from Thrace in 1928	3.6	7.8	0	0	0.1	16	38	141
Share of refugees from other regions in 1928	0.1	0.4	0	0	0.0	0	4	141
Panel B:				**Settlement characteristics**				
Population in 1928	485.4	565.5	0	84	329.0	1,010	10,438	7,769
Population in 1920	431.4	504.8	0	69	288.0	919	8,483	7,769
Share of refugees in 1928	10.0	24.4	0	0	0.0	44	100	7,769
Settlement with land grant (%)	17.5	38.0	0	0	0.0	100	100	7,769
Number of refugee families settled with land grant	19.9	68.6	0	0	0.0	61	1,123	7,769
Monastery in 1920 (%)	4.1	19.9	0	0	0.0	0	100	7,769
Log distance to reclaimed marshlands	2.7	1.2	0	1	2.9	4	5	7,769
Log distance to province capital	2.6	0.6	0	2	2.7	3	4	7,769
Log distance to coast	2.6	1.2	0	1	2.6	4	5	7,769
Log distance to railway stations	3.3	1.5	0	1	3.2	6	6	7,769
Log crop suitability	3.5	0.5	0	3	3.6	4	4	7,767
Log altitude	5.6	1.1	0	4	5.8	7	7	7,769
Log average precipitation	4.0	0.2	3	4	4.0	4	5	7,769
Log average temperature	2.7	0.1	2	3	2.8	3	3	7,769

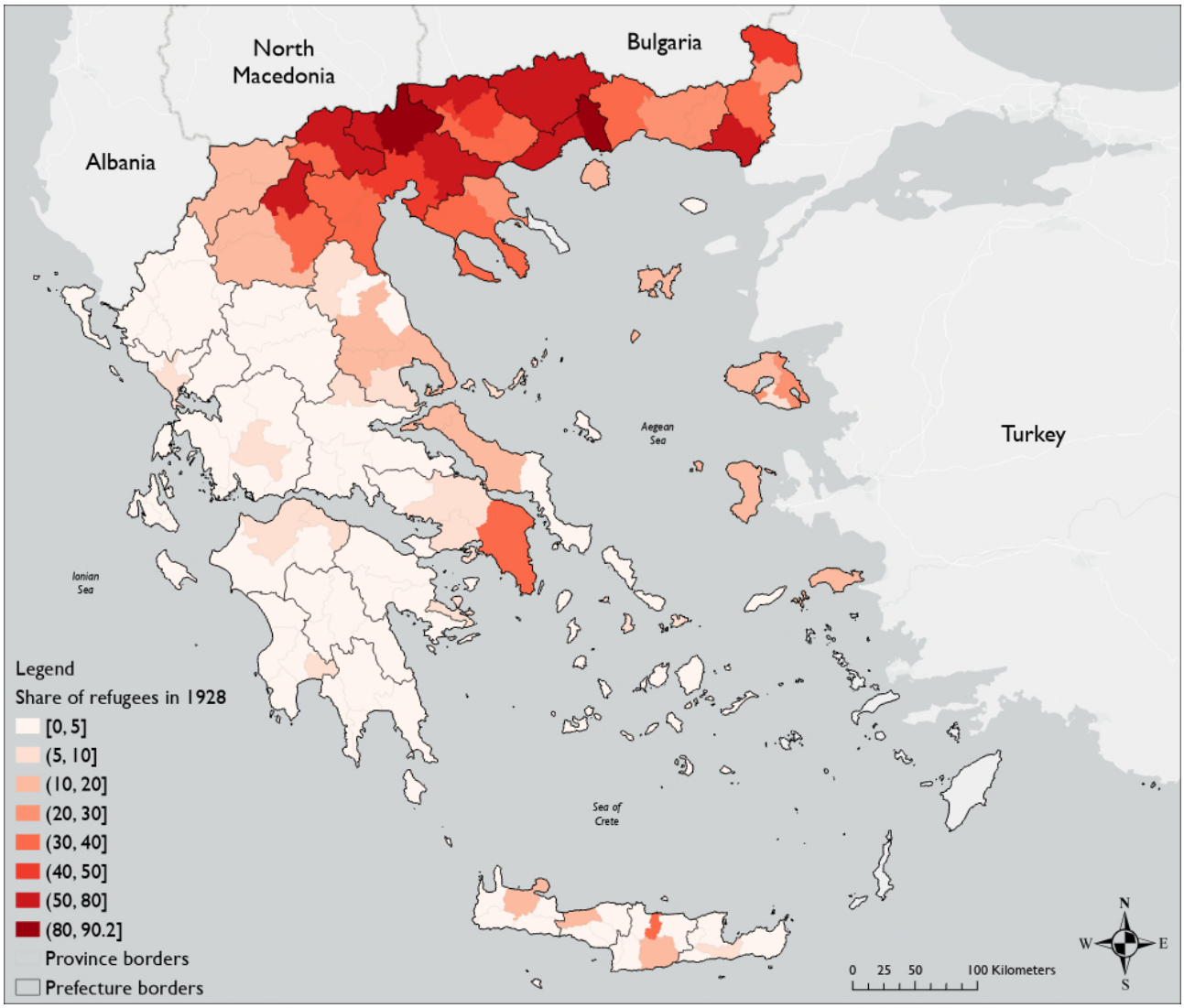
Notes. The table gives summary statistics of population and refugee shares across provinces and settlements, along with settlement geographical characteristics. Panel *A* gives statistics across 141 provinces (*eparchies*). Panel *B* gives statistics across all rural settlements, baseline sample, which excludes cities with a population exceeding 10,000 in 1928, province capitals, and settlements within 25 km distance to Athens, Piraeus, and Thessaloniki.

Figure A1: Refugee Share in 1923



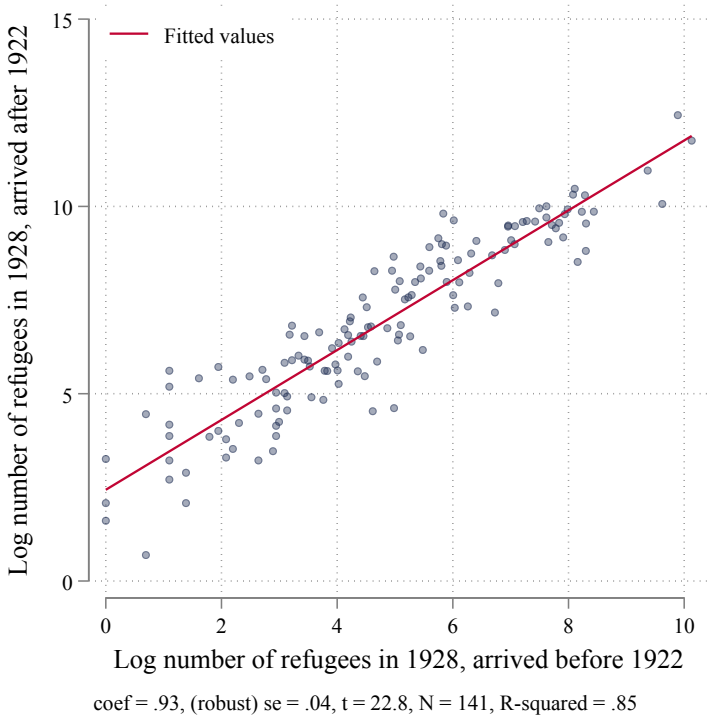
Notes. The figure maps the share of refugees in 1923 across 5,605 communities, as recorded in the specialized census of the Refugee Settlement Committee. (Community (*koinotita* in Greek) is an administrative unit above the settlement level, which we use in our analysis.) The partial Census includes 778,690 refugees.

Figure A2: Refugees in 1928 across Provinces (*Eparchies*)



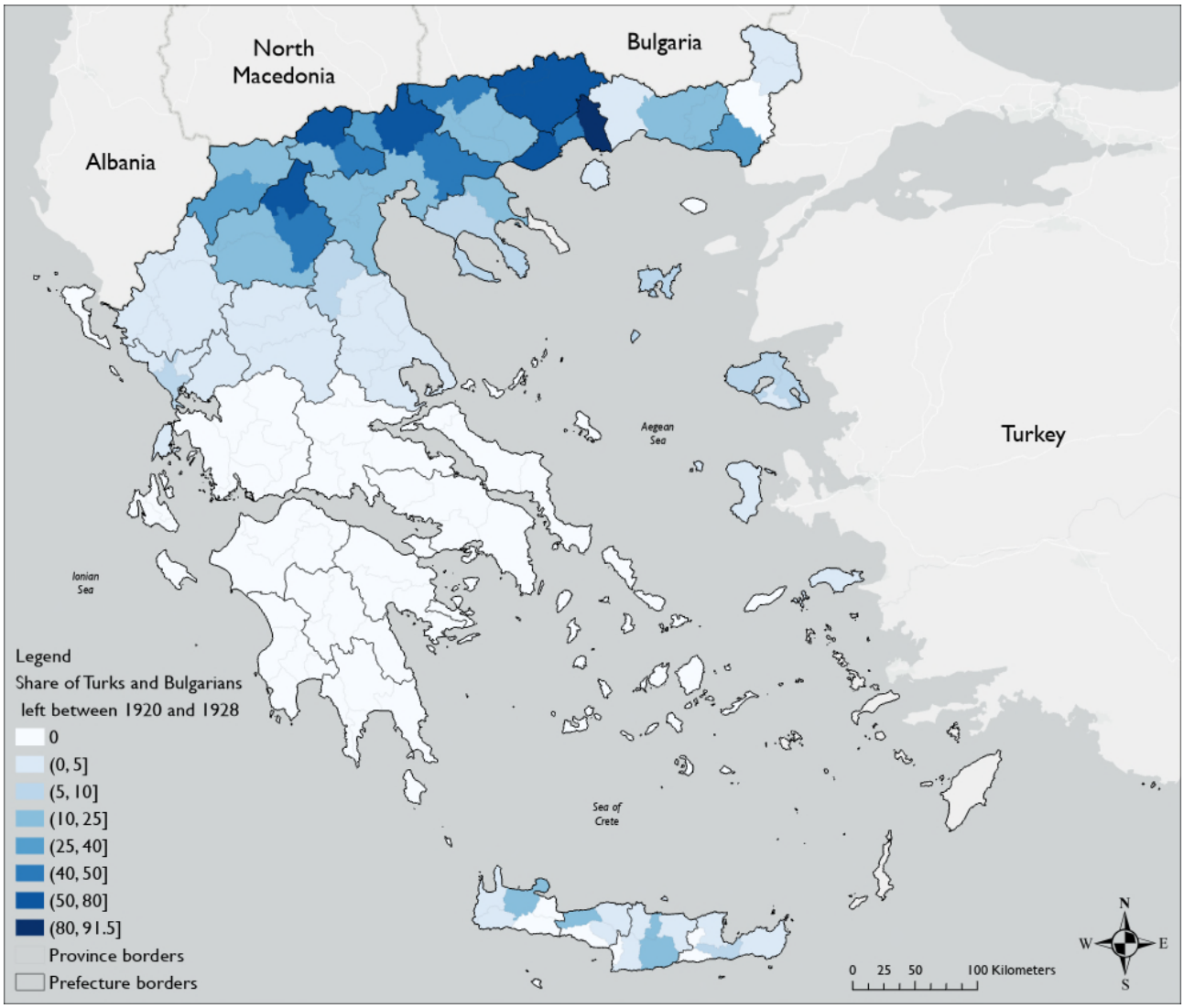
Notes. The figure maps the distribution of all refugees across 141 provinces (*eparchies*), excluding Mount Athos, as recorded in the 1928 General Population Census. There are 1,221,849 refugees in a total population of 6,204,684.

Figure A3: Correlation Refugees Arriving before and after the Asia Minor Catastrophe (1922–23)



Notes. The figure shows the OLS association between the log number of refugees arriving after 1922 (vertical axis) and before 1922 (horizontal axis) across 141 provinces (admin-3 level units). Data come from the 1928 General Population Census.

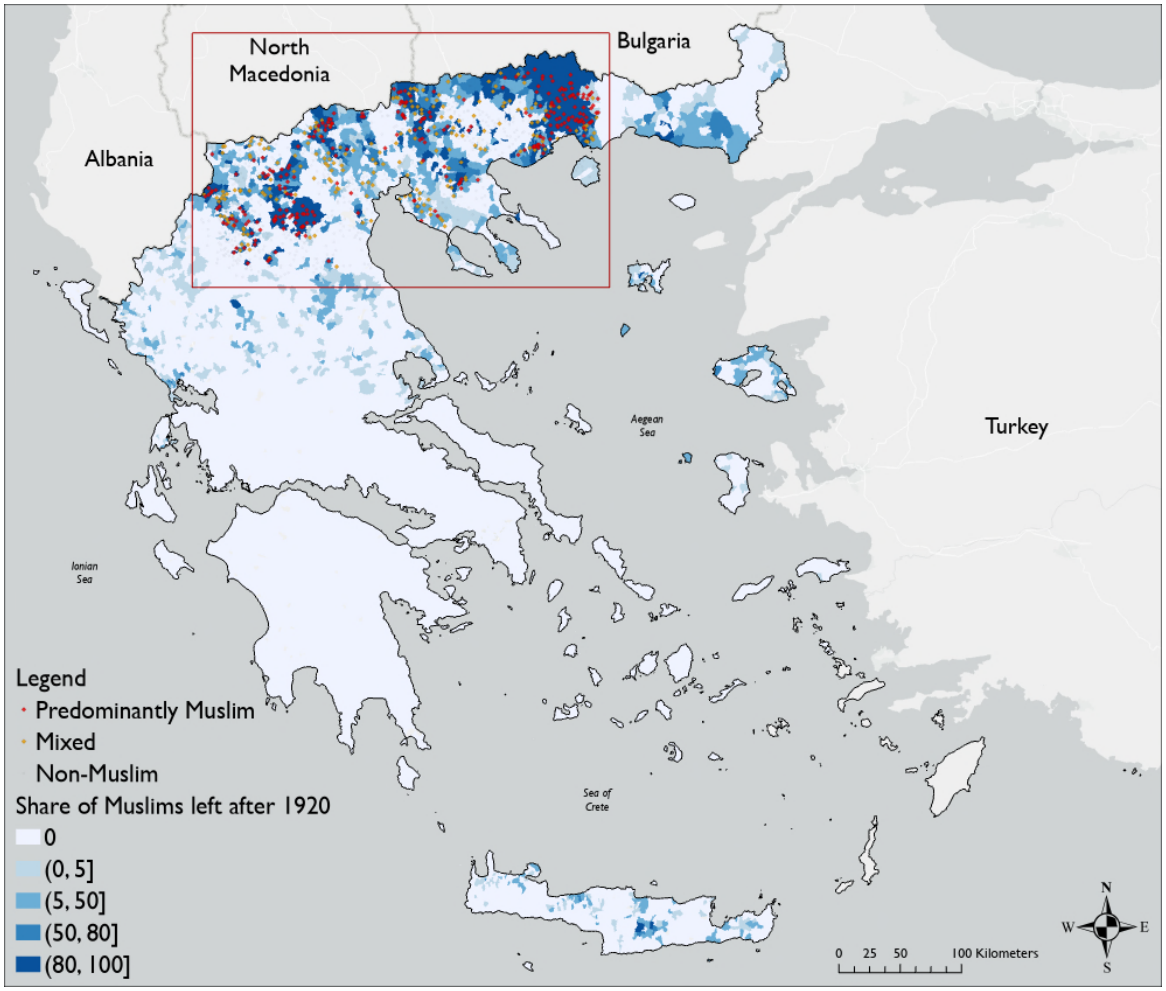
Figure A4: Muslim Population (Turks and Bulgarians) before the Population Exchange



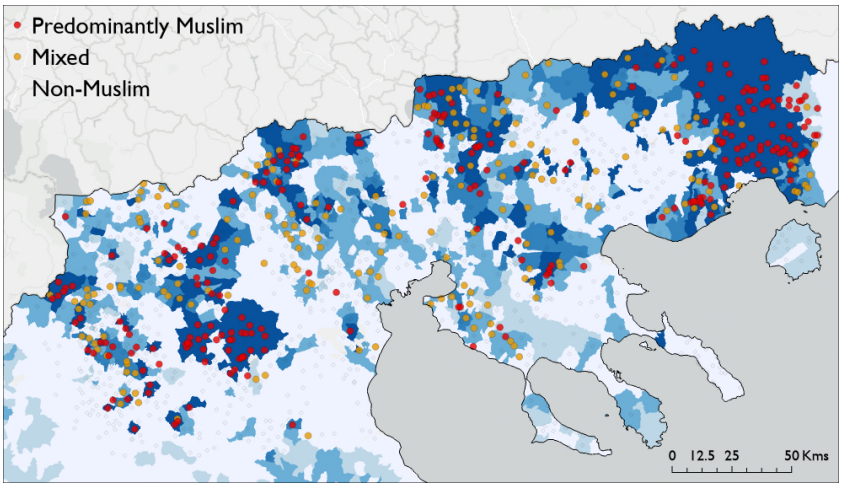
Notes. The figure maps the share of Turks and Bulgarians who left Greece between 1920 and 1928 across 141 provinces (*eparchies*), excluding Mount Athos, as recorded in the 1928 General Population Census. In total, 415,942 Turks and Bulgarians left Greece in this period.

Figure A5: Muslim Share in 1920 across Settlements

(a) All Greece

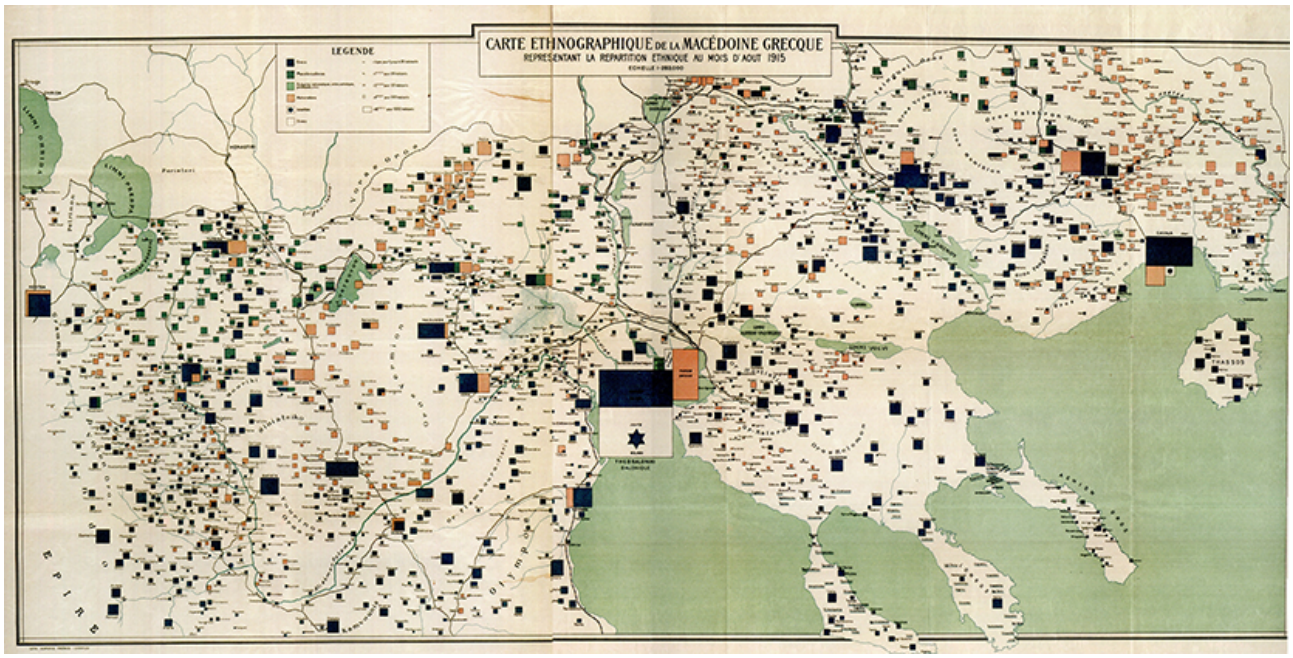


(b) Macedonia



Notes. The panels map the imputed share of Muslim (Turkish and Bulgarian) refugees who left Greece after 1920 across 5,605 communities in Greece in Panel *A*, and across communities in Macedonia in Panel *B*. (Community (*koinotita* in Greek) is an administrative unit above the settlement level, which we use in our analysis.) The red rectangle in Panel *A* marks the regions covered by the ethnographic map of 1915 (Güvenç (2010)).

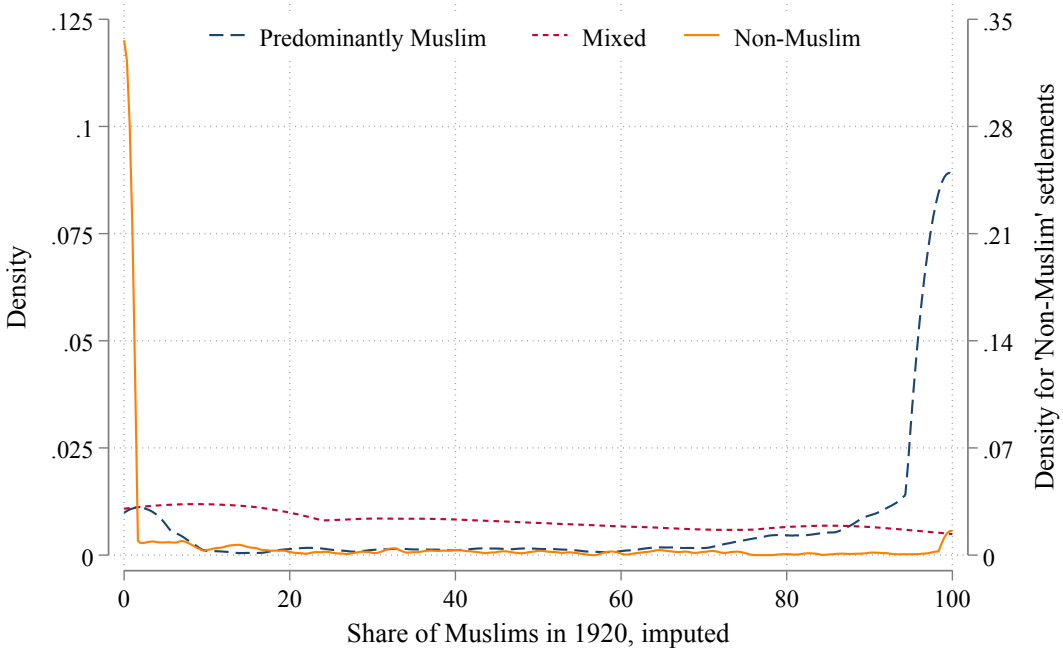
Figure A6: Ethnographic Map of Macedonia (1915)



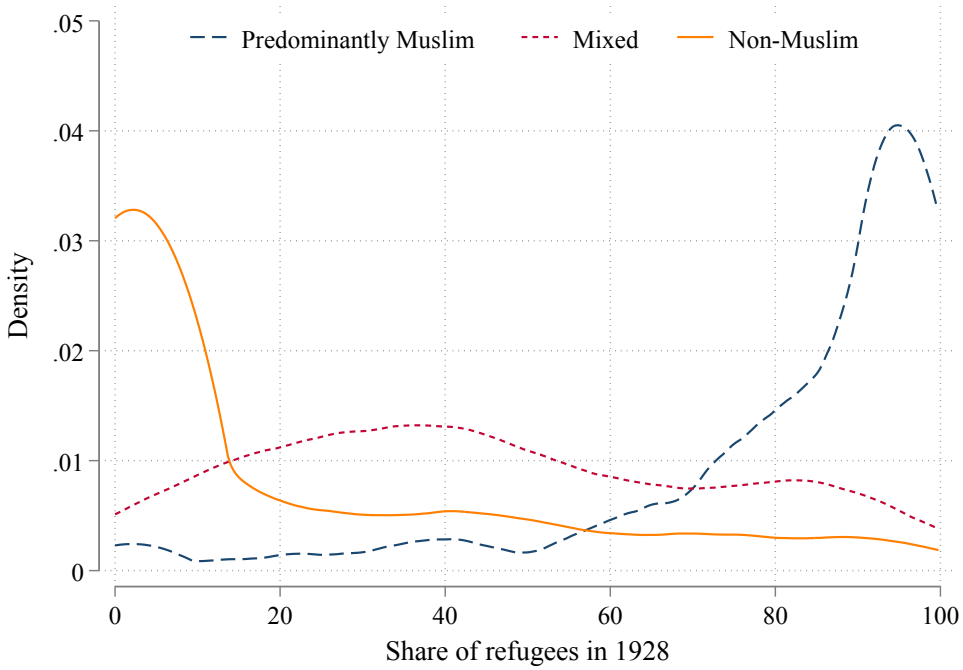
Notes. The figure presents the population distribution by ethno-religious groups at the settlement level in Macedonia in 1915. The title of the map reads: *"Carte ethnographique de la Macedoine Grecque: presentant la repartition ethnique au mois d'aout 1915"* [Ethnographic map of Greek Macedonia: presenting the ethnic distribution in August 1915]. The map gives ethnoreligious groups for localities in Drama, Florina, Halhidiki, Kavala, Kozani, Pella, Thessaloniki, and Serres prefectures and some localities in Elassona province of Larissa prefecture. Data source: Güvenç (2010).

Figure A7: Distribution of Greek-Orthodox and Muslim Refugees across Settlements

(a) Departing Muslims (after 1920)



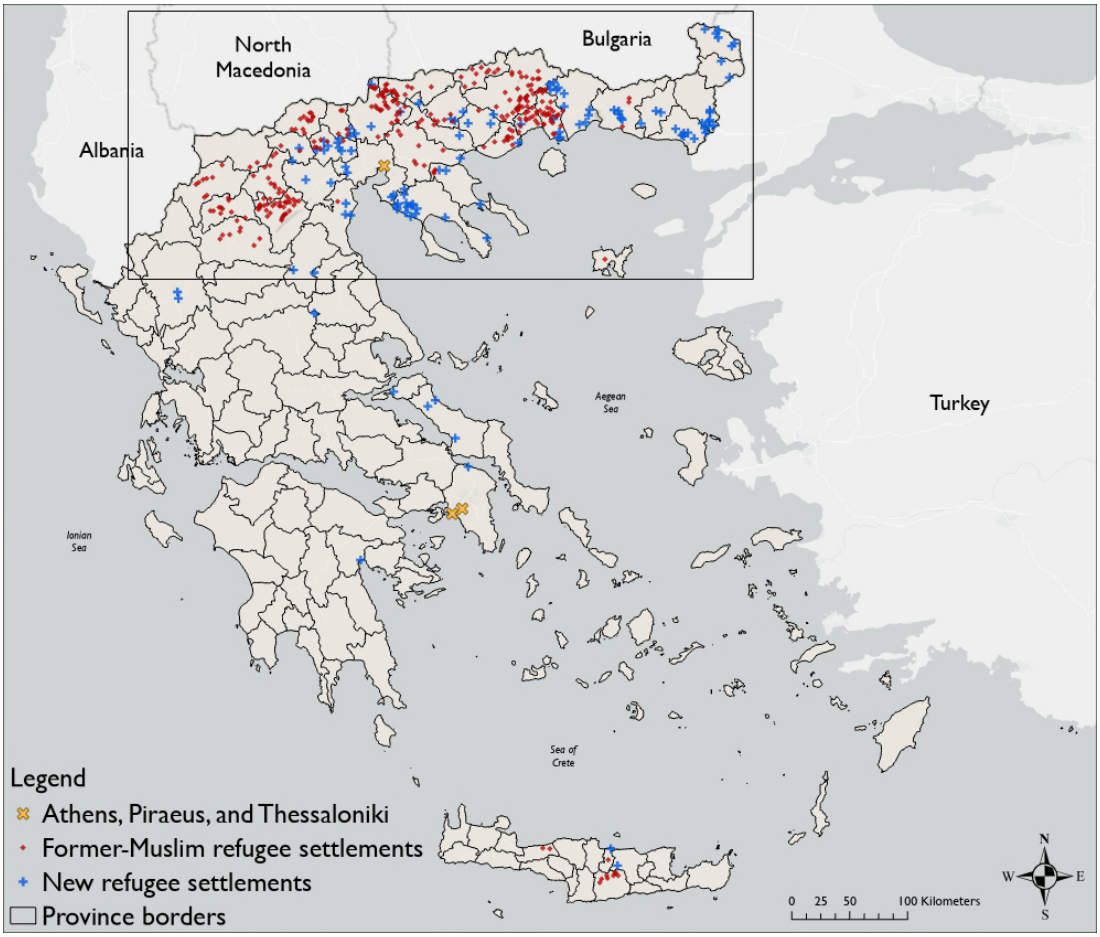
(b) Greek-Orthodox Refugees (in 1928)



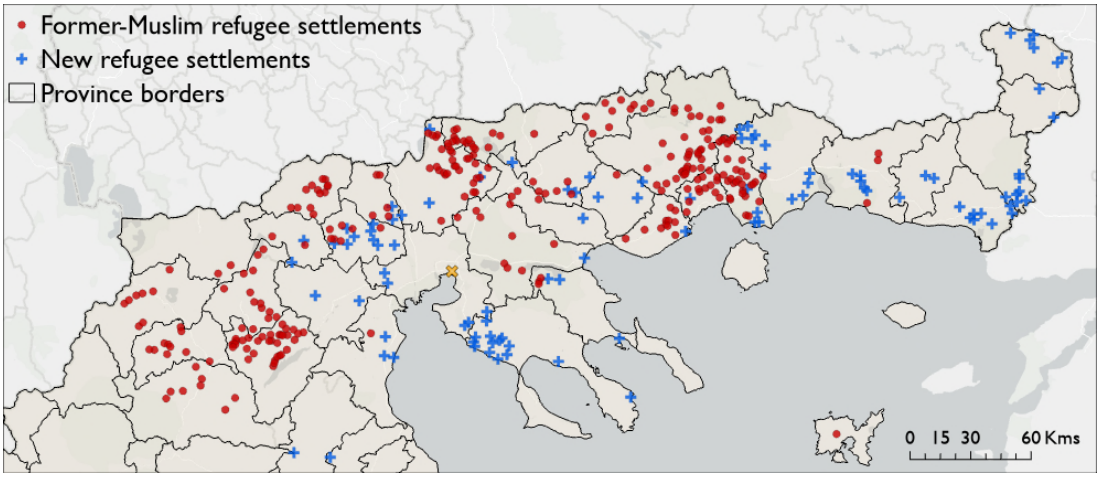
Notes. The figures plot the density of the imputed share of Muslims departing after 1920 in Panel *A* and the share of refugees in 1928 in panel *B* across "Predominantly Muslim" villages (share of Muslims in 1915 exceeds 80%), "Mixed" villages (share of Muslims in 1915 between 5% and 80%), and "Non-Muslim" villages (share of Muslims in 1915 less than 5%) according to the ethnographic map of Macedonia in 1915 (Güvenç (2010)).

Figure A8: Rural Refugee Settlements by Type in 1928

(a) Greece



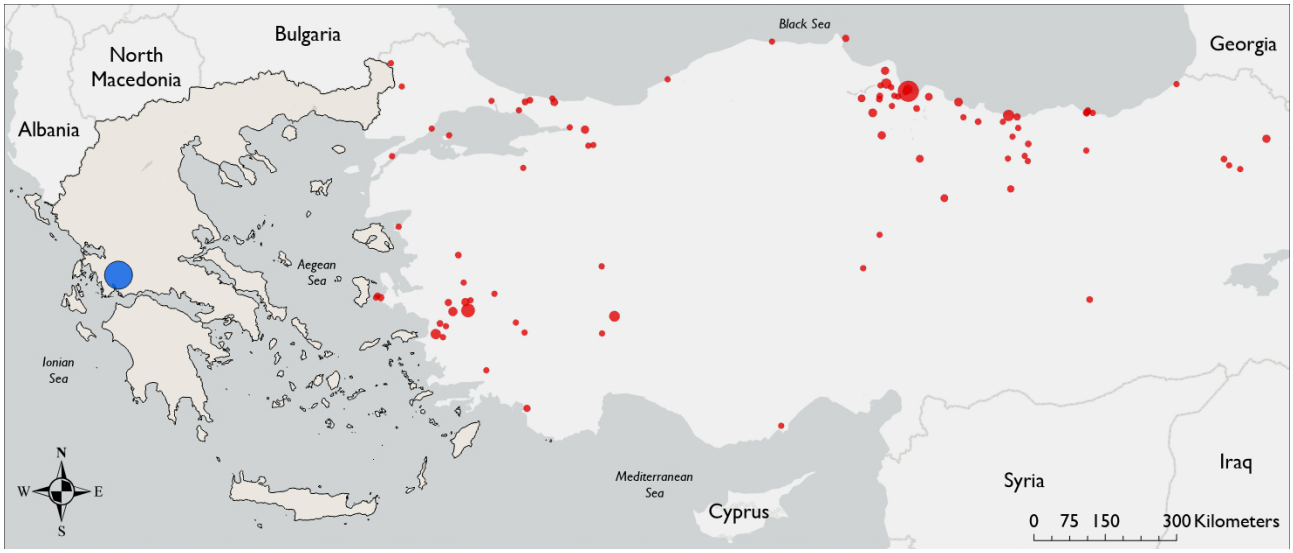
(b) Macedonia and West Trace



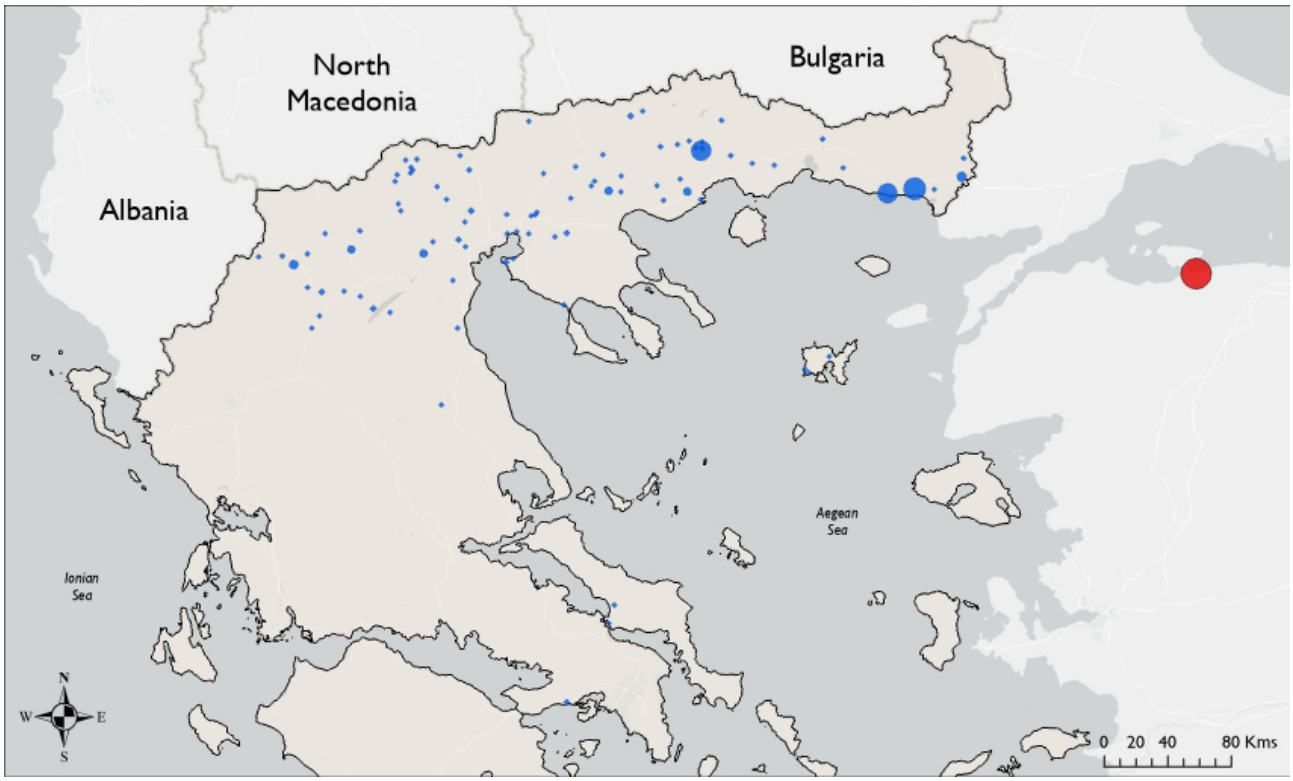
Notes. The panels map 405 predominantly-refugee rural settlements, where the share of refugees in 1928 exceeds 80% by settlement type: 291 former-Muslim refugee settlements, where the imputed share of Muslims in 1920 exceeds 80% (in red dots), and 114 new refugee settlements, where the imputed share of Muslims in 1920 is below 80% (in blue plus signs). Panel *A* plots settlements across Greece. Panel *B* zooms into Macedonia and West Thrace.

Figure A9: Rural Refugee Catalogue Examples: Settlement of Destination and Origin

(a) Origin of Refugees in Agios Konstantinos (Aetolia-Acarnania)



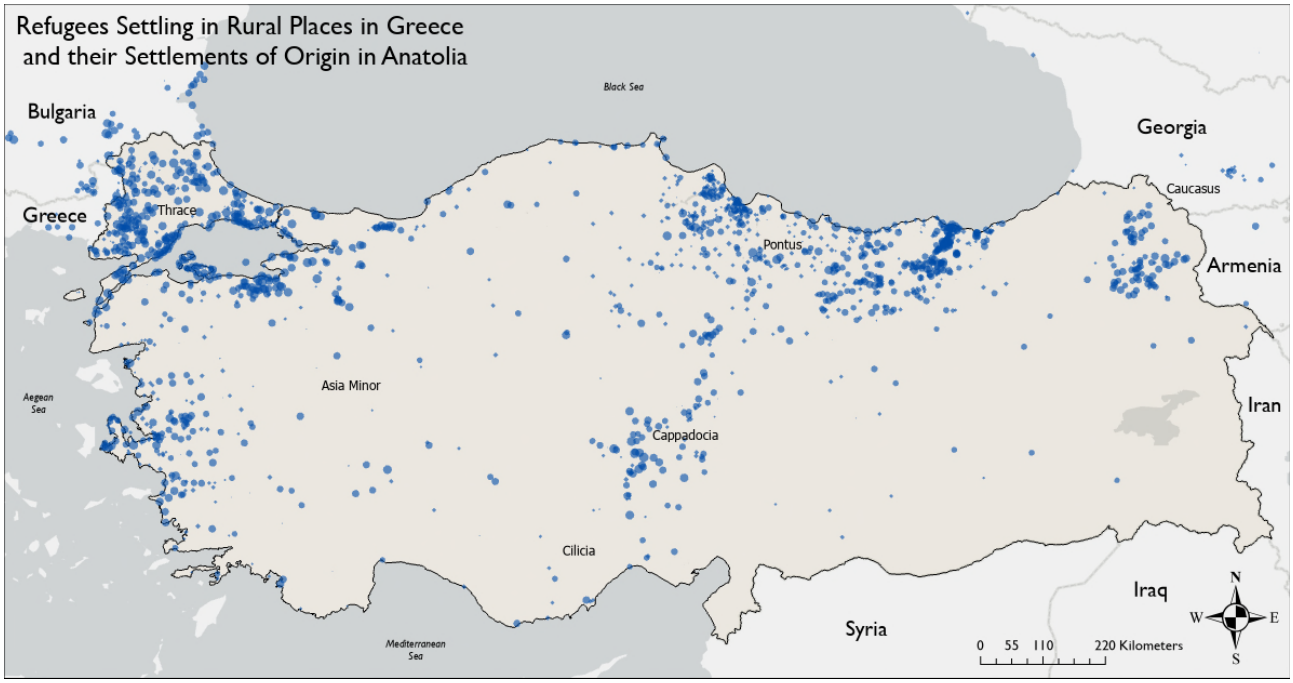
(b) Settlements of Refugees from Panormos (Bandırma, Balikesir)



Notes. The two panels illustrate the origin-destination information of the Rural Refugee Catalog compiled by RSC in the mid-1920s. Panel *A* gives the origin (in contemporary Turkey) of refugees settling in Agios Constantinos (Aetolia-Akarnania). Panel *B* gives the destination settlements of all refugees from Panormos (Band). The dot size presents the (log) number of refugee households from (in) that location.

Figure A10: Settlements of Greek-Orthodox Refugees in Turkey before the Population Exchange

(a) Rural Refugee Catalog

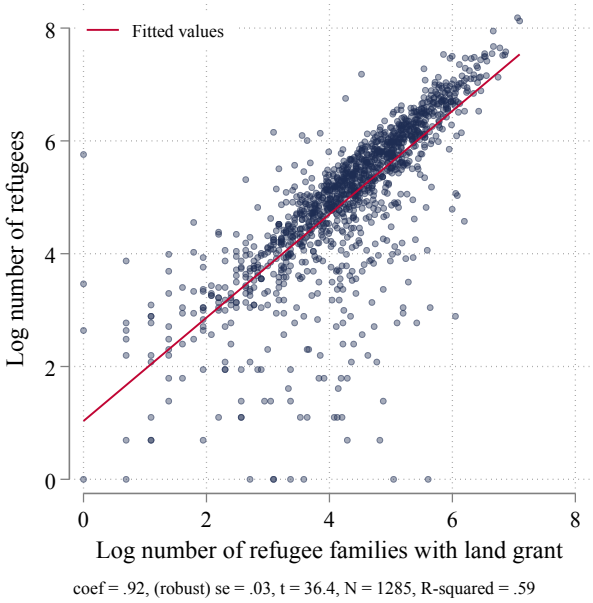


(b) Urban Refugee Catalog

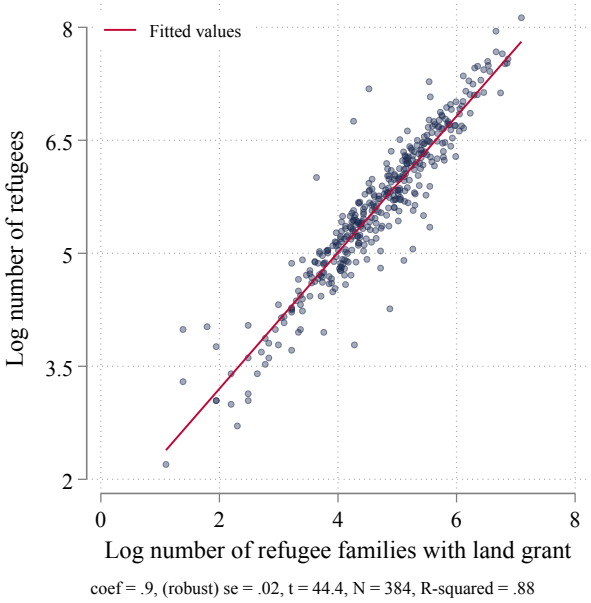
Notes. The figure plots the origin of all Greek refugees in the rural (agricultural) and urban refugee catalogs compiled by the Refugee Settlement Committee in the mid-1920s. The information is based on 211,056 household entries in the rural catalog across 1,850 settlements and 164,617 household entries in the urban catalog across 923 settlements). 1,735 origin settlements are in contemporary Turkey and 115 are outside Turkey. Blue (red) dots indicate the origin settlements of refugees primarily settling in rural (urban) areas in Greece in the mid-1920s. The dot size presents the (log) number of refugee households from that location.

Figure A11: Refugees in the 1928 Population Census and Refugee Families Receiving Land Grants in the Rural Refugee Catalog

(a) All rural settlements with land grant



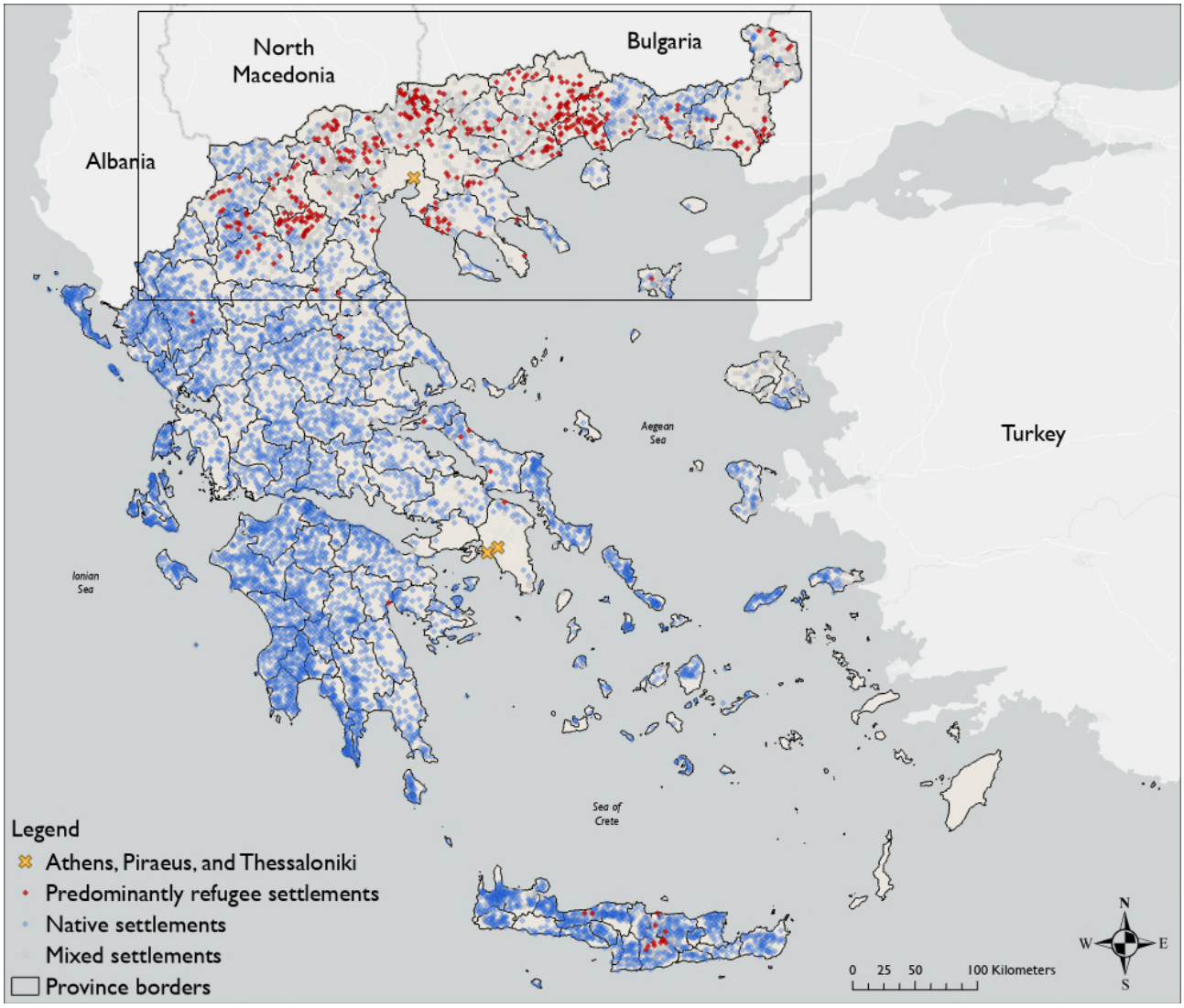
(b) Predominantly refugee settlements with land grant



Notes. The panels plot the OLS correlation between the log number of refugees based on the 1928 Population Census and the log number of refugee families that received a land grant based on the Rural Refugee Catalog for all rural settlements with at least one land grant (Panel *A*) and among rural predominantly refugee settlements, where the share of refugees in 1928 exceeds 80%, with a land grant (Panel *B*).

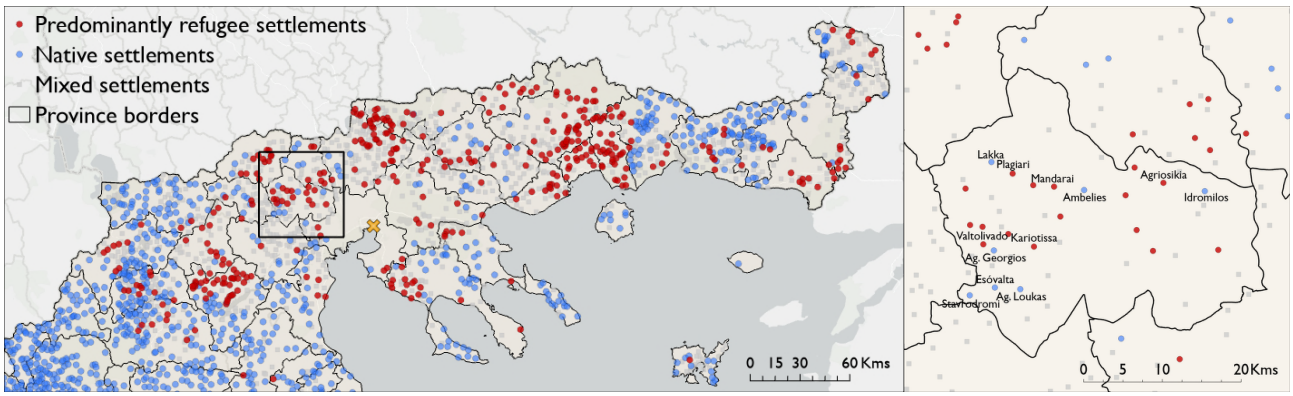
Figure A12: Rural Refugee and Native Settlements in 1928

(a) Greece



(b) Macedonia and West Trace

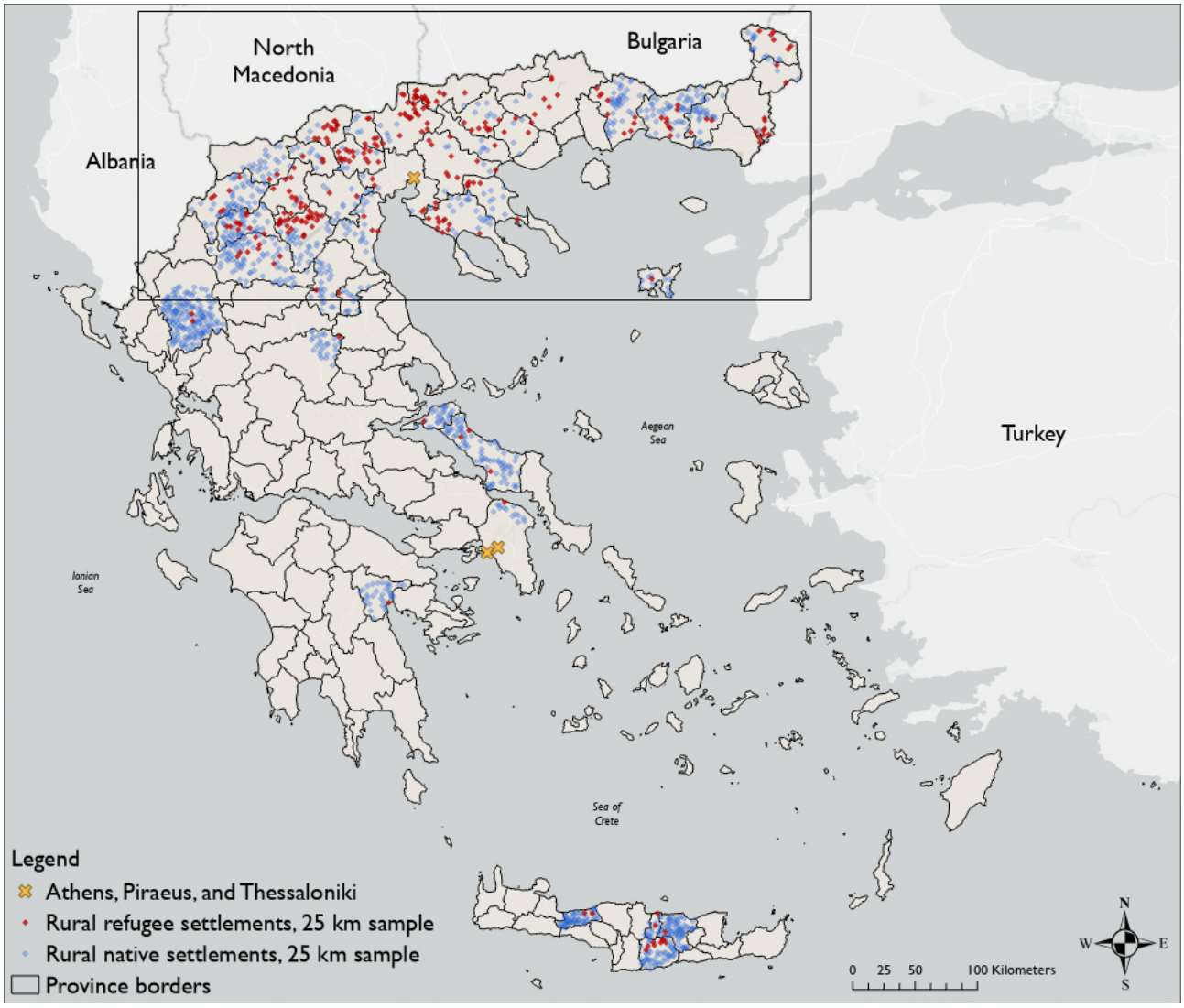
(c) Giannitsa province



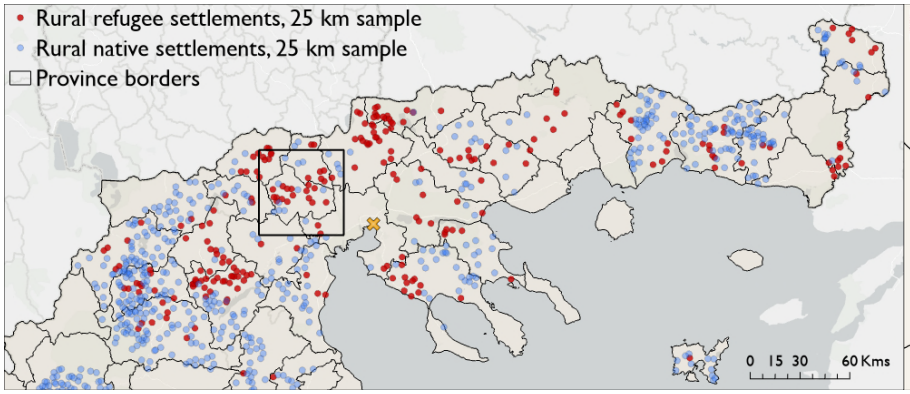
Notes. The panels map 7,769 rural settlements, distinguishing between predominantly-refugee settlements, where the share of refugees in 1928 exceeds 80% (in red dots), and native settlements, where the share of refugees is below 5% (in blue dots), and mixed settlements, where the share of refugees is between 5% and 80% (in gray squares). Panel *A* plots settlements across Greece. Panel *B* zooms into Macedonia and West Thrace. Panel *C* zooms into the area surrounding Giannitsa province in Pella, Central Macedonia.

Figure A13: Local Analysis Sample

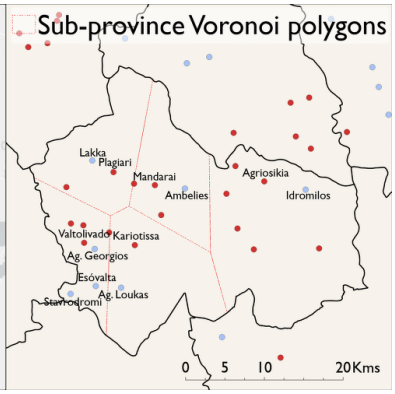
(a) Greece



(b) Macedonia and West Trace

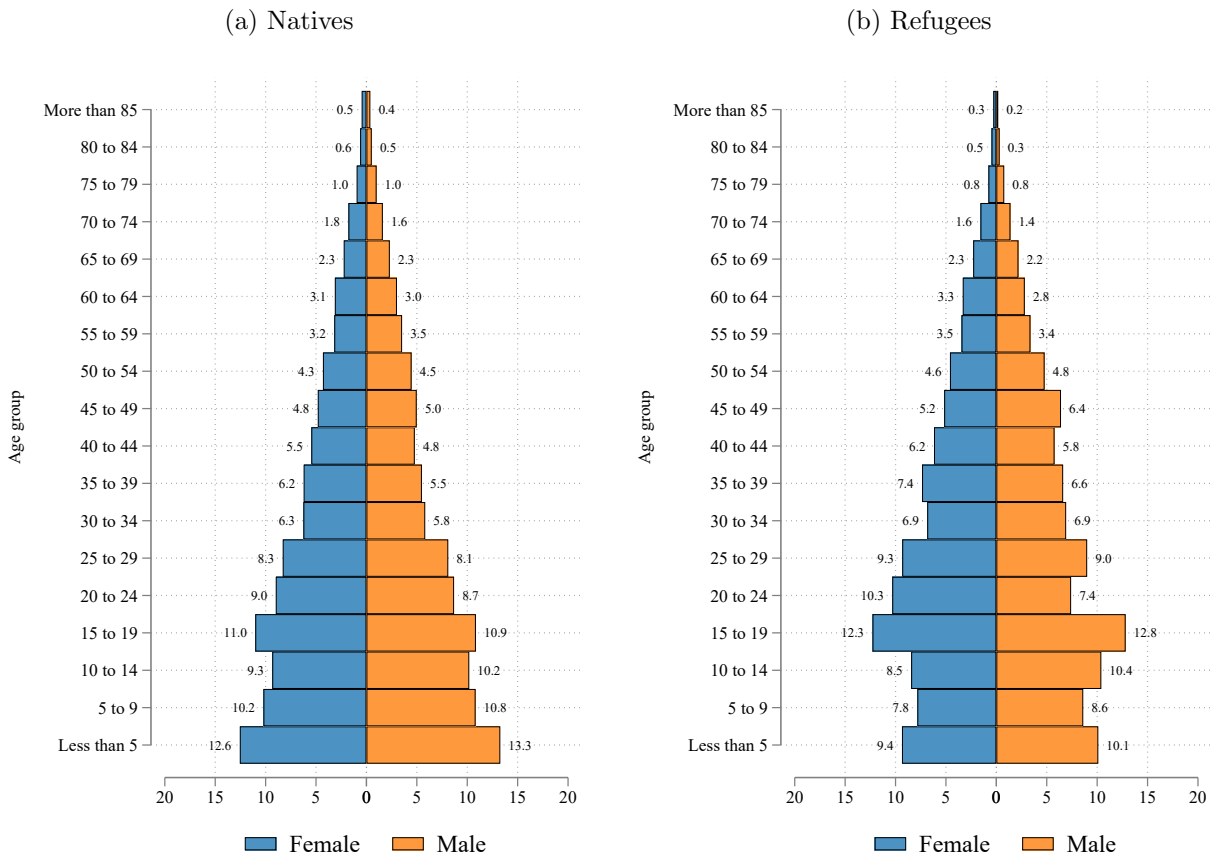


(c) Giannitsa province



Notes. The panels map the native and refugee settlements in the local sample. Each native settlement is matched to at least one nearby refugee settlement within a 25 km radius. In the refugee settlements, the share of refugees in 1928 exceeds 80% (in red dots); in native settlements, the share of refugees is below 5% (in blue dots). Panel *A* plots nearby (within 25 km) refugee and native settlements across Greece. There are 262 refugee and 1,042 native settlements. Panel *B* zooms into Macedonia and West Thrace. Panel *C* zooms into the area surrounding Giannitsa province in Pella, Central Macedonia.

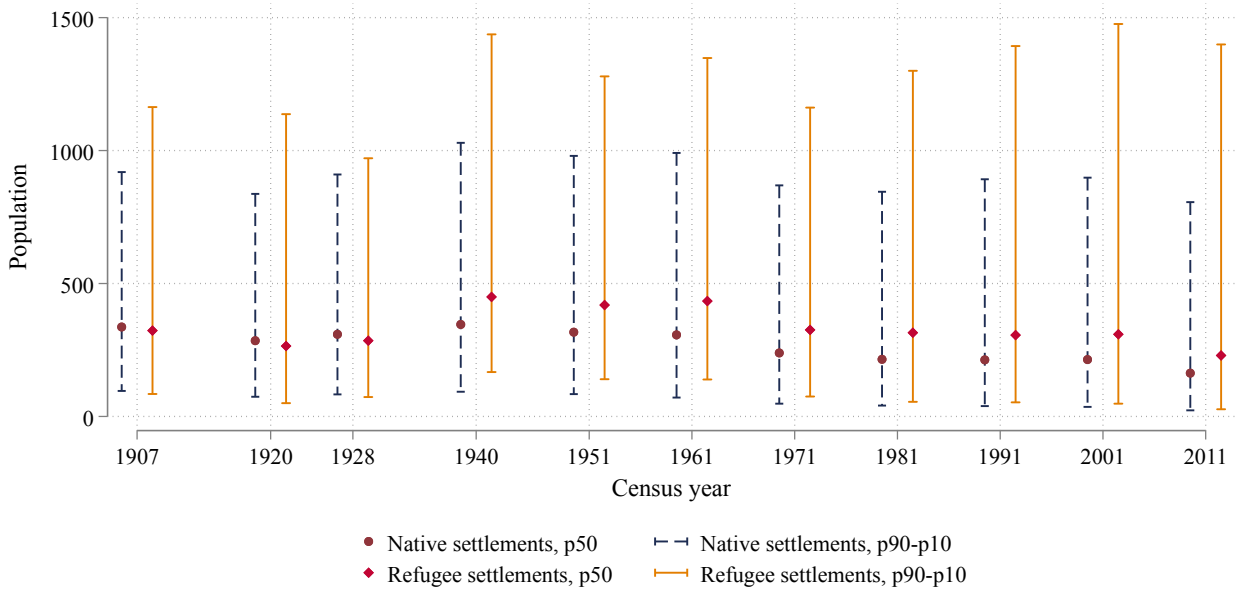
Figure A14: Population Distribution in 1928 Across Refugees and Natives



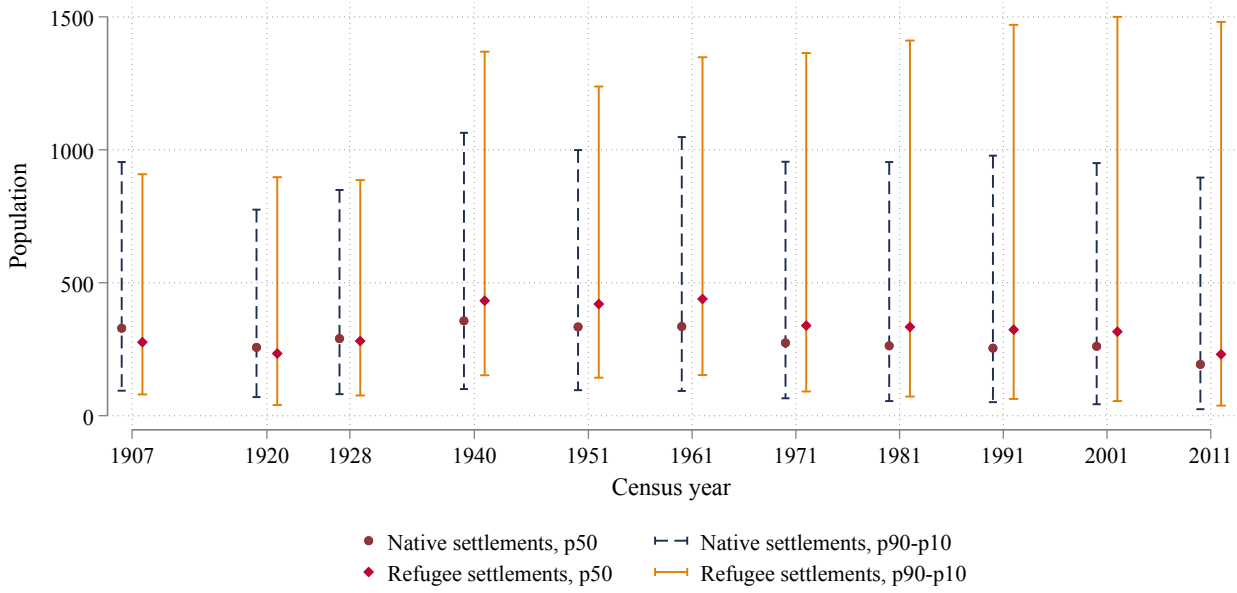
Notes. The figures plot the age distribution by gender among refugees and natives using the 1928 General Population Census.

Figure A15: Population Dynamics. Rural Native and Refugee Settlements

(a) All rural native and refugee settlements



(b) Local sample



Notes. The figures plot the evolution of the median and the 90th- and 10th-percentiles of the population in predominantly refugee and rural native settlements between 1907 and 2011, as recorded in the respective Censuses. Panel *A* plots the population statistics for all rural native and refugee settlements. Panel *B* focuses on the native and refugee settlements in the local sample.

A.2 Correlates of Refugee Share and Balancedness

Here, we report on the correlates of refugee settlement across the Greek countryside, providing details to the discussion in Section 3.3 of the main paper.

Refugee Presence Appendix Table A2 reports the OLS estimates associating the share of refugees in the total population in 1928 with various geographic, location, and ethnographic features across 7,769 rural settlements. First, by far, the most significant correlate of refugee settlement is the share of Muslims forced to leave Macedonia—and, to a lesser extent, Crete and Thessaly, with the Convention for the Population Exchange of 1923. A ten percentage point increase in the share of Muslims in 1920, just before the population exchange, is associated with an eight percentage point higher share of refugees in 1928. The share of Muslims explains about 60% of the variation in refugee share. When conditioning on geography and location features (columns (2)–(5)) or exploit within prefecture (column (4)) or province (column (5)) variation, the coefficient on the share of Muslims in 1920 drops somewhat from 0.83 to 0.67, implying, however, still considerable effects. Second, among the geographic features, the log distance to reclaimed marshlands enters with a highly significant coefficient, supporting the historical narrative that refugees quite often settled in swamps. Besides, refugees settled mainly in the lowlands, but if anything, in areas with lower-quality land. Appendix Figure A16 illustrates the strong correlation between the share of refugees in 1928 and the share of Muslim populations in the Greek countryside in 1920. Panel *A* provides the bin scatter plot showing the unconditional correlation (column (1), Appendix Table A2). Panel *B* plots the correlation conditional on province constants and other covariates (column (4)).

Balance Tests Appendix Figure A17 and A18 report "balance tests" that examine mean differences in geographic, location, and demographic features between predominantly refugee settlements, where the share of refugees in 1928 exceeded 80%, and native settlements, where the refugees' share in 1928 was below 5%. Appendix Figure A17 gives results in the baseline sample, consisting of 405 refugee and 6,091 native settlements. Appendix Figure gives "balancedness tests" in the local sample (Appendix Figure A13), which comprises 262 refugee and 1,042 settlements. In the local sample, refugee settlements are less than 25 km away than native settlements. All panels give standardized coefficients and 90% confidence intervals from OLS regressions that associate the geographic, ecological, demographic, and locational features to an indicator variable that takes the value of one for predominantly refugee settlements and zero for native settlements. Panel *A* gives unconditional specification estimates. Panel *B* gives estimates conditional on log altitude, a significant correlate. Specification 1 (left) does not include any (additional) controls. Specification 2 includes prefecture (admin-2 unit, *nomos*) fixed-effects. Specification 3 includes province (admin-3 units, *eparchies*) fixed-effects. Specification 4 (in Appendix Figure A17) includes sub-province fixed effects, based on Voronoi polygons restricting the comparisons to native and refugee settlements within a 25 km radius. In addition to the significantly higher share of Muslims in refugee settlements, there are significant differences in proximity to the

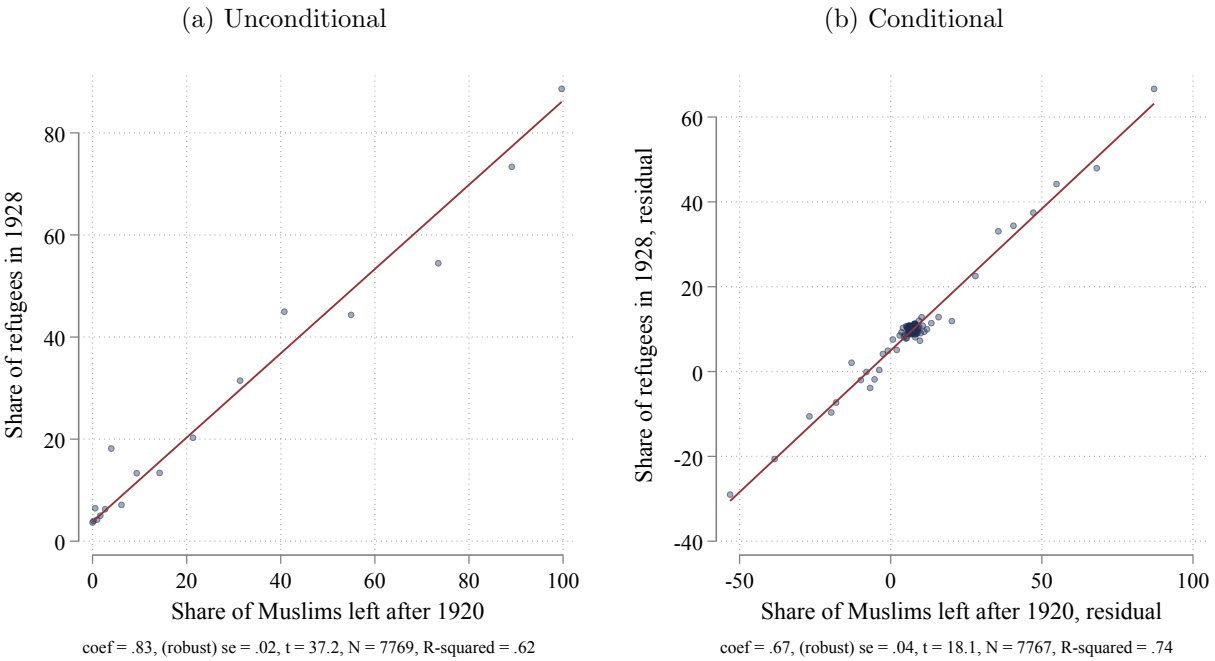
coast, marshlands, and crop suitability. Refugee settlements had a smaller population before the exchange, as many refugees settled in new villages, built by the RSC (see the discussion in Section 2). However, these differences shrink and turn statistically indistinguishable from zero when comparing settlements in the same prefecture or province, and when we condition on log altitude. When we restrict comparisons to proximate settlements in the local sample (Appendix Figure A17), geographical, location, and ecology differences are tiny.

Table A2: Correlates of Refugee Share across Rural Settlements

Dependent Variable:	Refugee share in 1928					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Muslims in 1920, imputed	0.825***	0.738***	0.682***	0.668***		
	(0.022)	(0.026)	(0.033)	(0.037)		
Log population in 1920		-2.317***	-2.271***	-2.290***	-1.776***	-1.862***
		(0.341)	(0.342)	(0.356)	(0.393)	(0.361)
Monastery in 1920		0.152	0.083	0.030	0.037	0.219
		(1.390)	(0.880)	(0.851)	(0.952)	(0.892)
Log distance to reclaimed marshlands		-2.675***	-1.058***	-1.001**	-1.099	-0.573
		(0.341)	(0.398)	(0.406)	(0.677)	(0.685)
Log distance to province capital		-0.294	-0.434	-0.318	0.358	0.135
		(0.538)	(0.444)	(0.517)	(0.800)	(0.841)
Log distance to coast		0.761	0.417	0.673	0.872	0.534
		(0.466)	(0.384)	(0.520)	(0.741)	(0.853)
Log distance to railway stations		-1.009***	-1.122***	-0.951*	-0.974*	-0.919
		(0.282)	(0.342)	(0.506)	(0.546)	(0.591)
Log crop suitability		-1.594**	-1.194*	-1.824**	-0.130	0.172
		(0.666)	(0.643)	(0.737)	(1.326)	(1.607)
Log altitude		-4.327***	-2.657***	-3.061***	-0.729	-1.776
		(0.683)	(0.656)	(0.902)	(1.056)	(1.167)
Log average precipitation		-12.087***	-1.927	-0.214	-18.187**	-4.843
		(1.874)	(2.997)	(4.257)	(8.199)	(5.115)
Log average temperature		-18.978***	-0.597	3.990	0.604	3.571
		(6.116)	(4.662)	(5.314)	(7.930)	(9.297)
R-squared	0.620	0.710	0.738	0.750	0.482	0.544
Observations	7,769	7,767	7,767	7,767	7,767	7,767
Average outcome	10.0	10.0	10.0	10.0	10.0	10.0
Prefecture FE			✓		✓	
Province FE				✓		✓

Notes. The table gives OLS estimates associating the share of refugees after the Asia Minor Catastrophe (in the fall of 1922) across 7,769 rural settlements, as recorded in the 1928 General Population Census, to various location, geographic, and demographic features. The specifications in columns (3) and (5) include prefecture fixed-effects and the specifications in column (4) and (6) province fixed-effects. Standard errors are clustered at the province level. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

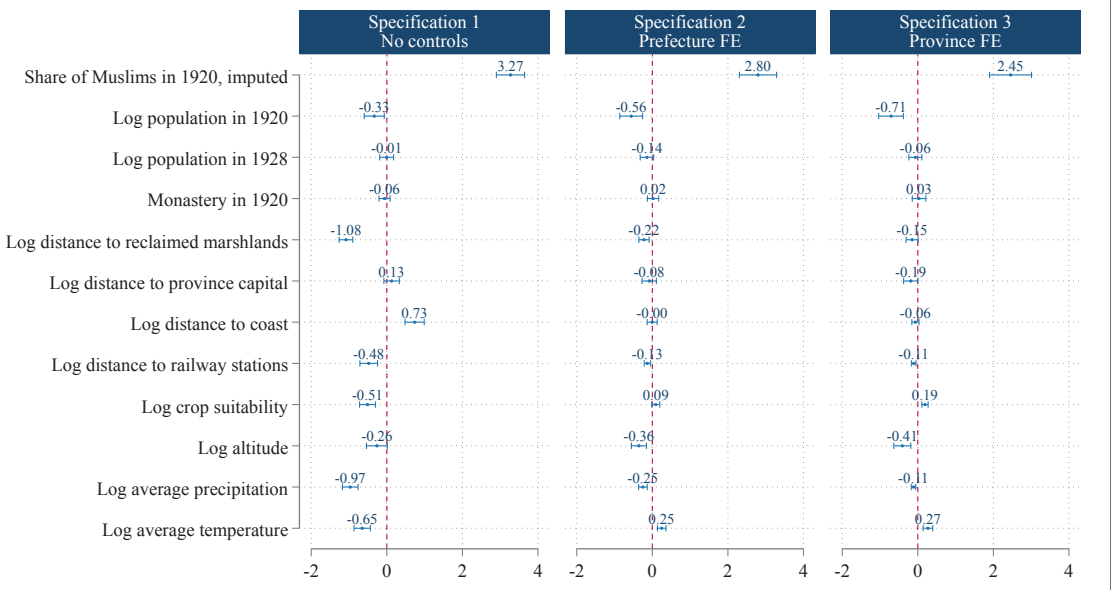
Figure A16: Binscatter Share of Refugees in 1928 and the Share of Muslims in 1920



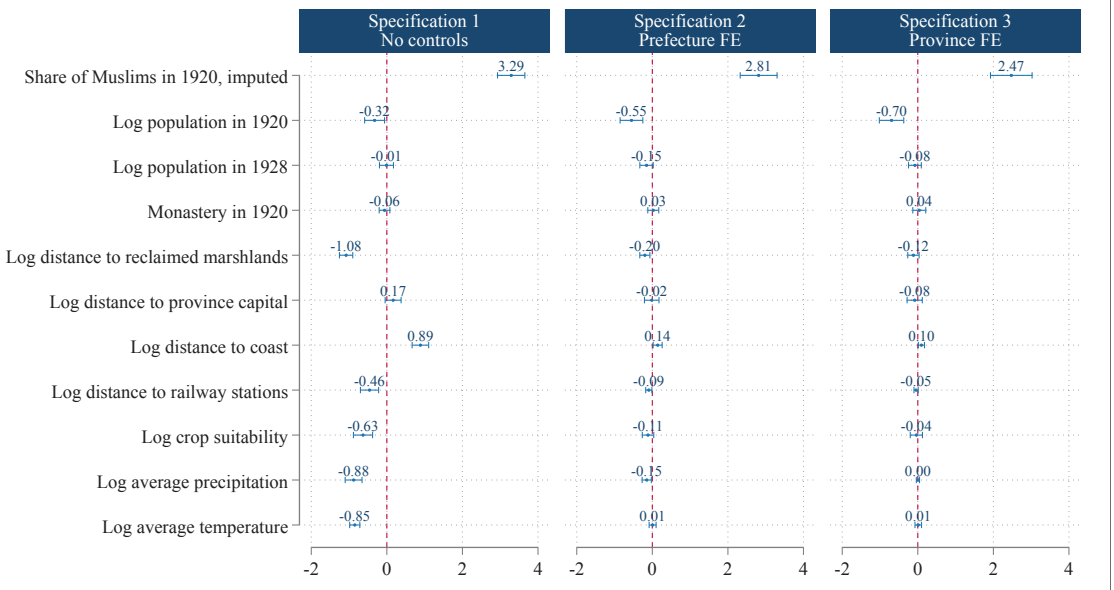
Notes. The panels provide the bin scatter plot across 100 bins visualizing the correlation between the share of Muslims departing after 1920 and the share of refugees in 1928. Panel *A* presents he unconditional correlation presented in column 1 of Appendix Table A2 while Panel *B* presents the correlation conditional on province fixed effects and covariates presented in column 4. Standard errors are clustered at the province level.

Figure A17: Balancedness across All Rural Native and Refugee Settlements

(a) Unconditional



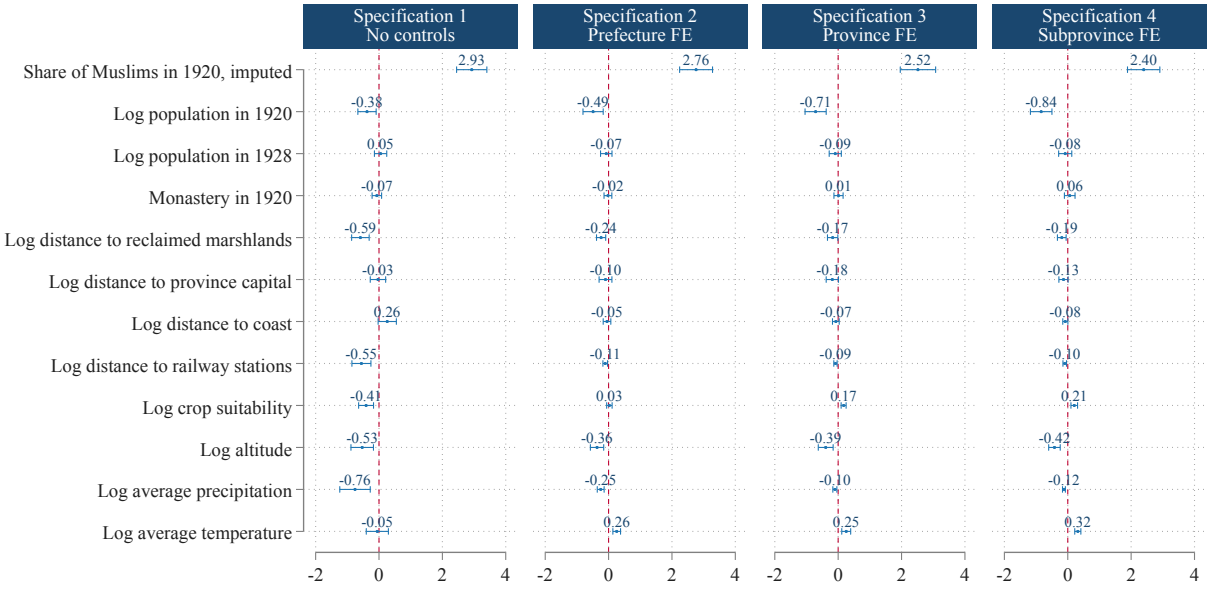
(b) Conditional on log altitude



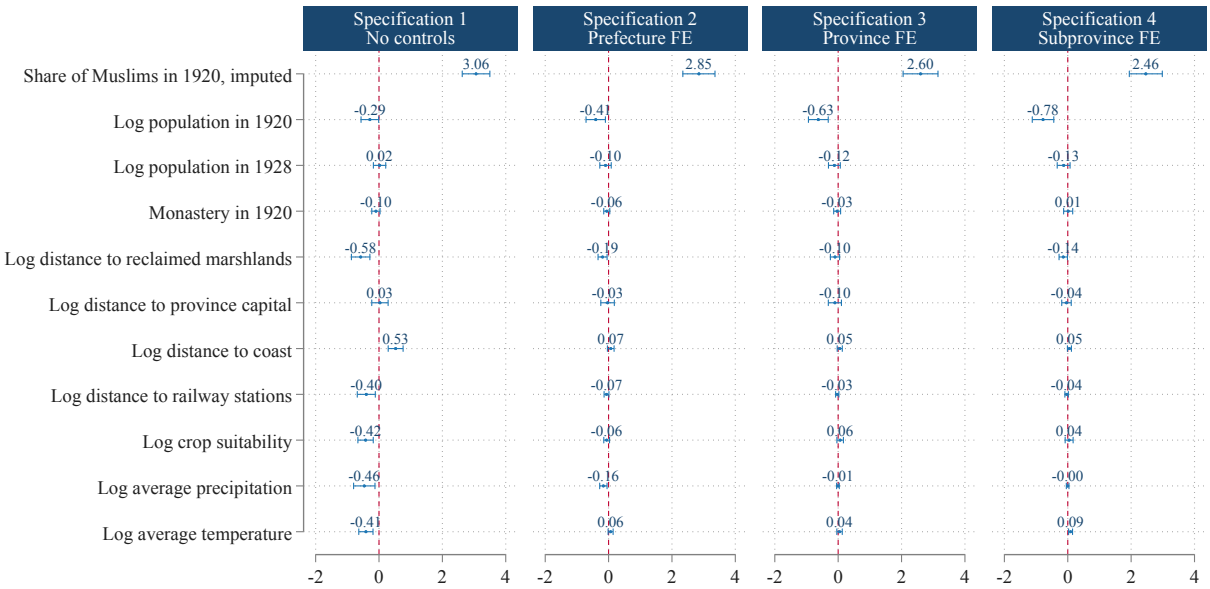
Notes. The panels explore whether various geographic, location, and population features differ between predominantly refugee and native settlements in the baseline sample of rural settlements. All specifications in both Panels give standardized coefficients and 90% confidence intervals obtained from regressing standardized settlement characteristics on a dummy variable indicating predominantly refugee settlements, where the share of refugees in 1928 exceeds 80%; the indicator equals zero for native settlements, where the share of refugees in 1928 is below 5%. Panel *A* presents unilateral unconditional specifications. Panel *B* gives the regression estimates conditioning on settlements log altitude. Specification 1 does not include any controls. Specification 2 conditions on prefecture fixed effects. Specification 3 conditions on province fixed-effects. Standard errors are clustered at the province level.

Figure A18: Balancedness across Rural Native and Refugee Settlements in the Local Sample

(a) Unconditional



(b) Conditional on log altitude



Notes. The panels explore whether various geographic, location, and population features differ between predominantly refugee and native settlements in the local sample of rural settlements. All specifications in both Panels give standardized coefficients and 90% confidence intervals obtained from regressing standardized settlement characteristics on a dummy variable indicating predominantly refugee settlements, where the share of refugees in 1928 exceeds 80%; the indicator equals zero for native settlements, where the share of refugees in 1928 is below 5%. Panel *A* presents unilateral unconditional specifications. Panel *B* gives the regression estimates conditioning on settlements log altitude. Specification 1 does not include any controls. Specification 2 conditions on prefecture fixed effects. Specification 3 conditions on province fixed-effects. Specification 4 conditions on sub-province fixed effects, based on Voronoi polygons. Standard errors are clustered at the province level.

B Displacement and Human Capital

Appendix Section B complements the descriptive and regression analyses linking human capital to forced displacement in Section 4 of the paper. First, we report descriptive evidence on (trends in) literacy rates for refugees and non-refugees before the displacement. Second, we give additional results on the association between (changes in) education and refugee lineage. Third, we report further evidence and various robustness checks, revealing the stability of the baseline estimates to alternative definitions of refugee and native settlements and different cutoffs of proximity in the local analysis. Fourth, we report results linking the share of refugees in 1928 to subsequent human capital across all rural settlements: besides predominantly refugee and autochthonous ones, we include mixed refugee-native ones. Fifth, we report gender-specific estimates.

B.1 Descriptive Patterns. Initial Differences and Dynamics

Initial Differences (1971-1981 Censuses) Appendix Figure B1 plots the mean literacy (in Panel *A*) and primary school completion rates (in panel *B*) for those born before 1902 from predominantly refugee and native settlements using the 1971 and 1981 General Population Censuses. The figure, which complements the descriptions in Figure 4 in the main paper, shows the lower levels of education for Greeks in predominantly refugee settlements compared to those from native settlements.

Pre-Displacement Differences in Education. Rural Refugees and Non-Refugees Appendix Figure B2 plots the differences in the literacy of refugees and natives by birth-cohort, as recorded in the 1928 General Population Census. Panel *A* plots the differences in literacy among the rural population across the country. Panel *B* zooms into Macedonia and West Thrace as eight out of ten rural refugees settled there. Both panels illustrate the drop in literacy for the Greek-Orthodox communities in Asia Minor following the repression of the Ottoman Empire, which accelerated after 1908 and further intensified during the Greco-Turkish hostilities after the Great War.

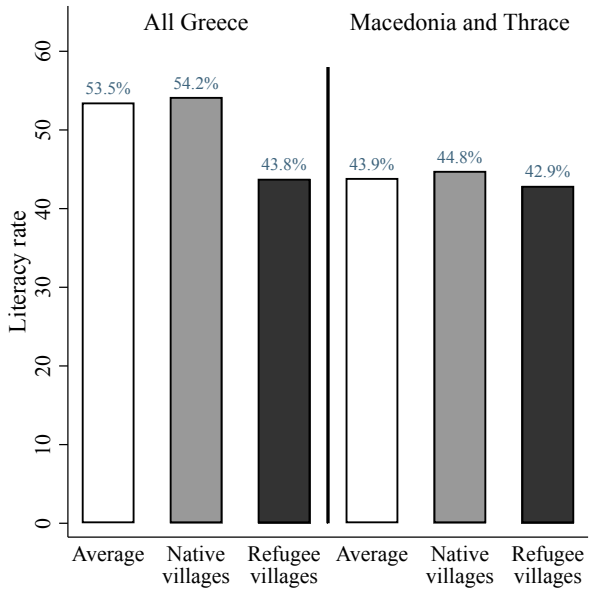
Pre-Trends in Literacy. Refugees and Non-Refugees Appendix Table B1 reports specifications exploring whether the literacy increase during the 19th century is similar for refugees and non-refugees. We retrieve from the 1928 General Population Census five-year birth-cohort level statistics on literacy rates for residents in the countryside and then regress the literacy rates on a linear time trend, an indicator variable for refugees, and an interaction between the time trend and the refugee indicator. As revealed by the trend estimate, literacy has increased steadily for natives and refugees (when the latter resided in Anatolia). The refugee indicator enters with a negative coefficient, telling of the refugees' lower literacy by about 2.5–5.7 percentage points. However, the coefficient on the refugee dummy drops in absolute value and becomes statistically indistinguishable from zero in Macedonia and Thrace when we omit the cohorts impacted by the Greco-Turkish warfare and the ethnic violence against minorities in the Ottoman Empire after the Young Turks rose in 1908. Most importantly, the coefficient on the interaction term that estimates the differential trend between natives and refugees

is small and statistically insignificant, both when we analyze the patterns across the country (columns (1)–(2)) and when we zoom into Macedonia and West Thrace (columns (3)–(4)).

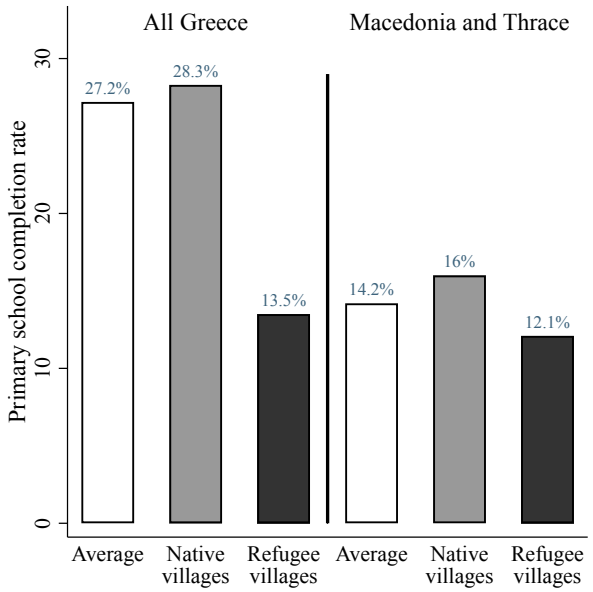
Educational Dynamics. Rural Refugee and Native Settlements Appendix Figure B3 plots the dynamics of schooling years (left) and junior high school completion rates (right) for Greeks living in predominantly refugee and native settlements, pooling the micro samples of the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. Panel *A* considers all predominantly refugee and native settlements, while Panel *B* focuses on Macedonia and West Thrace. The descriptive analysis uncovers two patterns. First, in line with the tabulations of the 1928 General Population Census, human capital was lower in predominantly refugee settlements compared to native settlements for those who started schooling prior to the population exchange. Those born outside Greece before WWI who settled in the countryside after the population exchange had lower completion rates of junior high school. When we zoom into Macedonia and West Thrace, differences become muted because the human capital of non-refugees was lower than that of natives in Central and Southern Greece. Second, upon arriving in Greece, mainly in Macedonia and West Thrace, refugees in the countryside caught up and soon surpassed the natives' education. Second, the refugee-native differences in junior high completion rates for individuals in Macedonia and West Thrace in the 1950s, 1960s, and 1970s are a staggering 15 percentage points.

Figure B1: Education across Refugee and Native Settlements. Greeks born before 1902. 1971–1981 Censuses

(a) Literacy Rates



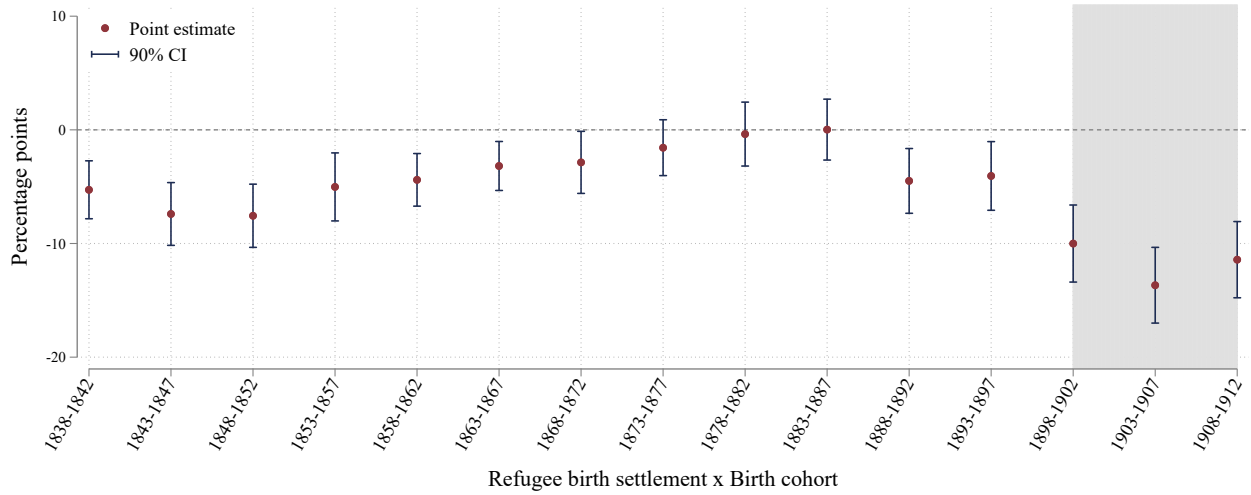
(b) Primary School Completion



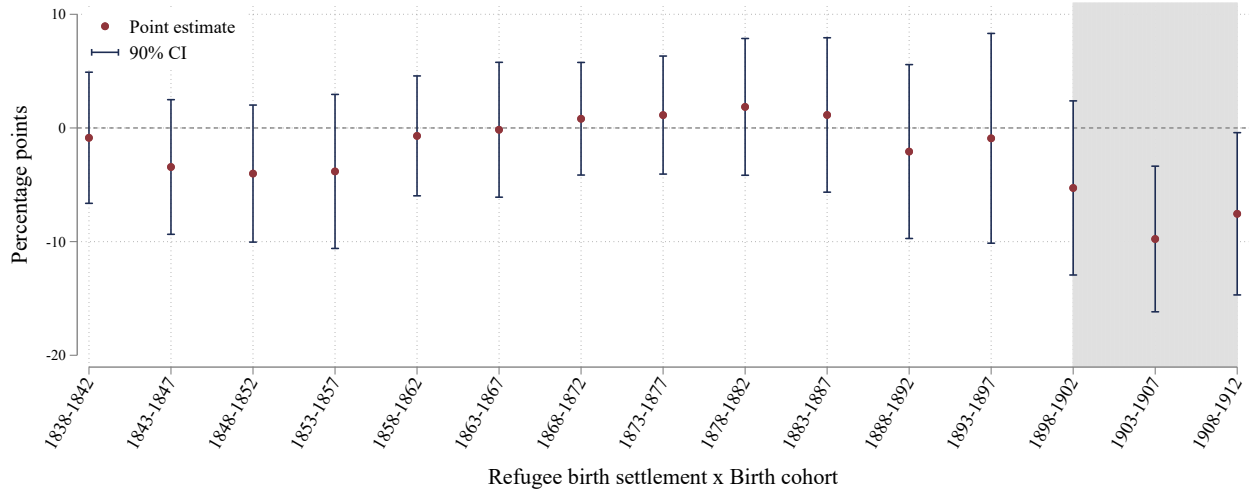
Notes. Panel *A* plots literacy rates for all Greeks (white bars) born before 1902 living in the countryside (rural), in native settlements (gray bars), and in refugee settlements (black bars). Panel *B* plots primary school completion rates for all Greeks (white bars) born before 1902 living in the countryside (rural), in refugee settlements (gray bars), and in native settlements (black bars). Data come from the 1971 and 1981 General Population Censuses. Both panels report means across Greece (left), and Macedonia and West Thrace (right).

Figure B2: Gap in Literacy between Refugees and Natives in Rural Settlements in 1928

(a) All Greece



(b) Macedonia and West Thrace



Notes. All panels plot the difference in literacy rate in 1928 between rural natives and refugees across provinces, retrieved from OLS estimates, weighted by province population. Panel *A* plots the unconditional difference across all 141 provinces. Panel *B* plots the difference across all provinces in Macedonia and West Thrace, as 79.8% of rural refugees settled there. Literacy rates are retrieved from the 1928 General Population Census. Gray indicates the cohorts affected by Ottoman repression, the Balkan Wars, the Great War, and the Greco-Turkish War. Standard errors are clustered at the prefecture level.

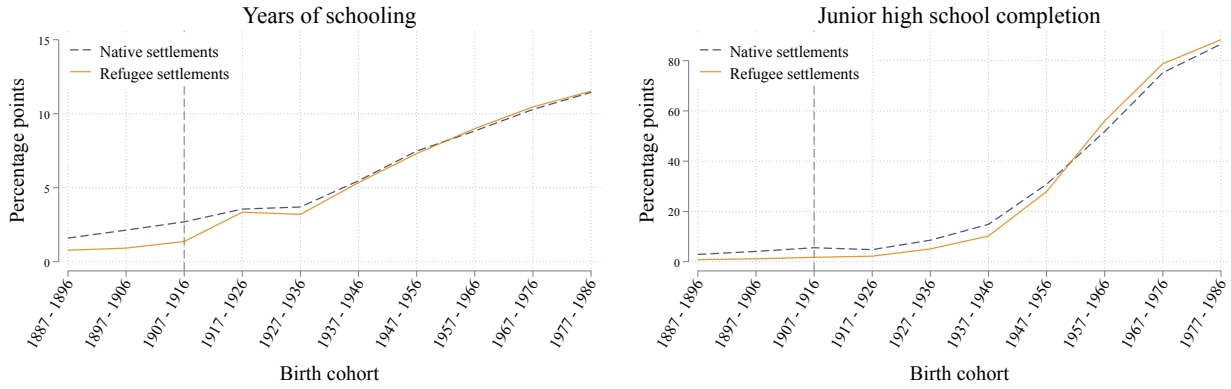
Table B1: Testing for Differential Trends across Refugees and Natives before the Population Exchange

Sample:	All sample		Macedonia & Thrace	
Cohorts:	All	Excluding 1898–1912	All	Excluding 1898–1912
	(1)	(2)	(3)	(4)
Refugee	-5.704***	-2.757**	-2.527***	-0.197
	(1.055)	(1.084)	(0.866)	(0.914)
Time trend	3.039***	2.441***	2.979***	2.542***
	(0.225)	(0.229)	(0.185)	(0.199)
Refugee × Time trend	-0.287	0.436	-0.286	0.289
	(0.263)	(0.301)	(0.238)	(0.296)
R-squared	0.957	0.958	0.966	0.960
Observations	30	24	30	24

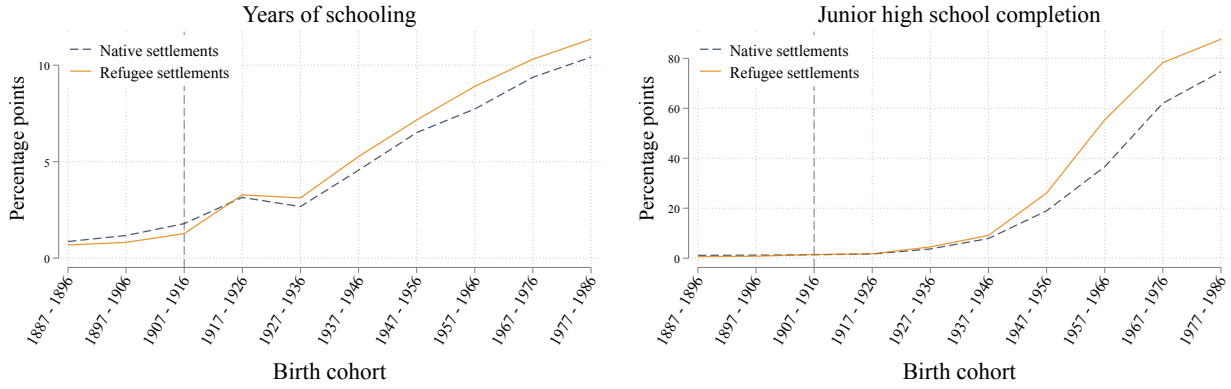
Notes. The table gives estimates testing for differential trends in literacy rates between refugees and natives, as observed in the 1928 General Population Census. All columns report OLS estimates associating average literacy rates, by 10-year birth cohorts and refugee status, to a linear time trend, an indicator variable that takes the value of one for refugees and zero for non-refugees, and an interaction between the linear time trend and the refugee indicator. Columns (1) and (2) consider all provinces; columns (3) and (4) focus on Macedonia and Western Thrace provinces. Columns (2) and (4) exclude the cohorts born between 1898 and 1912, affected by Ottoman repression, the Balkan Wars, the Great War, and the Greco-Turkish War. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Figure B3: Human Capital Dynamics across Rural Refugee and Native Settlements. 1971–2011 Censuses

(a) All provinces



(b) Macedonia and West Thrace



Notes. All panels plot the evolution in years of schooling and junior high school completion rates across all predominantly refugee settlements (in solid orange) and native (non-refugee) settlements (in dashed blue) across 10-year birth cohorts. Panel A considers the rural predominantly refugee and native settlements across all provinces, while Panel B restricts the sample to those in Macedonia and West Thrace. Data come from the 1971, 1981, 1991, 2001, and 2011 Population Censuses. Predominantly refugee settlements are those where the share of refugees from Anatolia in 1928 exceeds 80%; native settlements are those where the share of refugees in 1928 is lower than 5%.

B.2 Additional Results

Here, we report additional results on the link between forced displacement and subsequent human capital investments across the Greek countryside (Sections 4.2 and 4.3).

B.2.1 Difference-in-Difference Estimates. Local Sample

Appendix Figure B4 plots the event-study graphs showing the within-settlement correlation between human capital and living in refugee settlements across birth cohorts in the local sample; the comparison is between predominantly refugee villages and towns (where the refugee share exceeded 80% in 1928) and native settlements (where the share of refugees was below 5%) in the same province, no more than 25 kilometers away; regression equation (1). While standard errors are wider compared to the baseline sample estimates, as the sample is considerably smaller, the patterns are similar to the within-province estimates in Figure 6. First- and second-generation refugees have a four to five percentage points higher likelihood of completing primary education. Over time, second-generation refugees born in the 1940s–1970s have a four to six percentage points higher chance of completing junior high school; second- and even third-generation refugees born after WWII have a two to six percentage point higher likelihood of completing high school, while the latest cohorts of Greeks born in refugee settlements have a one to two percentage points higher propensity finishing tertiary education.

B.2.2 Conditioning on Current Residence

Appendix Figure B5 plots cohort-specific estimates on the refugee indicator that shed light on the differential dynamics of schooling between Greeks born in predominantly refugee and native settlements, using the 2011 General Population Census that gives details on the respondent's family birthplace. These estimates thus complement the regression estimates in Table 2. Two results emerge. First, years of schooling are considerably higher for Greeks from refugee settlements compared to those born in native settlements in the same prefecture (admin-2 level unit), province (admin-3 level unit), and sub-province level in the local sample. Second, while the cohort-specific estimates are somewhat noisy, they illustrate that the educational gain of second-, third-, and fourth-generation refugees is considerable and stable over time.

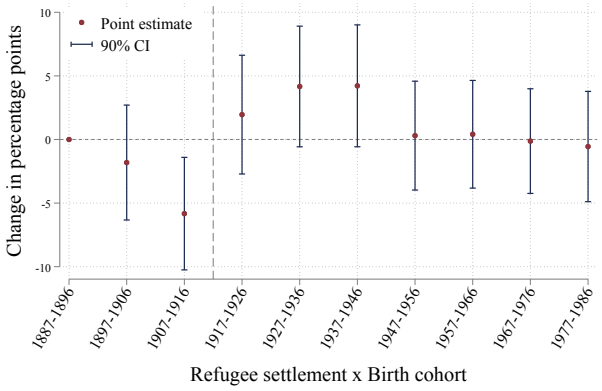
Appendix Table B2 associates primary school, junior high school, high school, and tertiary education completion rates with the refugee indicator that takes the value of one for Greeks from predominantly refugee settlements and zero for individuals from native settlements, conditioning on municipality of residence in 2011 (regression equation (2)). Panel A gives the coefficients across Greek citizens, aged 25–90, born in all rural predominantly refugee and native settlements; Panels B and C report the estimates for individuals who continue residing in their birthplace (non-migrants) and for Greeks who moved out (internal migrants), respectively. The results complement the ones in Table 2–Panel A in the paper, which looks at schooling years. The educational gain of second-, third-, and fourth-generation refugees is present both when looking at Greeks living in their families' home-

towns (Panel B) and at internal migrants (Panel C), who mainly moved to Athens, Thessaloniki, and Piraeus. Besides, the patterns are similar to the difference-in-difference specifications (using all post-1971 Censuses) in Table 1 in the main paper.

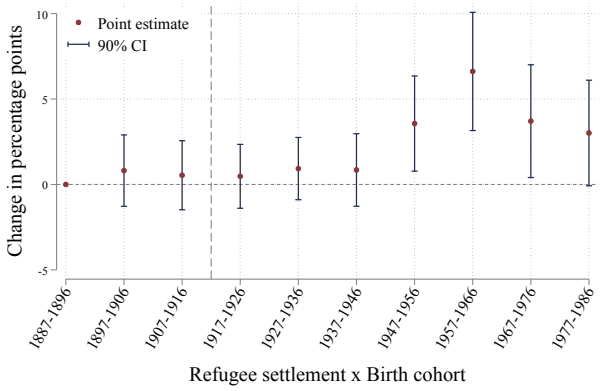
Appendix Figure B6 plots the corresponding cohort-specific estimates that allow exploring the dynamics of educational attainment rates for the two groups. We condition on the municipality of residence. The patterns illustrate the dynamics of uprootedness over three generations. First- and second-generation refugees have a much higher propensity, about ten percentage points, to complete primary schooling than Greeks born in native settlements (Panel A). As primary school completion reaches 100% for all Greeks, then the educational advantage of Greeks born in refugee settlements emerges for junior high and high school completion (Panels B and C). Gradually, the refugee-native educational differences also become evident, strengthening over time, when looking at college/university completion (Panel D). In addition, the patterns in Appendix Figure B6 also echo the difference-in-difference estimates in the main paper (Figure 6).

Figure B4: Education Dynamics across Refugee and Native Settlements. Local Sample. 1971–2011

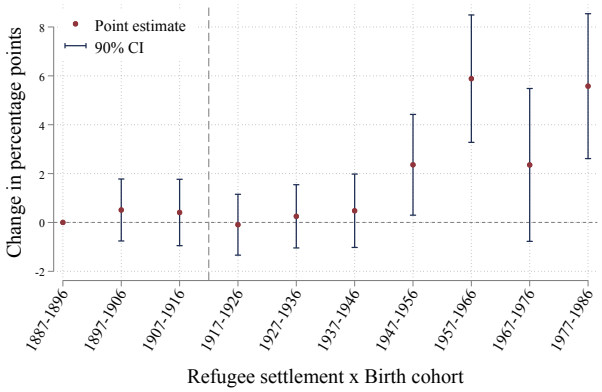
(a) Primary school completion



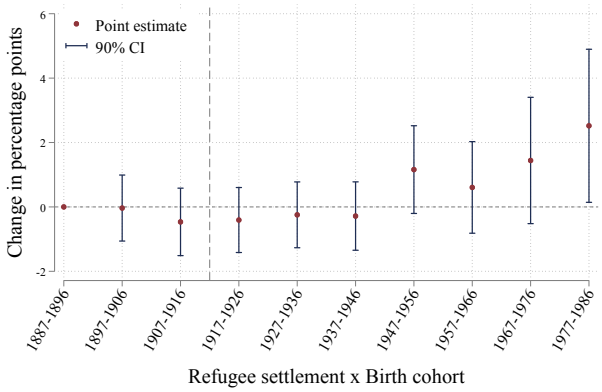
(b) Junior high school completion



(c) High school completion



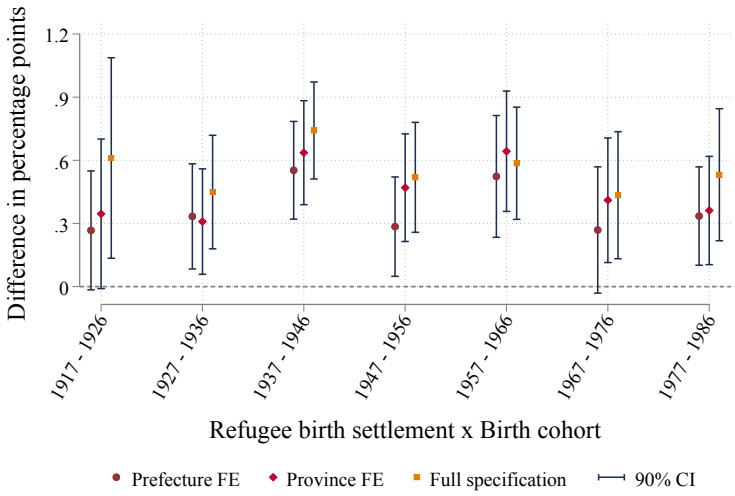
(d) Tertiary education completion



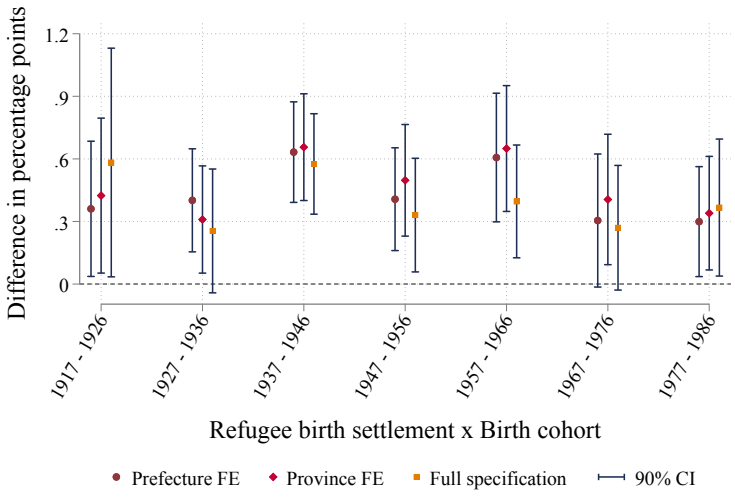
Notes. The panels plot the event-study OLS (linear probability model) estimates obtained from associating indicator variables for completing primary school (in Panel *A*), for competing junior high school (in Panel *B*), for competing high school (in Panel *C*), and for completing tertiary education (in Panel *D*) on a birth-cohort varying dummy variable indicating whether an individual resides in a predominantly refugee settlement, where the share of refugees in 1928 exceeds 80%. Estimation is in the local sample that comprises native and refugee settlements in the same sub-province unit (Voronoi polygon) within a 25 km radius. All specifications include settlement fixed effects, interactions of census-year and birth-cohort fixed effects, interactions of sub-province fixed effects and birth-cohort fixed effects, interactions of individual characteristics (age and gender) and birth-cohort fixed effects, and interactions of various settlement characteristics and birth-cohort fixed effects. The settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1897–1906 birth cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level.

Figure B5: Years of Schooling across Rural Refugee and Native Birth Settlements, cond. on Residence (2011)

(a) All Settlements



(b) Local Sample



Notes. The panels present the cohort-level estimates obtained from regressing years of schooling on a dummy variable indicating whether an individual was born in a predominantly refugee settlement, where the share of refugees in 1928 exceeds 80%, for each birth-cohort separately; regression equation (2). Panel *A* gives estimates across all rural refugee and native birth settlements. Panel *B* focuses on the local analysis sample, where refugee birth settlements are matched to native ones no more than 25 km away in the same sub-province unit. Maroon circles and red diamonds plot the estimates obtained from parsimonious specifications exploiting within-birth-prefecture (*nomos*) and within-birth-province (*eparchia*) variation; these specifications only condition on interactions of the birth prefecture or the birth province fixed effects. Orange squares plot the estimates obtained from a rich (full) specification that also controls for individual characteristics (age and gender) and birth settlement characteristics, and also for sub-province fixed effects in Panel *B*. Standard errors are clustered at the birth-settlement and the municipality-of-residence levels.

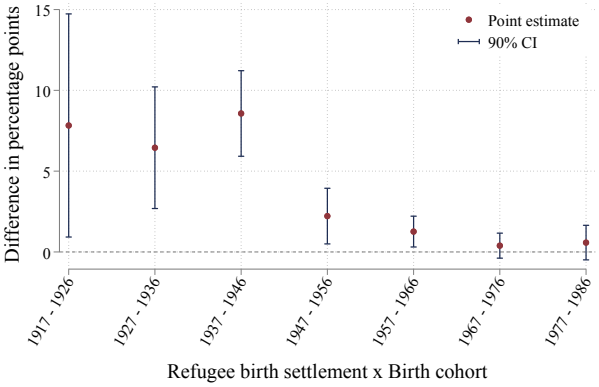
Table B2: Educational Attainment across Native and Refugee Birth Settlements, cond. on Residence (2011)

Dependent Variable:	Completed education level:							
	Primary school		Jun. high school		High school		Tertiary education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A:	**All sample**							
Born in a refugee settlement	4.365***	3.961***	6.788***	5.401***	5.327***	4.076***	2.191***	1.423**
	(0.804)	(0.795)	(1.311)	(1.240)	(1.176)	(1.134)	(0.601)	(0.568)
R-squared	0.297	0.311	0.323	0.360	0.227	0.267	0.059	0.088
Observations	549,480	549,474	549,480	549,474	549,480	549,474	549,480	549,474
Average outcome in native settlements	86.2	86.2	45.6	45.6	33.4	33.4	11.4	11.4
# Refugee observations	44,115	44,114	44,115	44,114	44,115	44,114	44,115	44,114
# Refugee birth settlements	357	357	357	357	357	357	357	357
# Native observations	505,365	505,360	505,365	505,360	505,365	505,360	505,365	505,360
# Native birth settlements	5,788	5,788	5,788	5,788	5,788	5,788	5,788	5,788
# Birth provinces for FE	141	141	141	141	141	141	141	141
# Residence-municipality FE		1,021		1,021		1,021		1,021
Panel B:	**Non-migrants**							
Born in a refugee settlement	3.780***	4.688***	5.238***	4.613***	3.901***	3.632**	1.858***	0.988
	(0.981)	(1.247)	(1.296)	(1.745)	(1.199)	(1.631)	(0.637)	(0.606)
R-squared	0.321	0.331	0.402	0.415	0.296	0.309	0.087	0.095
Observations	243,038	243,005	243,038	243,005	243,038	243,005	243,038	243,005
Average outcome in native settlements	83.2	83.2	40.0	40.0	26.8	26.8	6.7	6.7
# Refugee observations	18,259	18,255	18,259	18,255	18,259	18,255	18,259	18,255
# Refugee birth settlements	351	351	351	351	351	351	351	351
# Native observations	224,779	224,750	224,779	224,750	224,779	224,750	224,779	224,750
# Native birth settlements	5,688	5,688	5,688	5,688	5,688	5,688	5,688	5,688
# Birth provinces for FE	140	140	140	140	140	140	140	140
# Residence-municipality FE		943		943		943		943
Panel C:	**Internal migrants**							
Born in a refugee settlement	3.977***	3.043***	6.762***	4.996***	5.572***	3.961***	1.732**	1.154
	(0.905)	(0.884)	(1.542)	(1.424)	(1.421)	(1.317)	(0.805)	(0.740)
R-squared	0.286	0.298	0.298	0.328	0.215	0.249	0.066	0.088
Observations	289,213	289,208	289,213	289,208	289,213	289,208	289,213	289,208
Average outcome in native settlements	88.8	88.8	50.9	50.9	39.4	39.4	15.4	15.4
# Refugee observations	22,674	22,674	22,674	22,674	22,674	22,674	22,674	22,674
# Refugee birth settlements	355	355	355	355	355	355	355	355
# Native observations	266,539	266,534	266,539	266,534	266,539	266,534	266,539	266,534
# Native birth settlements	5,704	5,704	5,704	5,704	5,704	5,704	5,704	5,704
# Birth provinces for FE	140	140	140	140	140	140	140	140
# Residence-municipality FE		1,020		1,020		1,020		1,020
Cohort × Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth settlement controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth province FE	✓	✓	✓	✓	✓	✓	✓	✓
Residence municipality FE		✓		✓		✓		✓

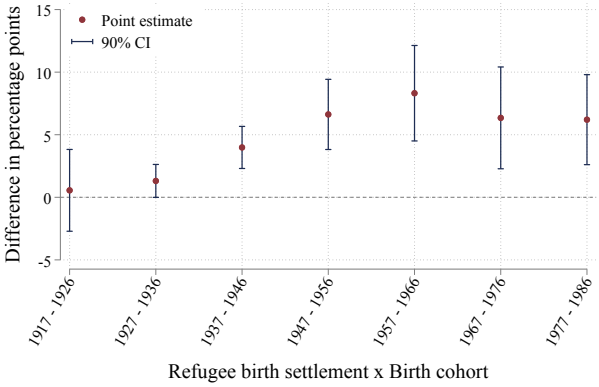
Notes. The Panels report OLS estimates associating being born in a predominantly refugee settlement and educational attainment across all rural predominantly refugee and native birth settlements. The dependent variable is primary school completion in columns (1)–(2), junior high school completion in columns (3)–(4), high school completion in columns (5)–(6), tertiary education completion in columns (7)–(8). Panel *A* considers all residents of rural refugee and native birth settlements, while Panel *B* restricts estimation to individuals who reside in their birth settlement and who did not migrate abroad, and Panel *C* restrict estimation to those who reside in a municipality other than their birth settlement. All reported regressions control for the interactions of the birth-province fixed effects and birth-cohort fixed effects, the interactions of birth-settlement characteristics and birth-cohort fixed effects and the interactions of individual characteristics (age and gender) and birth-cohort fixed effects. Columns (2), (4), (6), and (8) control for municipality-of-residence fixed effects. Birth settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. Standard errors are clustered at the birth-settlement and the municipality-of-residence levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Figure B6: Educational Attainment across Rural Native and Refugee Birth Settlements, cond. on Residence (2011)

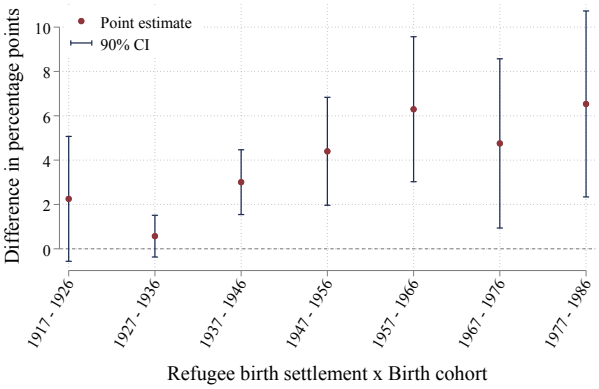
(a) Primary school completion



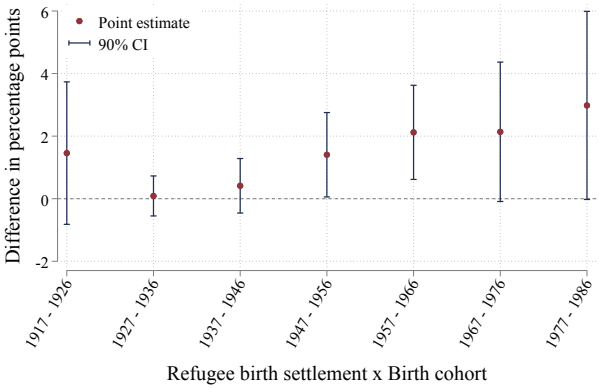
(b) Junior high school completion



(c) High school completion



(d) Tertiary education completion



Notes. The panels present the cohort-level estimates obtained from regressing indicator variables for completing primary school (in Panel *A*), for competing junior high school (in Panel *B*), for competing high school (in Panel *C*), and for completing tertiary education (in Panel *D*) on a dummy variable indicating whether an individual resides in a predominantly refugee settlement, where the share of refugees in 1928 exceeds 80%, for each birth-cohort separately. The analysis considers all rural refugee and native birth settlements. All panels control for birth-province fixed effects, birth-settlement characteristics, and individual characteristics (age and gender). Birth settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. Standard errors are clustered at the birth-settlement and the municipality-of-residence levels.

B.3 Further Evidence

B.3.1 Sectoral Employment

Appendix Table B3 reports conditional on residence estimates (regression equation 1), which also account for individuals' sector of employment. The sample consists of respondents between the ages of 25 and 55 in the 2011 General Population Census (25% micro sample). We focus on this age interval as employment drops considerably among Greeks after the mid-fifties. Column (1) gives the estimates on the refugee indicator that takes the value of one for respondents from predominantly refugee settlements (where the share of refugees in 1928 exceeded 80%) and zero for native settlements (where the refugee share was below 5% in 1928) among 25-55-year-olds. When we condition on the respondents' current residence, in (2), the coefficient on the refugee indicator is somewhat smaller than the estimate in the main paper, estimated across 25-90-year-olds (0.53 vs 0.58). In columns (3)-(4), we augment the specification with broad sector of employment fixed effects (15 industries), which, conservatively, we allow to differ across birth cohorts. In columns (5)-(6), we use the most granular sectoral disaggregation of the 2011 Cesnus that distinguishes across 260 sectors. The estimate in (6) suggests that third and fourth-generation refugees have about 0.3 years more schooling than Greeks from native settlements in the same *eparchia*, even when looking among Greeks working in the same industry and living in the same municipality. The coefficient is approximately a third lower than the full sample estimate (0.53) in column (2). This suggests that sectoral employment is a significant underlying factor behind refugee families' higher propensity to invest in their children's education.

B.3.2 Spatial Spillovers

Appendix Table B4 reports difference-in-difference specifications testing for spatial spillovers from refugee to native settlements. For each native settlement, we calculate the share of refugees (as recorded in the 1928 General Population Census) within 5 km, considering all predominantly refugee and native villages in the baseline sample. Columns (1) and (3) report OLS estimates across all rural native settlements, while columns (2) and (4) focus on rural native settlements in the local analysis sample. The specifications in (1) and (2) show that the presence of refugees within 5 km does not affect natives' years of schooling after the resettlement of refugees (controlling for the log of total population among refugee and native settlements within 5 km). Likewise, the estimates in (3) and (4), where we control for the share of refugees in 1928 and log the total population within a 5 to 25 km radius, also yield little evidence of spillovers.

B.3.3 Gender

Below, we report descriptive evidence, difference-in-difference specifications (regression equation (1)), and conditional-on-residence (regression equation (2)) results distinguishing by respondent's gender.

Before-After Mean Differences Appendix Figure B7–Panel *A* reports mean literacy rates for men and women refugees and non-refugees living in the Greek countryside using the breakdown provided by the 1928 General Population Census. These results thus complement the patterns in Figure 4, which does not distinguish by gender. Literacy rates are lower for both men and women refugees as compared to non-refugee men and women across the country (left) and in Macedonia and West Thrace (right); differences in Macedonia and West Thrace were smaller as literacy rates were overall higher in Sterea Ellada, the Peloponnese, the Ionian Sea Islands, and Crete. Panel *B* reports the mean completion rates of junior high school for men and women born in predominantly refugee and native settlements using the 2011 General Population Census. The considerable advantage of natives in literacy before the Population Exchange has disappeared, with junior high completion rates being similar in 2011 for those born in refugee and native settlements (left). Zooming into Macedonia and West Thrace yields even more striking patterns (right). Women and men from predominantly refugee settlements have significantly higher human capital than Greeks from native settlements; the difference in junior high school completion rates is considerable, about ten percentage points for both men and women.

Difference-in-Difference Estimates Appendix Table B5 reports the difference-in-difference specification separately for men (odd-numbered columns) and women (even-numbered columns), complementing the analysis in Table 1 in the main paper. The coefficients for women are higher across all educational attainment categories (and years of schooling). However, statistical tests fail to reject the null hypothesis of equality of coefficients between gender. The somewhat stronger uprootedness effects for women may reflect their higher propensity to work in manufacturing and services (as compared to agriculture) and more substantial returns to schooling in the "modern sector."

Dynamics Appendix Figure B8 plots the cohort-specific estimates of the refugee indicator when regression equations (specification (1) in Panel *A*) and (specification (2) in Panel *B*) are estimated separately for men and women to explore dynamics. While the cohort-specific coefficients are noisy, the dynamic analysis shows that refugee men fare better than men born or living in native settlements in the initial post-displacement decade(s), and, over time, the refugee-native difference in schooling for women becomes more prominent and stable.

B.3.4 Sensitivity Analysis. Refugee and Native Settlements

Cutoffs Appendix Figure B9 reports OLS estimates associating years of schooling with the refugee indicator using alternative classifications for predominantly refugee and native settlements. Panel *A* reports difference-in-difference specifications, which compare the difference in years of schooling in refugee and native rural settlements before and after the forced population exchange (regression equation 1). Panel *B* reports OLS specifications that compare schooling years between those from predominantly refugee and non-refugee settlements residing in the same municipality in 2011 (regression equation (2)). We do two permutations. First, we define refugee settlements as those where, in 1928,

the share of refugees from Anatolia exceeded 50%, 60%, 70%, 80% (baseline), and 90%. Second, we define non-refugee (native) settlements as those where the share of refugees in 1928 is less than 10%, 5% (baseline), and even 1%. The significant educational gain for Greeks born in refugee compared to native settlements is present across all permutations.

Very Local Analysis Appendix Figure B10 gives difference-in-difference specifications (in Panel *A*) and specifications conditioning on residence in 2011 (in Panel *B*) restricting estimation to Greeks living in proximate refugee and native settlements in the same province, using increasingly narrower radii of 20 km, 15 km, 10 km, and 5 km to define the local sample. Even when comparing Greeks living in refugee to native settlements no more than 10 km—or even 5 km, we find a gain for second-, third-, and fourth-generation refugees with about 0.4 to 0.5 extra years of schooling.

Accounting for Spatial Autocorrelation Appendix Figure B11 presents inference robustness tests with alternative means to account the spatial autocorrelation. We compute Conley standard errors correcting for spatial autocorrelation using a uniform kernel, which is shown to approach nominal null rejection rates compared to a bartlett kernel (Conley (1999); Bester, Conley and Hansen (2011)). Following the suggestion of Colella et al. (2019), which provides Monte Carlo simulations showing that the null rejection rates approaches nominal rates when the error correction threshold approaches the underlying data generating process, we present Conley standard errors for several distance thresholds, ranging from 25 km to 200 km, using the Stata command `acreg` (Colella et al., 2023). Panels *A* and *B* plot the 90% confidence intervals obtained by correcting the standard errors of the difference-in-difference estimates for spatial autocorrelation in the baseline and the local samples, respectively. Panels *C* and *D* give the 90% confidence intervals adjusted for spatial autocorrelation across predominantly refugee and native birth settlements, conditioning on the current municipality of residence. The coefficients' statistical significance is not sensitive to alternative ways of accounting for spatial autocorrelation.

B.3.5 All Rural Settlements

Difference-in-Difference Estimates Appendix Table B6 reports difference-in-difference specifications, estimated across 7,507 rural settlements (stable geographic units), which associate the various educational outcomes to the share of refugees in 1928 We estimate the regression equation (1) using the share of refugees in 1928 rather than the binary refugee indicator on the RHS. Settlements with a higher share of refugees experienced a more substantial increase in schooling after the Population Exchange. Column (1) suggests that a rise in refugee share of 25 percentage points is associated with an extra 0.1 schooling years and a 1.0–1.5 percentage points higher likelihood of completing primary or junior high school. The estimates with the refugee share imply similar effects to the ones in the paper, where we limit comparisons between those born and living in predominantly refugee and native settlements. Moving from a settlement without any refugees to an exclusive refugee one in the same

province is associated with an increase of about 0.4 years of schooling.

Cond. on Residence Estimates Appendix Table B7 presents results from a variant of regression equation (2), replacing the refugee indicator with the share of refugees in 1928 across all rural birth settlements. So the sample now includes all types of rural birth settlements, predominantly refugee, native and mixed. We do so using the 2011 General Population Census, which, besides residence, also reports an individual's birthplace. Exploring information from Greeks who may have moved from their birth settlement—and conditioning on residence in 2011 yields similarly highly significant estimates of the refugee share on educational attainment.

Dynamics Appendix Figure B12 plots the event-study graphs showing the within-settlement correlation correlation between schooling and refugees' share in 1928 across birth cohorts. Panels *A* and *B* consider years of schooling as the outcome variable, while Panels *C*–*E* focus on the highest completed level of education. Panel *A* plots the cohort-specific estimates on the share of refugees of a parsimonious specification that conditions on settlement fixed effects, interactions between prefecture constants and birth-cohort fixed effects, and interactions between census and birth-cohort fixed effects. The coefficients in Panels *B*–*E* stem from a specification with a rich set of controls (e.g., the interactions of province constants and birth-cohort fixed effects, the interactions of settlement characteristics and birth-cohort fixed effects, and the individual characteristics and birth-cohort fixed effects, etc). The pre-displacement patterns show a drop in primary school attainment for refugee communities in Anatolia that faced Ottoman repression were impacted by WWI and the subsequent Greco-Turkish War. The correlation between human capital and the share of refugees turned positive for refugees arriving in Greece when they were about to start school (cohorts born in 1917–26). The correlation strengthens for Greeks born after that the resettlement was completed (cohorts born post-1927). Panels *C*–*F* illustrate the underlying dynamics. The likelihood of completing primary schooling is positively associated with the share of refugees for Greeks born in the 1920s and 1930s. As the share of Greeks completing primary education approached 100%, the completion rate of junior high school and then high school are higher in settlements with more substantial refugee presence. The correlation between the share of refugees and the completion of university education becomes significant and positive for those born after WWII.

Summary The analysis across all rural settlements reveals a strong association between human capital and the share of refugees in 1928 in the settlement of residence or birthplace. While these estimates may not have a causal interpretation due to mixing, and perhaps other factors related to initial settlement features (selection), they are in line, both qualitatively and quantitatively, with the baseline estimates (in Section 4) based on comparisons of nearby (in the same province and even sub-province level) almost exclusively refugee and native settlements, where the role of geography, location, and local features are muted.

Table B3: Displacement and Years of Schooling in 2011, cond. on Residence and Industry

	(1)	(2)	(3)	(4)	(5)	(6)
Born in a refugee settlement	0.604***	0.531***	0.369***	0.351***	0.352***	0.324***
	(0.162)	(0.167)	(0.105)	(0.114)	(0.097)	(0.105)
R-squared	0.108	0.162	0.363	0.379	0.422	0.434
Observations	154,461	154,453	153,617	153,609	154,395	154,387
Average outcome in native settlements	10.8	10.8	10.8	10.8	10.8	10.8
Cohort × Individual controls	✓	✓	✓	✓	✓	✓
Cohort × Birth settlement controls	✓	✓	✓	✓	✓	✓
Cohort × Birth province FE	✓	✓	✓	✓	✓	✓
Cohort × Industry (1-digit) FE			✓	✓		
Cohort × Industry (3-digit) FE					✓	✓
Residence municipality FE		✓		✓		✓
# Refugee observations	12,148	12,145	12,080	12,077	12,137	12,134
# Refugee birth settlements	353	353	353	353	353	353
# Native observations	142,313	142,308	141,537	141,532	142,258	142,253
# Native birth settlements	5,556	5,556	5,555	5,555	5,556	5,556
# Birth provinces for FE	139	139	139	139	139	139
# Residence-municipality FE		1,008		1,008		1,008

Notes. Both panels report OLS estimates associating years of schooling and a binary variable that takes the value of one for those born in a refugee settlement (where the share of refugees exceeded 80% in 1928) and zero for those born in a native settlement (where the share of refugees in 1928 was below 5%). The sample is restricted to Greek citizens, aged 25–55, who are employed. Columns (1)–(2) consider replicates the baseline estimation presented in Table 2–Columns (1) and (2), Columns (3)–(4) control for the interactions of 1-digit industry codes (N=15) with birth-cohort fixed effects. Columns (5)–(6) control for the interactions of 3-digit industry codes (N=260) with birth-cohort fixed effects. Specifications include interactions of individual characteristics (age and gender) and birth-cohort fixed effects and interactions of birth-settlement characteristics and birth-cohort fixed effects. Specifications in Panel *A* include interactions of the birth-province fixed effects and birth-cohort fixed effects. All specifications in Panel *B* include interactions of the birth-subprovince (Voronoi polygon) fixed effects and birth-cohort fixed effects. Columns (2), (4), and (6) control for municipality-of-residence fixed effects. Birth settlement characteristics are: log distances to the coast, to the provincial capital, to railway stops, and to reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. Standard errors are (double) clustered at the birth-settlement and the municipality-of-residence levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

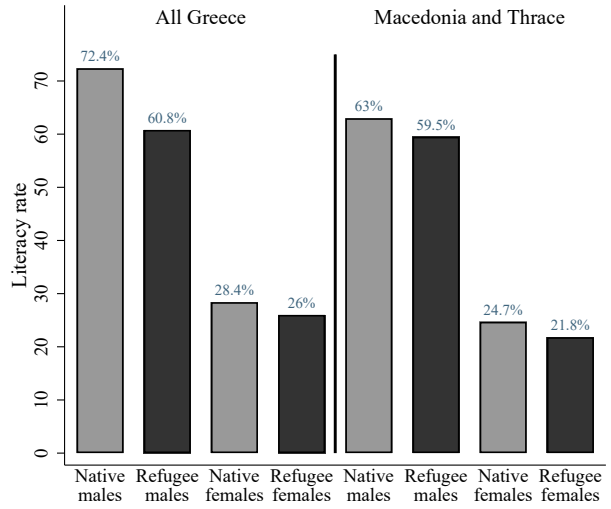
Table B4: Educational Attainment Spillovers to Native Settlements (1971–2011)

Dependent Variable:	Years of schooling			
	All sample	Local sample	All sample	Local sample
	(1)	(2)	(3)	(4)
Refugee share in 1928 within 5 kms × Post	0.002	-0.000	0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Refugee share in 1928 within 5 kms × War cohort	0.000	0.001	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)
Log total population within 5 kms × Post	0.017*	0.017	0.017	0.030
	(0.010)	(0.020)	(0.010)	(0.021)
Log total population within 5 kms × War cohort	0.019*	-0.005	0.018*	-0.001
	(0.011)	(0.019)	(0.011)	(0.020)
Refugee share in 1928 within 5 to 25 kms × Post			0.002	-0.001
			(0.006)	(0.005)
Refugee share in 1928 within 5 to 25 kms × War cohort			-0.002	-0.003
			(0.005)	(0.005)
Log total population within 5 to 25 kms × Post			0.062	0.402***
			(0.040)	(0.146)
Log total population within 5 to 25 kms × War cohort			0.022	0.112
			(0.037)	(0.123)
R-squared	0.458	0.485	0.458	0.485
Observations	1,348,231	266,541	1,348,231	266,541
Average outcome in native settlements: 1887-1906	2.0	1.5	2.0	1.5
Average outcome in native settlements: 1907-1916	2.7	2.1	2.7	2.1
Average outcome in native settlements: 1917-1986	6.4	6.1	6.4	6.1
Settlement FE	✓	✓	✓	✓
Cohort × Census year FE	✓	✓	✓	✓
Cohort × Individual controls	✓	✓	✓	✓
Cohort × Settlement controls	✓	✓	✓	✓
Cohort × Province FE	✓		✓	
Cohort × Sub-province FE		✓		✓
# Native settlements	5,922	1,001	5,922	1,001
# Admin units for FE	137	120	137	120

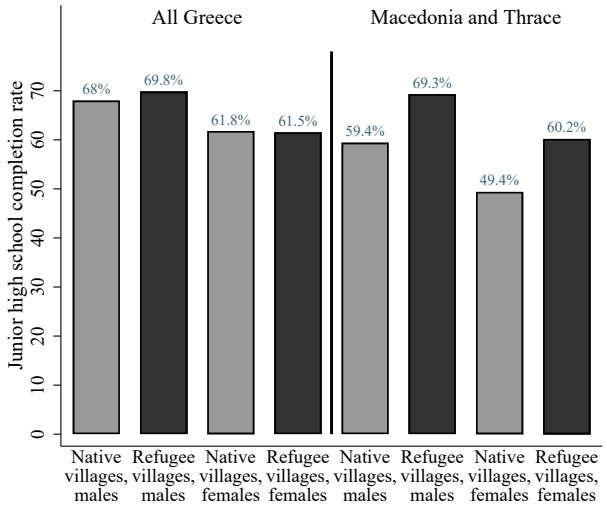
Notes. This table presents the OLS estimates associating the share of refugees in 1928 in the nearby residence settlement to years of schooling across all rural native settlements. *Post* is an indicator for cohorts born after 1917 and that started school in Greece after the Population Exchange. *War cohort* is an indicator for cohorts born between 1907 and 1916 who were of school age during the Balkan Wars, the Great War, and the Greco-Turkish War. All columns control for settlement fixed effects, the interactions of province (*eparchia*) fixed effects and birth-cohort fixed, the interactions of census-year and birth-cohort fixed effects, the interactions of individual characteristics (age and gender) and birth-cohort fixed effects, and the interactions of settlement characteristics and birth-cohort fixed effects. Settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1897–1906 birth-cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Figure B7: Education in 1928 and 2011 across Refugees and Natives by Gender

(a) Literacy in 1928



(b) Junior high school completion in 2011



Notes. Panel *A* plots literacy rates for Greeks, aged 15–64, in 1928 living in the countryside (rural) by gender: refugees (gray bars) and natives (black bars). Literacy rates are retrieved from the 1928 General Population Census. Panel *B* plots the share of Greeks, aged 15–64, who had completed junior high school (9 years of schooling) in 2011 using data from the 2011 Population Census: gray bars across native rural settlements and black bars across predominantly refugee settlements (villages and towns). Both panels report means across Greece (left), and Macedonia and West Thrace (right).

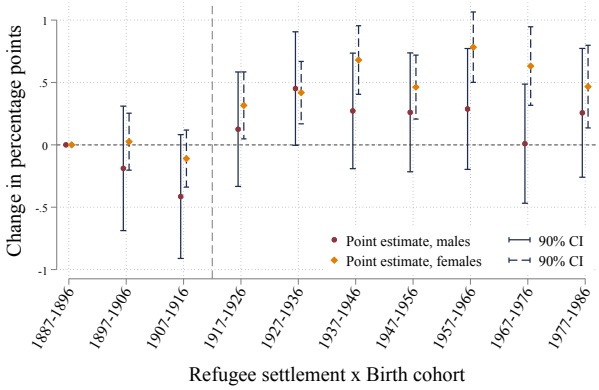
Table B5: Forced Displacement and Gender-Specific Human Capital (1971–2011)

Dependent Variable:	Years of schooling		Completed education level: Primary school		Jun. high school		High school		Tertiary education	
Sample:	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Refugee settlement × Post	0.421***	0.521***	4.210*	3.600***	2.601**	4.880***	2.314***	3.412***	0.515	1.413***
	(0.157)	(0.103)	(2.196)	(1.355)	(1.017)	(0.695)	(0.814)	(0.671)	(0.449)	(0.352)
Refugee settlement × War cohort	-0.265	-0.130	-3.911*	-2.317	-0.281	0.599	-0.302	0.013	-0.311	-0.237
	(0.163)	(0.100)	(2.302)	(1.496)	(0.968)	(0.483)	(0.667)	(0.459)	(0.450)	(0.187)
R-squared	0.359	0.528	0.236	0.411	0.339	0.464	0.239	0.372	0.065	0.121
Observations	737,006	739,464	737,006	739,464	737,006	739,464	737,006	739,464	737,006	739,464
Average outcome in native settlements: 1887-1906	3.3	0.9	49.6	14.0	7.1	0.9	2.0	0.5	0.7	0.2
Average outcome in native settlements: 1907-1916	3.9	1.5	58.0	24.4	9.7	1.6	3.5	0.7	0.9	0.2
Average outcome in native settlements: 1917-1986	7.0	5.7	84.8	70.3	34.0	24.0	22.3	17.9	6.5	5.1
Settlement FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Census year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Settlement controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Province FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
# Refugee observations	62968	65334	62968	65334	62968	65334	62968	65334	62968	65334
# Refugee settlements	359	358	359	358	359	358	359	358	359	358
# Native observations	674,038	674,130	674,038	674,130	674,038	674,130	674,038	674,130	674,038	674,130
# Native settlements	5,897	5,866	5,897	5,866	5,897	5,866	5,897	5,866	5,897	5,866
# Admin units for FE	142	140	142	140	142	140	142	140	142	140

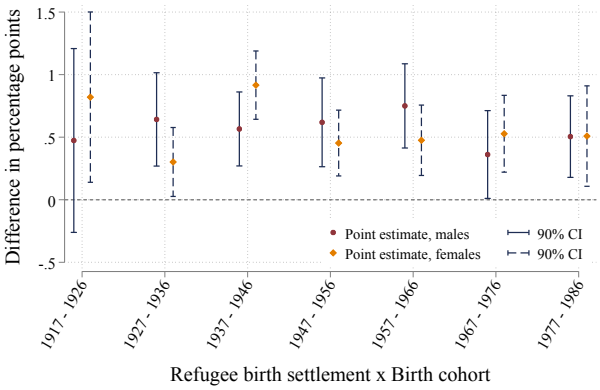
Notes. All specifications report OLS estimates associating various educational outcomes to a binary refugee settlement indicator, *Refugee settlement*, that identifies those residing in predominantly refugee villages where the share of refugees in 1928 exceeded 80%. The binary indicator equals zero for Greeks residing in native settlements, where the refugees' share in 1928 was below 5%. *Post* is an indicator for cohorts born after 1917 and that started school in Greece after the Population Exchange. *War cohort* is an indicator for cohorts born between 1907 and 1916 who were of school age during the Balkan Wars, the Great War, and the Greco-Turkish War. Odd-numbered columns restrict the sample to males, while even-numbered columns restrict it to females. The dependent variable in columns (1)–(2) is years of schooling; in columns (3)–(4) the outcome is an indicator for completing primary school; in columns (5)–(6) an indicator for completing junior high school; in columns (7)–(8) an indicator for completing high school; and in columns (9)–(10) an indicator for completing tertiary education. Odd-numbered columns focus on the sample of all rural refugee and native settlements, conditioning on birth-cohort specific province (*eparchia*) fixed effects. Even-numbered columns focus on the local sample, conditioning on birth-cohort specific sub-province (Voronoi polygon) fixed effects. In each Voronoi polygon, refugee villages and native settlements are less than 25 km away. All specifications include settlement fixed effects, interactions of census-year and birth-cohort fixed effects, interactions of individual characteristics (age and gender) and birth-cohort fixed effects, and interactions of various settlement characteristics with birth-cohort fixed effects. Settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1887–1906 birth cohort serves as the excluded generation. Standard errors are clustered at the settlement-of-residence level. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Figure B8: Forced Displacement and Schooling. Gender-Specific Estimates

(a) Diff-in-Diff (1971-2011)



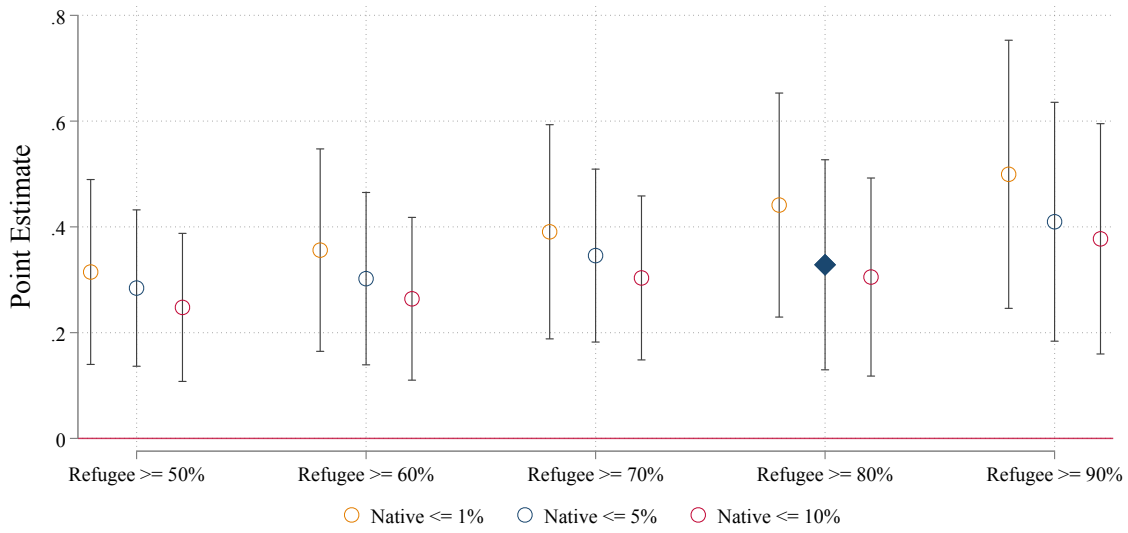
(b) Cond. on Residence (2011)



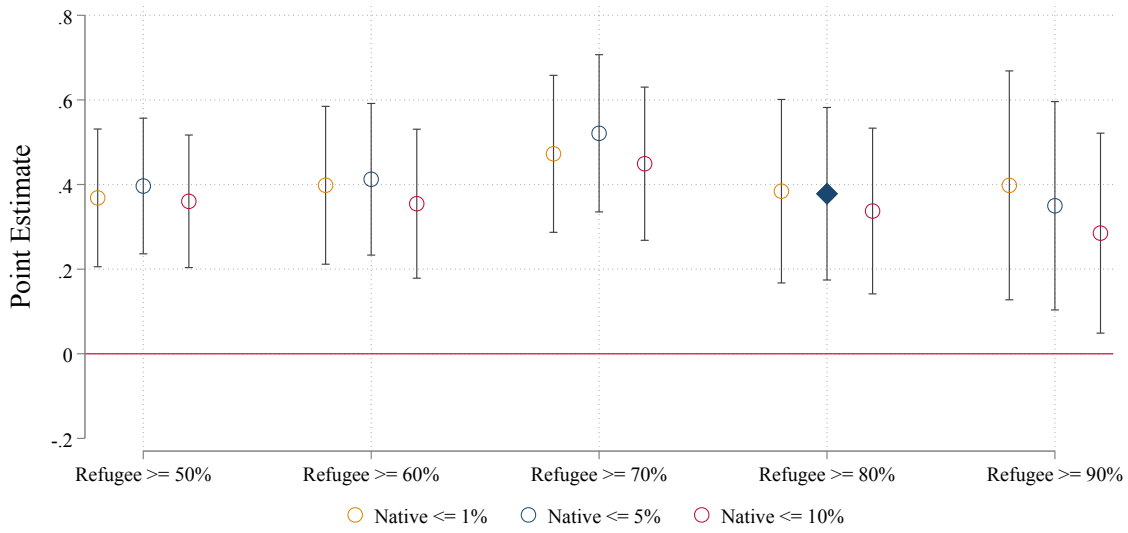
Notes. The panels explore the difference-in-difference estimates (Panel *A*) and the estimates conditioning on residence (Panel *B*) by gender. Refugee settlements are those where the share of refugees in 1928 exceeds 80%, and native settlements are those where the share of refugees from Anatolia in 1928 is below 5%. Maroon markers and solid lines present the estimated coefficients and 90% confidence intervals obtained from restricting the analysis to males, while orange markers and dashed lines present those from restricting the analysis to females. The confidence intervals in Panel *A* are based on standard errors clustered at the municipality-of-residence level, while in Panel *B* are based on standard errors double clustered at the birth-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure B9: Sensitivity Analysis I. Classification Predominantly Refugee and Native Rural Settlements

(a) Diff-in-Diff (1971–2011)



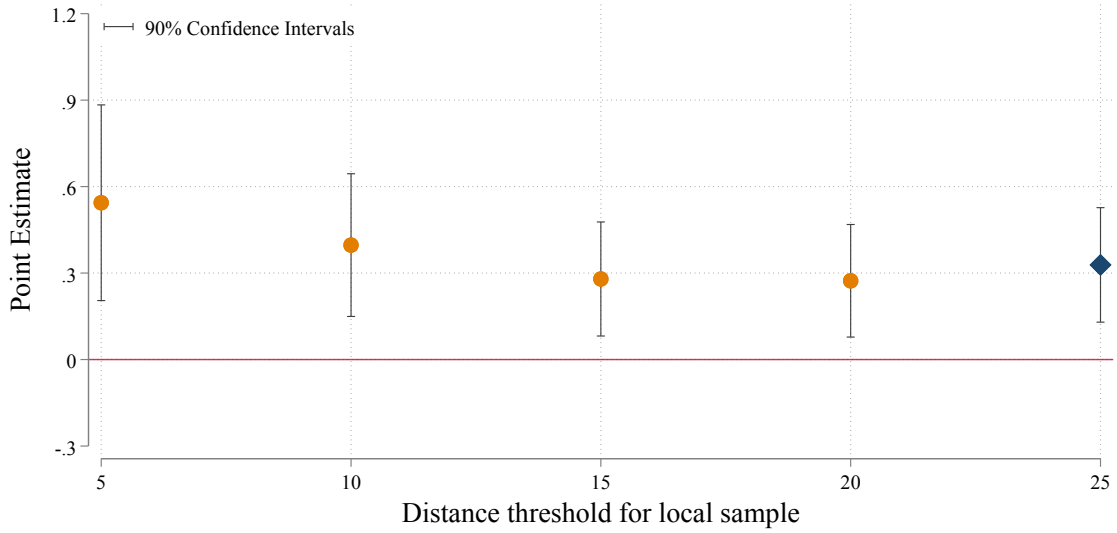
(b) Cond. on Residence (2011)



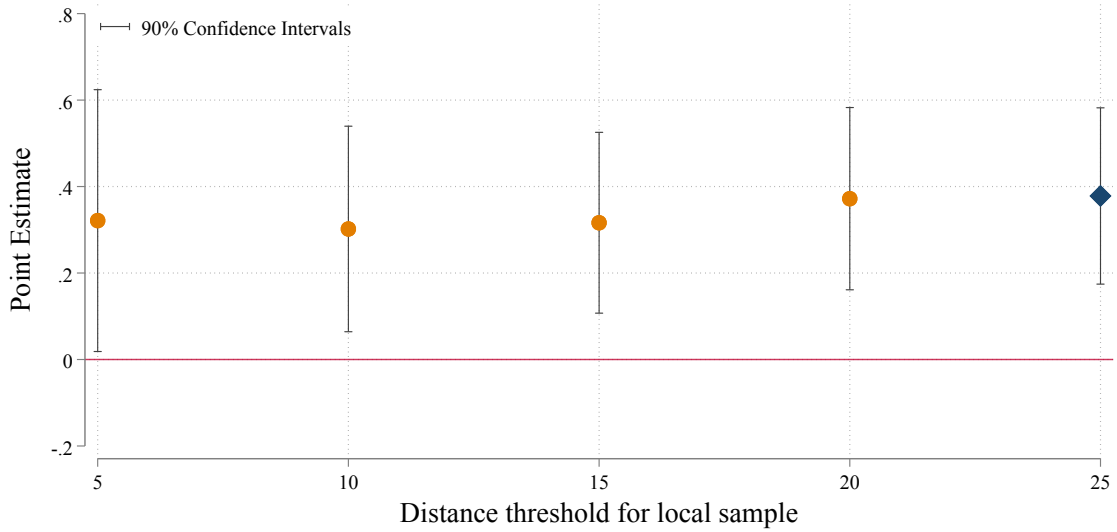
Notes. The panels explore the sensitivity of the difference-in-difference estimates (Panel *A*) and the estimates conditioning on residence (Panel *B*) to changing the definition of (predominantly) refugee and native settlements. Each panel gives 15 coefficients and 90% confidence intervals: refugee settlements are those where the share of refugees in 1928 exceeds 50%, 60%, 70%, 80%, and 90% and native settlements are those where the share of refugees in 1928 is below 1%, 5%, and 10%. The diamond shape in solid blue gives the estimates obtained using the baseline definition of refugee settlements (the share of refugees in 1928 exceeds 80%) and native settlements (the share of refugees in 1928 is below 5%). The confidence intervals in Panel *A* are based on standard errors clustered at the municipality-of-residence level, while in Panel *B* are based on standard errors double clustered at the birth-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro sample from the 2011 General Population Census.

Figure B10: Sensitivity Analysis II. Proximity Refugee and Native Rural Settlements

(a) Diff-in-Diff (1971–2011)



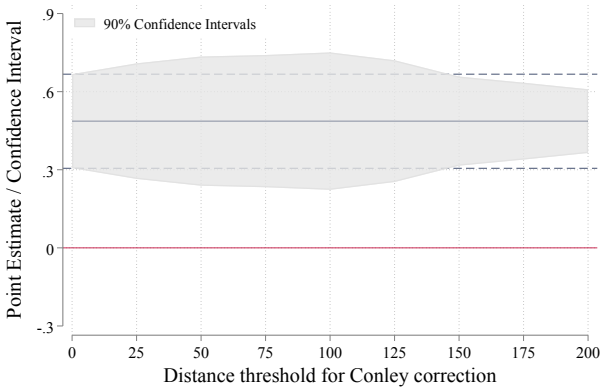
(b) Cond. on Residence (2011)



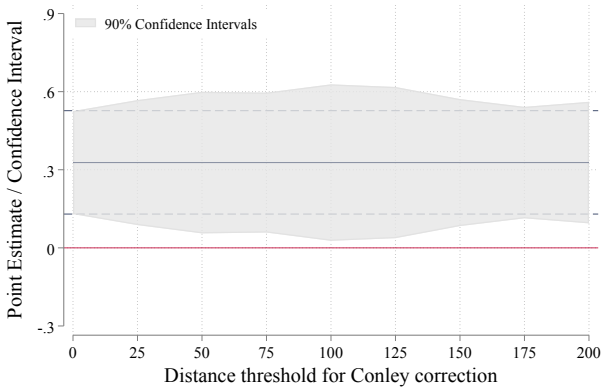
Notes. The panels explore the sensitivity of the difference-in-difference estimates (Panel *A*) and the estimates conditioning on residence (Panel *B*) to changing the local sample: the maximum distance between predominantly refugee and native settlements in the same sub-province unit (Voronoi polygons) in the local sample. Each panel gives 5 coefficients and 90% confidence intervals from alternative distance cutoffs: 25 kilometers (baseline), 20 km, 15 km, 10 km, and 5 km. Refugee settlements are those where the share of refugees in 1928 exceeds 80%, and native settlements are those where the share of refugees in 1928 is below 5%. The confidence intervals in Panel *A* are based on standard errors clustered at the residence-settlement level, while in Panel *B* are based on standard errors double clustered at the birth-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure B11: Sensitivity Analysis III. Conley Standard Errors with Varying Distance Thresholds

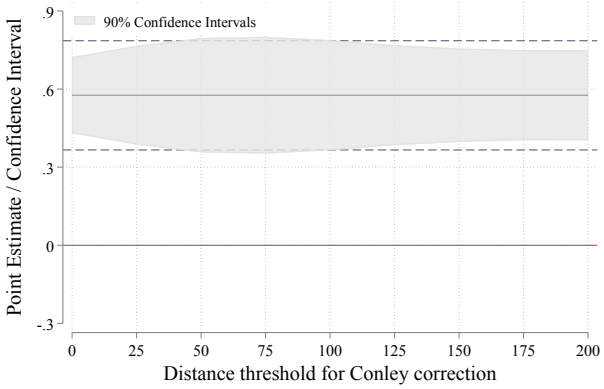
(a) Diff-in-Diff, All Rural Sample (1971–2011)



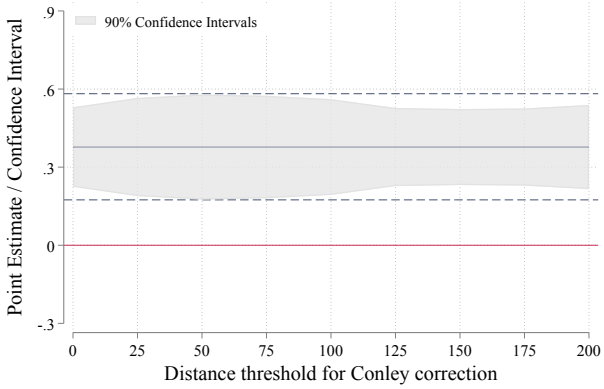
(b) Diff-in-Diff, Local Sample (1971–2011)



(c) Cond. on Residence, All Rural Sample (2011)



(d) Cond. on Residence, Local Sample (2011)



Notes. All panels explore the sensitivity of the difference-in-difference estimates (Panels A and B) and the estimates conditioning on residence (Panels C and D) to correcting the standard errors for spatial autocorrelation for several distance thresholds that vary between 0 and 200 kilometers, increasing incrementally by 25 kilometers. In each reported estimate, the settlements within the indicated distance threshold are assumed to be correlated to each other, while settlements that are farther away are not. The gray horizontal line represents the point estimates. The dashed navy horizontal lines represent the baseline 90% confidence intervals, while the shaded are represents confidence intervals correcting for spatial autocorrelation. Panels A and C consider all rural predominantly refugee and native (birth) settlements in the same province, while Panels B and D restrict the sample to the rural predominantly refugee and native (birth) settlements in the same sub-province unit (Voronoi polygons) within a 25 km radius. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Table B6: Educational Attainment across All Rural Settlements (1971–2011)

Dependent Variable:	Years of schooling	Completed education level: Primary school	Jun. high school	High school	Tertiary education
	(1)	(2)	(3)	(4)	(5)
Share of refugees in 1928 × Post	0.004***	0.042***	0.032***	0.024***	0.006*
	(0.001)	(0.010)	(0.007)	(0.006)	(0.003)
Share of refugees in 1928 × War cohort	-0.002**	-0.035***	0.004	0.002	-0.001
	(0.001)	(0.011)	(0.004)	(0.003)	(0.002)
R-squared	0.464	0.356	0.406	0.305	0.092
Observations	2,123,467	2,123,467	2,123,467	2,123,467	2,123,467
Average outcome in native settlements: 1887-1906	2.0	31.0	3.9	1.2	0.4
Average outcome in native settlements: 1907-1916	2.7	40.7	5.6	2.0	0.5
Average outcome in native settlements: 1917-1986	6.4	77.6	29.0	20.1	5.8
Settlement FE	✓	✓	✓	✓	✓
Cohort × Census Year FE	✓	✓	✓	✓	✓
Cohort × Individual controls	✓	✓	✓	✓	✓
Cohort × Settlement controls	✓	✓	✓	✓	✓
Cohort × Province FE	✓	✓	✓	✓	✓
# Settlements	7,507	7,507	7,507	7,507	7,507
# Admin units for FE	142	142	142	142	142

Notes. This table presents the OLS estimates associating the share of refugees in 1928 in the residence settlement to various educational outcomes across all rural settlements. *Post* is an indicator for cohorts born after 1917 and that started school in Greece after the Population Exchange. *War cohort* is an indicator for cohorts born between 1907 and 1916 who were of school age during the Balkan Wars, the Great War, and the Greco-Turkish War. The dependent variable is: years of schooling in column (1); an indicator for completing primary school in column (2); an indicator for completing junior high school in column (3); an indicator for completing high school in column (4); an indicator for completing tertiary education in column (5). All columns control for settlement fixed effects, the interactions of province (*eparchia*) fixed effects and birth-cohort fixed, the interactions of census-year and birth-cohort fixed effects, the interactions of individual characteristics (age and gender) and birth-cohort fixed effects, and the interactions of settlement characteristics and birth-cohort fixed effects. Settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1897–1906 birth-cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

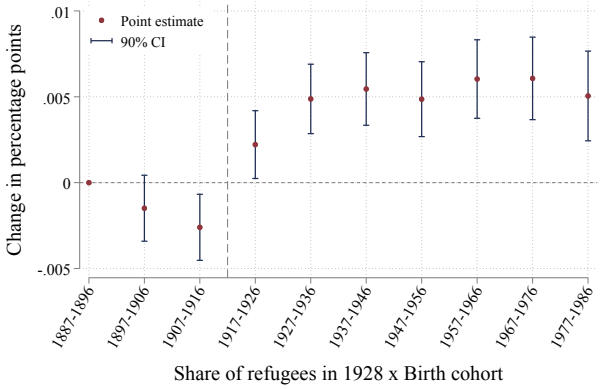
Table B7: Educational Attainment across Rural Birth Settlements, conditional on Residence (2011)

Dependent Variable:	Years of schooling		Completed education level: Primary school		Jun. high school		High school		Tertiary education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share of refugees in 1928, birth settlement	0.005***	0.003***	0.036***	0.026***	0.047***	0.028***	0.031***	0.015*	0.008	0.003
	(0.001)	(0.001)	(0.006)	(0.005)	(0.011)	(0.009)	(0.010)	(0.008)	(0.005)	(0.005)
R-squared	0.358	0.403	0.288	0.302	0.331	0.366	0.233	0.272	0.062	0.090
Observations	761,507	761,501	761,507	761,501	761,507	761,501	761,507	761,501	761,507	761,501
Average outcome in native settlements	8.0	8.0	86.2	86.2	45.6	45.6	33.4	33.4	11.4	11.4
Cohort × Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth settlement controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cohort × Birth province FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Residence municipality FE		✓		✓		✓		✓		✓
# Refugee observations	44,115	44,114	44,115	44,114	44,115	44,114	44,115	44,114	44,115	44,114
# Refugee birth settlements	357	357	357	357	357	357	357	357	357	357
# Native observations	505,365	505,361	505,365	505,361	505,365	505,361	505,365	505,361	505,365	505,361
# Native birth settlements	5,788	5,788	5,788	5,788	5,788	5,788	5,788	5,788	5,788	5,788
# Birth provinces for FE	141	141	141	141	141	141	141	141	141	141
# Residence-municipality FE		1,022		1,022		1,022		1,022		1,022

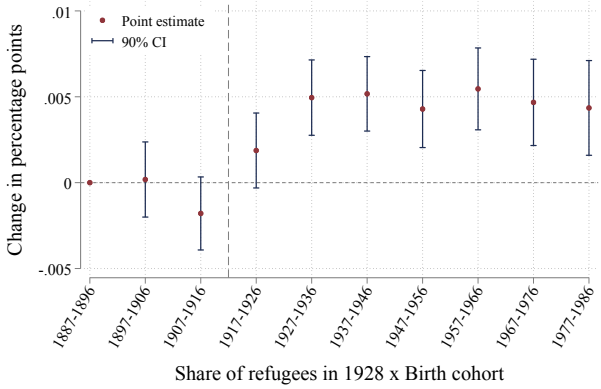
Notes. This table presents the OLS estimates associating the 1928 share refugees in the birth settlement to educational attainment across rural settlements. The dependent variable is years of schooling in columns (1)–(2), primary school completion in columns (3)–(4), junior high school completion in columns (5)–(6), high school completion in columns (7)–(8), tertiary education completion in columns (9)–(10). All reported regressions control for the interactions of the birth-province fixed effects and birth-cohort fixed effects, the interactions of birth-settlement characteristics and birth-cohort fixed effects and the interactions of individual characteristics (age and gender) and birth-cohort fixed effects. Columns (2), (4), (6), (8), and (10) control for municipality-of-residence fixed effects. Birth settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. Standard errors are clustered at the birth-settlement and the municipality-of-residence levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1%, confidence level.

Figure B12: Human Capital and Refugee Share across All Rural Settlements (1971–2011)

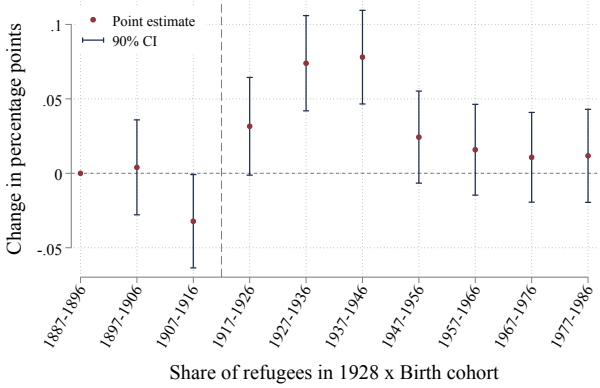
(a) Years of schooling, Prefecture FEs



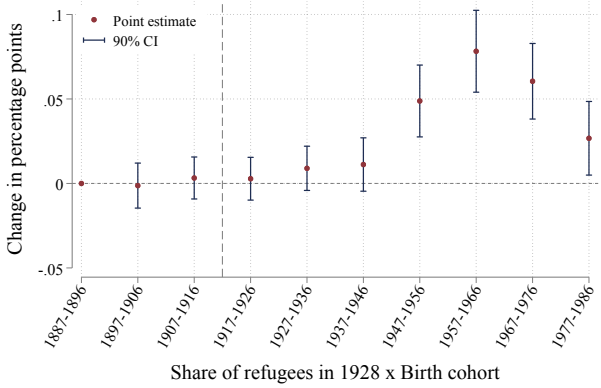
(b) Years of schooling, Full specification



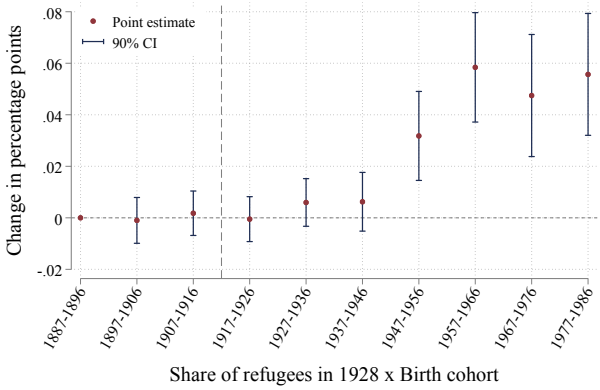
(c) Primary school completion



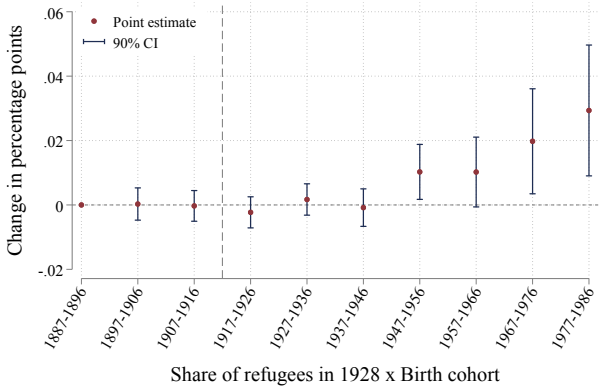
(d) Junior high school completion



(e) High school completion



(f) Tertiary education



Notes. The panels present the event-study estimates obtained from regressing educational outcomes on the share of refugees in 1928 across all rural settlements. Outcome variable in consideration is years of schooling in Panels *A* and *B*; completing primary school in Panel *C*, junior high school in Panel *D*, high school in Panel *E*, and tertiary education in Panel *F*. Panel *A* controls for settlement fixed effects, the interactions between prefecture constants and birth-cohort fixed effects, and the interactions between census-year and birth-cohort fixed effects. Panels *B*-*F* controls for the interactions of province constants and birth-cohort fixed effects, the interactions of settlement characteristics and birth-cohort fixed effects, and the individual characteristics and birth-cohort fixed effects in addition. Settlement characteristics are: log distance to the coast, the provincial capital, railway stops, and reclaimed marshlands; geographic characteristics: crop suitability, altitude, precipitation, and temperature; and the log population in 1920, an indicator for settlements with the monastery in 1920. The 1887–1896 birth-cohort serves as the excluded cohort. Standard errors are clustered at the settlement-of-residence level.

C Human Capital Portability

This Appendix Section reports descriptive statistics and sensitivity checks on the link between forced displacement and subsequent human capital investments in portable skills among college graduates, complementing the analysis in Section 5 of the main paper.

Degree Portability Appendix Table C1 lists the 50 most popular degrees in the Greek tertiary education system, tabulating the 2011 Population Census (sample 25%) alongside our classification of the skills' transferability based on Chat-GPT 4. The table gives the tabulations for the country and the rural areas in our baseline sample of predominantly refugee and native settlements, respectively.

Descriptive Patterns Appendix Figure C1 presents the average skill portability of tertiary education degrees among individuals born in native and refugee settlements. The first two sets of bars show that those from refugee settlements are more likely to complete a transferable degree by 3.5 percentage points in the baseline sample and by 5 percentage points in the local sample, where refugee and native villages are within 25 km radius in the same province. The last three sets of bars present the relative distribution of lawyers, doctors, and engineers among these professions. Lawyers are overrepresented among individuals from native settlements, while the share of doctors and engineers is significantly higher among those from predominantly refugee settlements.

Sensitivity Appendix Figure C2 reports sensitivity checks of the association between the portability of tertiary degree skills and being born in a predominantly refugee settlement, reported in Table 3. Panel *A* reports 15 specification coefficients altering the cutoff of refugees' share in 1928 to classify predominantly refugee and native settlements. Across all permutations, Greeks born in majority or predominantly refugee settlements are considerably more likely to complete tertiary education with portable skills. Panel *B* gives five permutations of the baseline specification linking an indicator of skill portability of university education to the refugee indicator that takes the value of one for an individual whose birthplace is a predominantly refugee settlement (refugee share in 1928 exceeds 80%) and zero for Greeks born in native settlements, where the share of refugees in 1928 is below 5% in the local sample. The panel gives estimates restricting comparisons between native and refugee settlements in the same province (*eparchia*) no more than 25 km, 20 km, 15 km, 10 km, and 5 km away. The estimates are stable even though the sample drops.

Spatial Autocorrelation Appendix Figure C3 explores inference accounting for spatial correlation, using Conley standard errors with a uniform kernel (Conley (1999) and Bester, Conley and Hansen (2011)). Following the suggestion of Colella et al. (2019), we present Conley standard errors for several distance thresholds: ranging from 25 km to 200 km, increasing by 25 km in each iteration, using the Stata command `acreg` (Colella et al., 2023). Panels *A* and *B* plot the 90% confidence intervals adjusted for spatial autocorrelation, conditioning on the municipality of residence across all rural

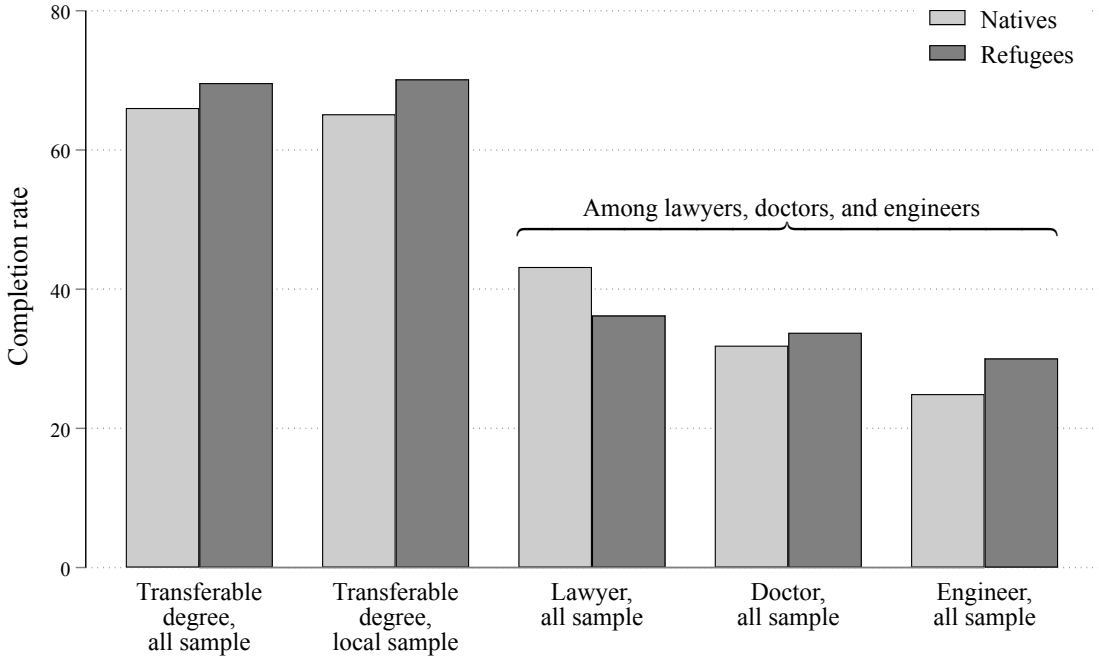
predominantly refugee and native birth settlements and those in the local sample where refugee and native birth settlements are within 25 km in the same province, respectively. The estimates' statistical significance is not sensitive to the particular cutoff used in the spatial autocorrelation adjustment.

Table C1: Top 50 Most Popular College Degrees and Skill Portability (2011)

		All Greece			Main rural sample		
Study field	Transferable Degree	Number	Share	Cumulative share	Number	Share	Cumulative share
Economic Sciences	1	23,469	6.9	6.9	4,070	6.6	6.6
Law	0	19,209	5.7	12.6	3,069	4.9	11.5
Primary Education Teachers	0	18,400	5.4	18.1	6,167	9.9	21.5
Business Administration	1	16,736	5.0	23.0	2,815	4.5	26.0
Medicine	1	15,450	4.6	27.6	2,304	3.7	29.7
Accounting	1	10,632	3.1	30.7	2,081	3.4	33.1
Civil Engineering	1	10,530	3.1	33.9	1,663	2.7	35.7
Mathematics	1	8,360	2.5	36.3	1,754	2.8	38.6
Greek Literature	0	8,303	2.5	38.8	1,713	2.8	41.3
Nursing	1	8,126	2.4	41.2	2,013	3.2	44.6
Physical Education and Sport Science	1	7,906	2.3	43.5	1,473	2.4	47.0
Military Officers	0	6,007	1.8	45.3	1,582	2.6	49.5
Physics	1	5,915	1.8	47.1	1,077	1.7	51.2
English Language and Literature	1	5,647	1.7	52.1	656	1.1	52.3
Electrical and Computer Engineering	1	5,754	1.7	48.8	746	1.2	53.5
Kindergarten Teachers	0	5,662	1.7	50.4	1,353	2.2	55.7
Mechanical Engineering	1	5,253	1.6	55.3	689	1.1	56.8
Engineering	1	5,505	1.6	53.7	1,107	1.8	58.6
Information Technology	1	4,941	1.5	56.8	572	0.9	59.5
Architectural Engineers	1	4,770	1.4	58.2	406	0.7	60.2
Political Sciences	1	4,657	1.4	59.5	1,091	1.8	61.9
Philosophy	1	4,465	1.3	62.2	923	1.5	63.4
Electrical Engineering	1	4,522	1.3	60.9	954	1.5	64.9
Agronomist	0	4,194	1.2	63.4	1,025	1.7	66.6
Dentistry	1	3,719	1.1	66.8	561	0.9	67.5
Theology	0	3,884	1.1	64.6	1,024	1.7	69.2
Pharmacists	1	3,781	1.1	65.7	626	1.0	70.2
French Language and Literature	1	3,487	1.0	67.8	249	0.4	70.6
Tourism Enterprises	1	3,229	1.0	68.8	456	0.7	71.3
Chemical Engineers	1	2,890	0.9	72.4	387	0.6	71.9
Chemistry	1	3,107	0.9	70.6	491	0.8	72.7
Policemen	0	2,953	0.9	71.5	743	1.2	73.9
History and Archaeology	1	3,135	0.9	69.7	493	0.8	74.7
Psychology	1	2,758	0.8	73.2	253	0.4	75.1
Business Organisation and Management	1	2,211	0.7	74.5	297	0.5	75.6
Sociology	1	2,272	0.7	73.9	393	0.6	76.2
Biology	1	1,966	0.6	76.4	253	0.4	76.6
Public Administration	0	2,163	0.6	75.8	525	0.8	77.5
Electronics	1	2,171	0.6	75.2	379	0.6	78.1
Physiotherapy	1	1,618	0.5	78.9	240	0.4	78.5
Infant and Child Care	0	1,696	0.5	77.4	343	0.6	79.0
Naval Officers	0	1,665	0.5	77.9	273	0.4	79.5
Plant Production	0	1,737	0.5	76.9	480	0.8	80.3
Fine Arts	1	1,644	0.5	78.4	140	0.2	80.5
Midwives	1	1,599	0.5	79.3	375	0.6	81.1
Accounting and Finance	1	1,506	0.4	79.8	286	0.5	81.5
Social Work	1	1,467	0.4	80.2	339	0.5	82.1
Electronic Computing Systems	1	1,426	0.4	80.6	255	0.4	82.5
Geology	1	1,405	0.4	81.5	197	0.3	82.8
Medical Laboratories	1	1,421	0.4	81.1	289	0.5	83.3

Notes. This table lists the 50 most popular degrees studied in tertiary education along with their skill portability status and prevalence. The degrees are sorted in descending order by their popularity.

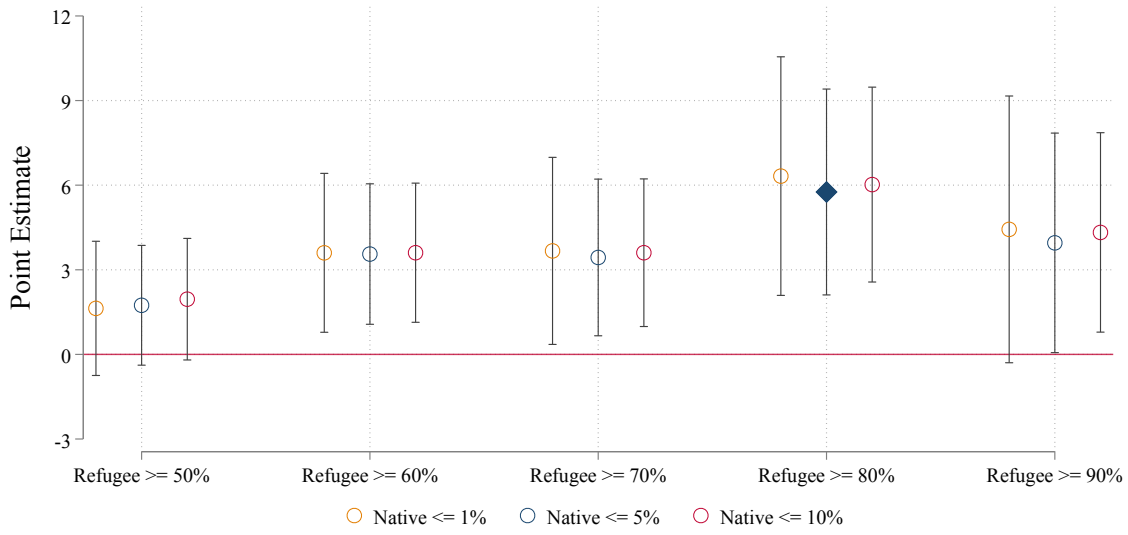
Figure C1: Skill Transferability across Refugee and Native Settlements (2011 Census)



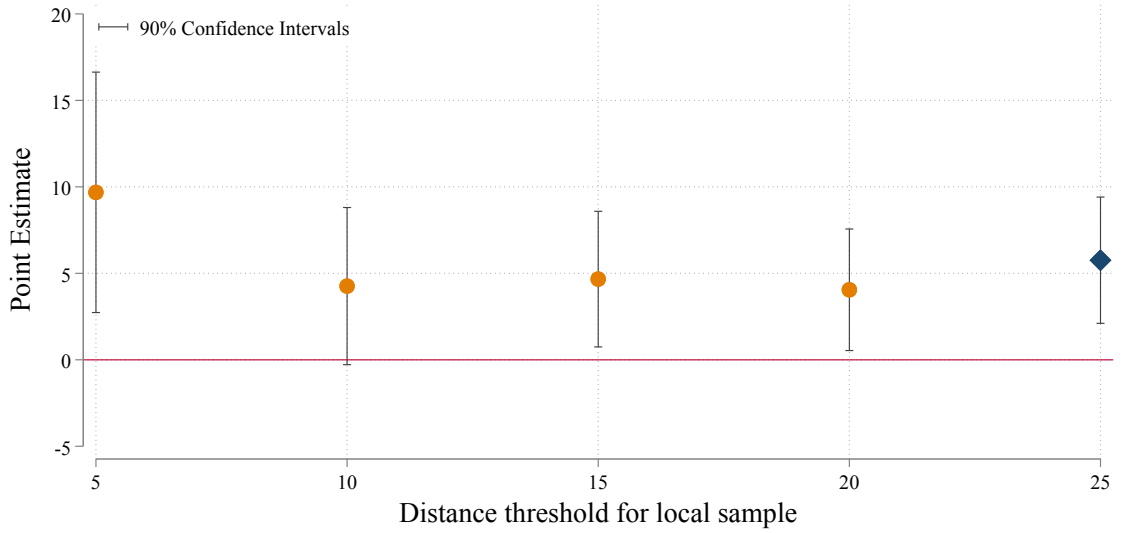
Notes. This figure presents the average skill transferability of tertiary degrees across refugee and native birth settlements among those with a tertiary education. It also shows the relative distribution of lawyers, doctors, and engineers among those with a degree in law, medicine, and engineering (excluding civil engineering) from universities and polytechnics across all rural refugee and native birth settlement. All sample stands for all rural predominantly refugee and native birth settlements; local sample focuses on a sample where native and refugee birth settlements are within a 25 km radius in the same province.

Figure C2: Sensitivity Checks. Forced Displacement and Skill Portability (2011)

(a) Classification of Native and Refugee Settlements



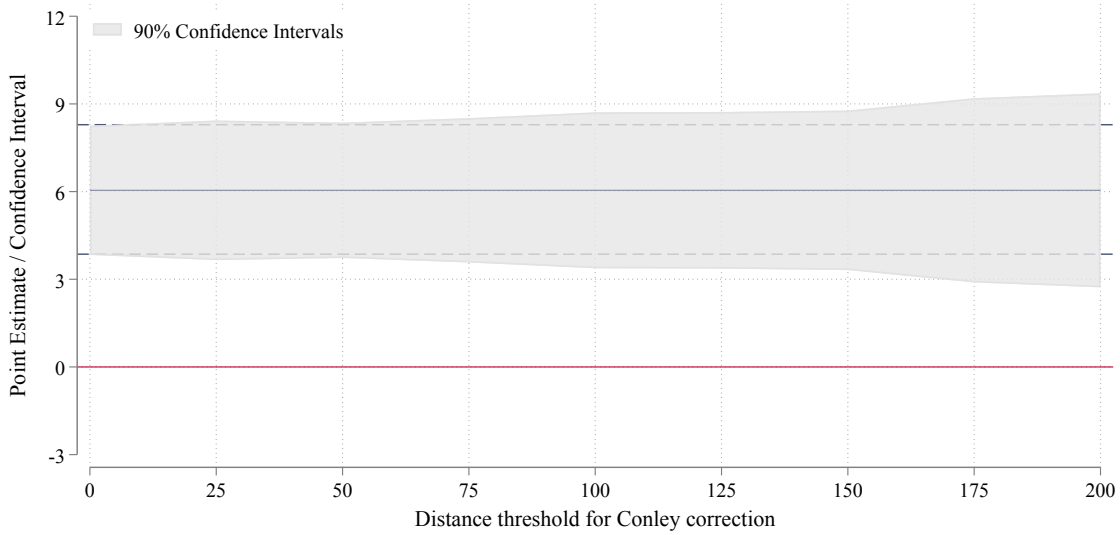
(b) Distance Threshold. Local Sample Analysis



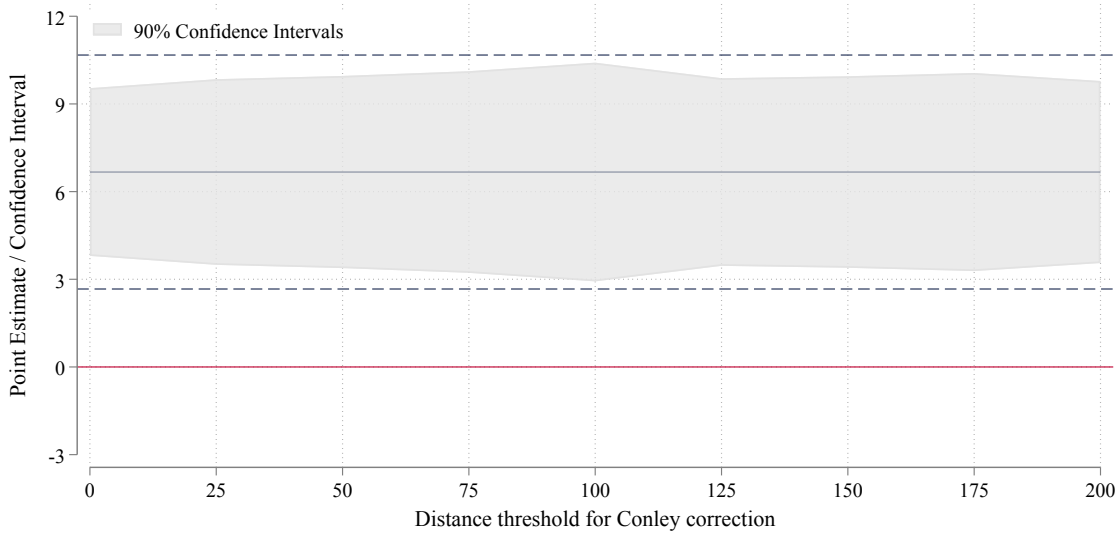
Notes. Panel *A* presents 15 estimates changing the definition of (predominantly) refugee and native settlements. Panel *B* plots 5 estimates changing the maximum distance between refugee and native birth settlements to build the sub-province clusters in the local analysis sample. In all OLS specifications, the dependent variable is an indicator for tertiary education degrees with portable skills. The diamond shape in solid blue presents the coefficient estimates, from Table 3–column (4), based on the baseline definition of refugee birth settlements (share of refugees in 1928 exceeds 80%) and native birth settlements (share of refugees in 1928 is below 5%), and the baseline definition of local sample, where refugee birth settlements are matched to native ones no more than 25 km away in the same sub-province unit. The round hollow shapes in Panel *A* represent the estimates obtained from using alternative definitions of refugee settlements (share of refugees in 1928 exceeds 50%, 60%, 70%, 80%, and 90%) and native settlements (share of refugees in 1928 is below 1% and 10%). The round orange shapes in Panel *B* represent the estimates obtained using alternative distance thresholds: 5 km, 10 km, 15 km, 20 km, and 25 km. The two panels also give 90% confidence intervals based on standard errors double clustered at the birth-settlement and the municipality-of-residence levels. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure C3: Sensitivity Analysis III. Conley Standard Errors with Varying Distance Thresholds

(a) All sample



(b) Local analysis sample



Notes. The figure explores the sensitivity of the estimates conditioning on residence to correcting the standard errors for spatial autocorrelation for several distance thresholds that vary between 0 and 200 kilometers, increasing incrementally by 25 kilometers. In each reported estimate, the settlements within the indicated distance threshold are assumed to be correlated to each other, while settlements that are farther away are not. The gray horizontal line represents the point estimates. The dashed navy horizontal lines represent the baseline 90% confidence intervals, while the shaded are represents confidence intervals correcting for spatial autocorrelation. Panels *A* considers all rural predominantly refugee and native birth settlements in the same province (Table 3–column (2)), while Panel *B* restricts the sample to the predominantly refugee and native birth settlements in the same sub-province unit (Voronoi polygons) within a 25 km radius (Table 3–column (4)).

D Mechanisms: Additional Results

This Appendix Section complements the heterogeneity analysis that delves into additional mechanisms linking the initial forced displacement experience to subsequent human capital investments (in Section 6 of the paper). The appendix gives additional results, summary statistics, and descriptive evidence across all aspects we explore heterogeneity.

Refugees' Settlement Characteristics Appendix Table D1 gives summary statistics (mean, median, max, min, standard deviation, the 10th and the 90th percentile) of all demographic, geographic, locational, and ethnographic variables we use in the heterogeneity analysis. Panel *A* reports summary statistics across 384 or 405 predominately refugee settlements, defined as those where the share of refugees in 1928 was above 80% of the total population. Panel *B* gives the statistics weighted by settlements' population in 1928.

Displacement Trajectory by Origin Region The Greek-Orthodox people from Smyrna (Izmir), Ionia, and the coastal cities had to flee in chaotic conditions, given the swift advance of Turkish troops in September 1922. The Eastern Thrace Greeks abandoned their homes in the fall of 1923, crossing the Evros River to Western Thrace, but in a relatively organized manner, carrying movable property and animals. The Cappadocia-Cilicia communities left their homes after the Convention for the Population Exchange but without the ability to carry assets. As part of the voluntary population exchange, the Greek-Orthodox communities from Bulgaria, unaffected by the Greco-Turkish wars, moved relatively peacefully south with their belongings. Pontic Greeks, living on the southern shores of the Black Sea and the Caucasus, faced labor camp internments and large-scale massacres. Those who survived arrived in 1923 and several even later. Appendix Figure D5 plots the share of refugees by origin as recorded in the 1928 General Population Census and the Rural Refugee Catalog.

Heterogeneity by Educational Level Appendix Figures D1–D4 present the heterogeneity analysis for different levels of educational attainment: primary school completion, junior high and high school completion, and tertiary education. It complements the baseline results in the main paper using schooling years.

Origin Shares Appendix Figure D5–Panel *A* plots the share of refugees by origin as recorded in the 1928 General Population Census for the entire population of the country and for rural areas separately. Panel *B* gives the distribution based on the Rural Refugee Catalog for all rural settlements that received at least one land grant and for rural settlements in the local sample.

Origin Distribution at Settlement Level. Mappings Appendix Figure D6 maps the share of refugees from Asia Minor and Cappadocia (Panel *A*), Pontus and Thrace (Panel *B*), Eastern Thrace (Panel *C*), and Constantinople (Panel *D*) across 141 provinces based on the breakdown provided by

the 1928 General Population Census. The four maps also overlay rural settlements where majority of land grant recipients originated from each region, based on the Rural Refugee Catalog compiled by the RSC in the mid-1920s.

Balance Test. Cappadocian Settlements vs. Other Origin Settlements Appendix Figure D7 explores differences in geographic, locational, and ecological features between refugee settlements where the majority of the population is from Cappadocia and other refugee settlements (with a non-Cappadocian majority). The two types of refugee settlements are rather balanced in terms of their characteristics. Once we condition on prefecture (level-2 admin unit) and province (level-3 admin unit) fixed effects, there are no statistically significant differences.

Event-Study by Origin Region Appendix Figure D8 plots the cohort-specific estimates of the refugee indicators distinguishing the main origin region of refugees settled with a land grant (Asia Minor, Cappadocia, Pontus and Caucasus, East Thrace, and Bulgaria) to quantify the evolution of refugee-native differences in schooling . Panel A gives the difference-in-difference estimates that compare the evolution of schooling years of Greeks in refugee and native settlements after the forced displacement to the pre-population exchange, pooling together the micro samples of all post-1971 General Population Censuses. Panel B gives conditional on residence estimates that compare Greeks born in predominantly refugee and native settlements using the 2011 General Population Census. While the estimates are somewhat noisy, the analysis clearly shows that Greeks born in refugee settlements where the majority in 1928 were from Cappadocia, who did not speak Greek, do not experience an educational gain. In contrast, the estimates for second-, third-, and fourth-generation refugees from Asia Minor, Cappadocia, Pontus and Caucasus, Pontus, East Thrace, and mixed settlements are positive despite some fluctuations.

Balance Test. New Refugee Settlements vs. Former-Muslim Refugee Settlements Appendix Figures D9 explore differences in geographic, locational, and ecological features between the former-Muslim and predominantly new refugee settlements in the baseline sample. New refugee settlements were somewhat more populated in 1928, closer to the coast and marshlands and at a lower elevation. However, once conditioning on (log) altitude, differences become muted.

New and Former-Muslim Refugee Settlements. Population Over Time Appendix Figure D10 plots the evolution of the population in the two types of rural refugee settlements, former Muslim and new ones, and native villages and towns between 1907 and 2011, as recorded in the respective Censuses. Panel A gives the evolution of the median population, along with 90th- and 10th-percentiles for all rural and native refugee settlements, while Panel B focuses on the native and refugee settlements in the local analysis sample.

Table D1: Summary Statistics across Settlements

	Mean	SD	Min	10p	Median	90p	Max	Obs.
Panel A:	**Settlement characteristics**							
Majority of refugees from Constantinople (%)	0.3	5.1	0.0	0.0	0.0	0.0	100.0	384
Majority of refugees from Asia Minor, coastal (%)	19.5	39.7	0.0	0.0	0.0	100.0	100.0	384
Majority of refugees from Cappadocia (%)	7.0	25.6	0.0	0.0	0.0	0.0	100.0	384
Majority of refugees from Pontus and Caucasus (%)	38.3	48.7	0.0	0.0	0.0	100.0	100.0	384
Majority of refugees from Thrace (%)	24.5	43.1	0.0	0.0	0.0	100.0	100.0	384
Majority of refugees from Bulgaria (%)	1.8	13.4	0.0	0.0	0.0	0.0	100.0	384
No majority region of origin (%)	8.6	28.1	0.0	0.0	0.0	0.0	100.0	384
Share of refugees from province centres in Turkey (%)	5.0	10.1	0.0	0.0	1.7	11.4	95.3	384
Origin diversity at settlement	0.7	0.2	0.0	0.3	0.7	0.9	1.0	384
Origin (province) diversity at settlement	0.5	0.2	0.0	0.1	0.5	0.8	0.9	384
Former-Muslim settlement (%)	71.9	45.0	0.0	0.0	100.0	100.0	100.0	405
Population in 1928	449.1	488.5	9.0	73.0	285.0	971.0	4,041.0	405
Agricultural suitability	27.5	10.6	11.5	16.1	25.8	40.2	71.0	405
Agricultural similarity to origin	85.2	6.2	61.3	78.1	86.0	92.2	99.3	384
Agricultural-suitability gain	-6.2	15.5	-37.4	-23.0	-8.8	14.6	60.3	384
Distance to province capital	15.6	7.8	1.9	6.5	14.1	27.7	39.0	405
Population density in 1920, province level	31.4	23.3	13.3	14.1	27.0	37.4	235.5	405
Share of manufacturing in 1920, province level	12.6	7.3	3.7	5.6	12.1	21.4	35.6	405
Share of commerce in 1920, province level	4.6	2.8	1.2	1.7	4.2	8.0	17.4	405
Share of public administration in 1920, province level	1.6	1.1	0.4	0.5	1.4	2.8	7.6	405
Number of teachers per 100 school-age children, province level	1.8	0.3	0.9	1.4	1.8	2.1	3.5	405
Panel B:	**Settlement characteristics, population weighted**							
Majority of refugees from Constantinople (%)	0.2	4.1	0.0	0.0	0.0	0.0	100.0	177,974
Majority of refugees from Asia Minor, coastal (%)	22.4	41.7	0.0	0.0	0.0	100.0	100.0	177,974
Majority of refugees from Cappadocia (%)	6.1	23.9	0.0	0.0	0.0	0.0	100.0	177,974
Majority of refugees from Pontus and Caucasus (%)	26.5	44.1	0.0	0.0	0.0	100.0	100.0	177,974
Majority of refugees from Thrace (%)	26.9	44.4	0.0	0.0	0.0	100.0	100.0	177,974
Majority of refugees from Bulgaria (%)	4.6	20.9	0.0	0.0	0.0	0.0	100.0	177,974
No majority region of origin (%)	13.3	34.0	0.0	0.0	0.0	100.0	100.0	177,974
Share of refugees from province centres in Turkey (%)	4.4	7.5	0.0	0.0	2.1	9.6	95.3	177,974
Origin diversity at settlement	0.7	0.3	0.0	0.2	0.8	0.9	1.0	177,974
Origin (province) diversity at settlement	0.6	0.3	0.0	0.1	0.6	0.8	0.9	177,974
Former-Muslim settlement (%)	66.1	47.3	0.0	0.0	100.0	100.0	100.0	181,882
Population in 1928	979.1	826.6	9.0	213.0	734.0	2,088.0	4,041.0	181,882
Agricultural suitability	27.0	9.7	11.5	16.1	26.4	36.4	71.0	181,882
Agricultural similarity to origin	85.9	5.3	61.3	80.0	86.0	91.8	99.3	177,974
Agricultural-suitability gain	-7.4	13.6	-37.4	-25.0	-8.8	10.2	60.3	177,974
Distance to province capital	14.1	7.5	1.9	6.2	12.3	25.6	39.0	181,882
Population density in 1920, province level	34.2	27.6	13.3	16.5	29.0	37.4	235.5	181,882
Share of manufacturing in 1920, province level	12.9	7.7	3.7	4.6	12.1	21.4	35.6	181,882
Share of commerce in 1920, province level	4.7	3.0	1.2	1.7	4.2	9.4	17.4	181,882
Share of public administration in 1920, province level	1.6	1.2	0.4	0.5	1.4	2.8	7.6	181,882
Number of teachers per 100 school-age children, province level	1.7	0.3	0.9	1.3	1.7	2.1	3.5	181,882

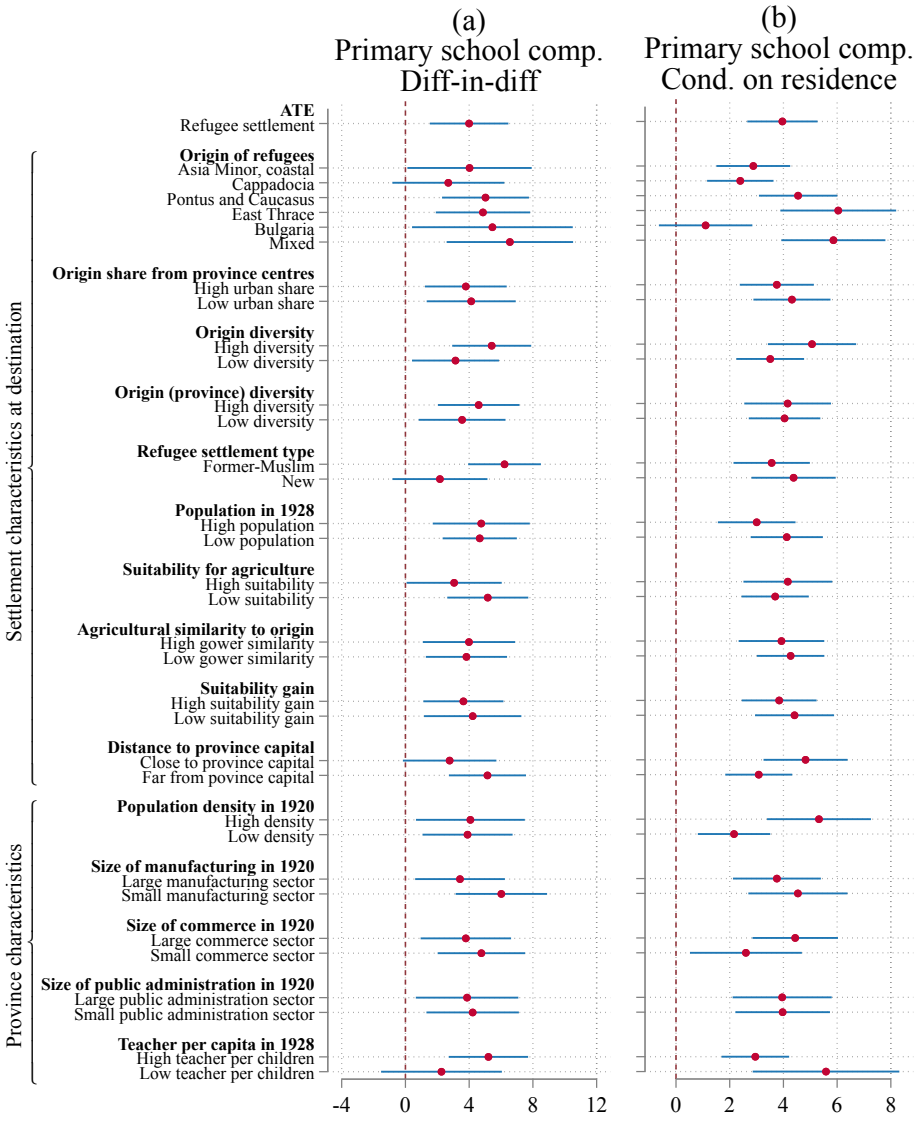
Notes. The table gives summary statistics of various ethnographic, geographic, and locational features across predominantly refugee settlements, defined as those where the share of refugees in 1928 exceeded 80%. Panel *A* gives unweighted summary statistics across all rural predominantly refugee settlements. Panel *B* gives statistics weighting the settlements by their population in 1928.

Table D2: Refugee Origin Settlement Characteristics and Displacement Trajectories

	Constantinople	**Eastern Thrace**	**Asia Minor**	**Pontus**	**Cappadocia**
Settlement	Mainly in Athens connections with Greek cities (trading)	West Thrace	Aegean Sea Islands Athens Sterea Elada	Rural Macedonia, Thessaloniki	Central Northern Greece
Trajectory	Peacefully, transferring assets Trade, Services	Cross Evros Fall 1922 carry movables and animals	Repression since 1908 Exodus, Fire of Smyrna Attacks by Turkish Army	Post Population Exchange Convention (1923); Some via Caucasus, Russia	Post Population Exchange Convention mid/late 1923 & 1924
Main Activity					
Agriculture	0.43	0.71	0.63	0.71	0.71
Herding	0.00	0.00	0.02	0.09	0.03
Fishing	0.29	0.00	0.08	0.00	0.00
Merchants	0.29	0.14	0.20	0.10	0.21
Mining	0.00	0.14	0.08	0.04	0.05
Main Agriculture					
Tobacco	0.00	0.00	0.12	0.21	0.00
Cereal	0.40	0.33	0.29	0.59	0.68
Silk Cocoons	0.00	0.33	0.39	0.01	0.00
Vineyards	0.00	0.33	0.21	0.02	0.32
Language					
Turkish Spoken	0.26	0.25	0.37	0.49	0.75
Only Turkish	0.00	0.00	0.14	0.31	0.57
Only Pontic	0.00	0.12	0.08	0.26	0.04
Only Greek	0.74	0.62	0.53	0.21	0.00
Only Cappadocian	0.00	0.00	0.00	0.00	0.20
# Settlements	19	8	333	250	75

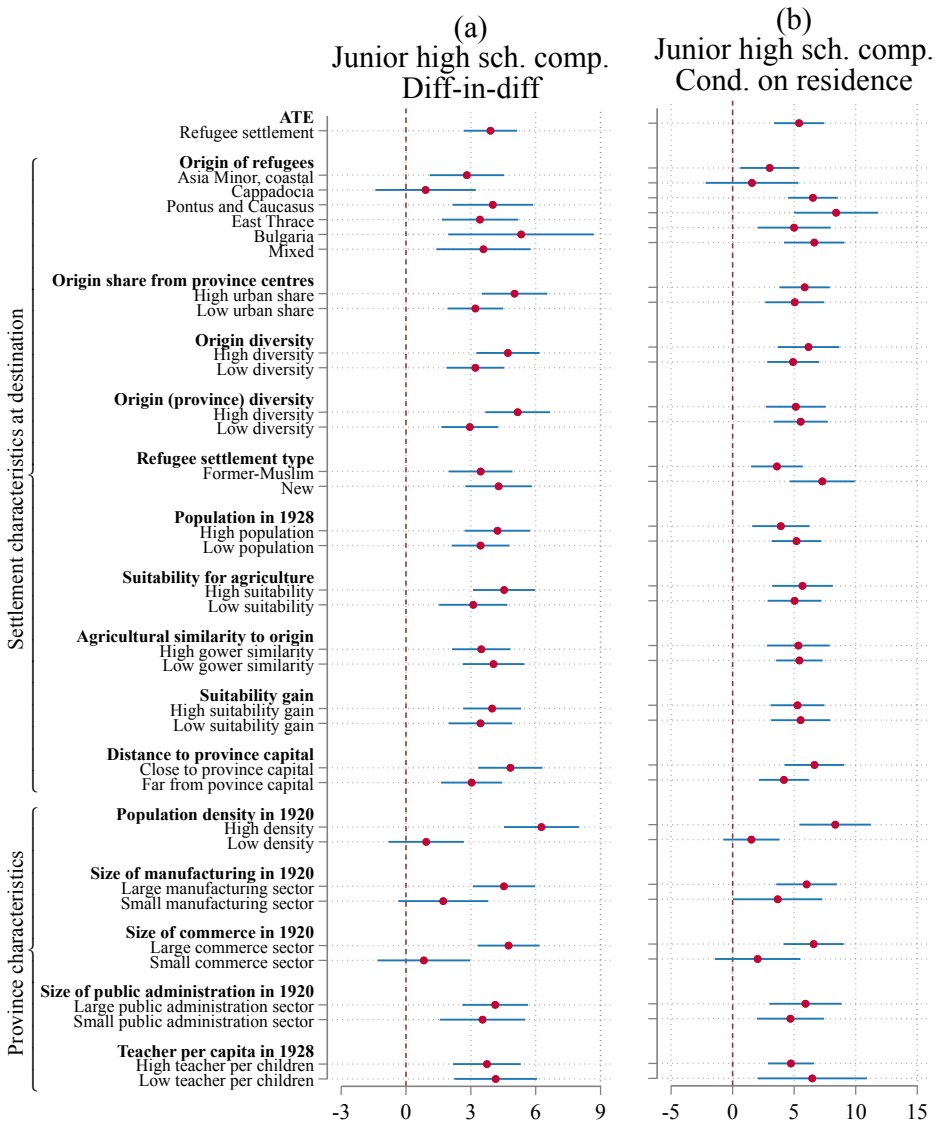
Notes. The table summarizes the main characteristics and displacement trajectories for Greek-Orthodox refugees from five broad regions in Anatolia: (i) Constantinople, (ii) Eastern Thrace, (iii) Coastal Asia Minor, (iv) Pontus, and (v) Cappadocia. The underlying data on main occupation, agricultural production, and language spoken are retrieved from the seconding of 1,500 interviews (oral histories) from the archive of *the Center for Asian Minor Studies*, *Xenophanes: Periodical Publication of the Association of Asia Minor Greeks "Anatoli"*, and Sia Anagnostopoulou (1997) monograph, "The Greek Orthodox Communities of Asia Minor."

Figure D1: Displacement and Primary School Completion by Settlement Characteristics (1971–2011)



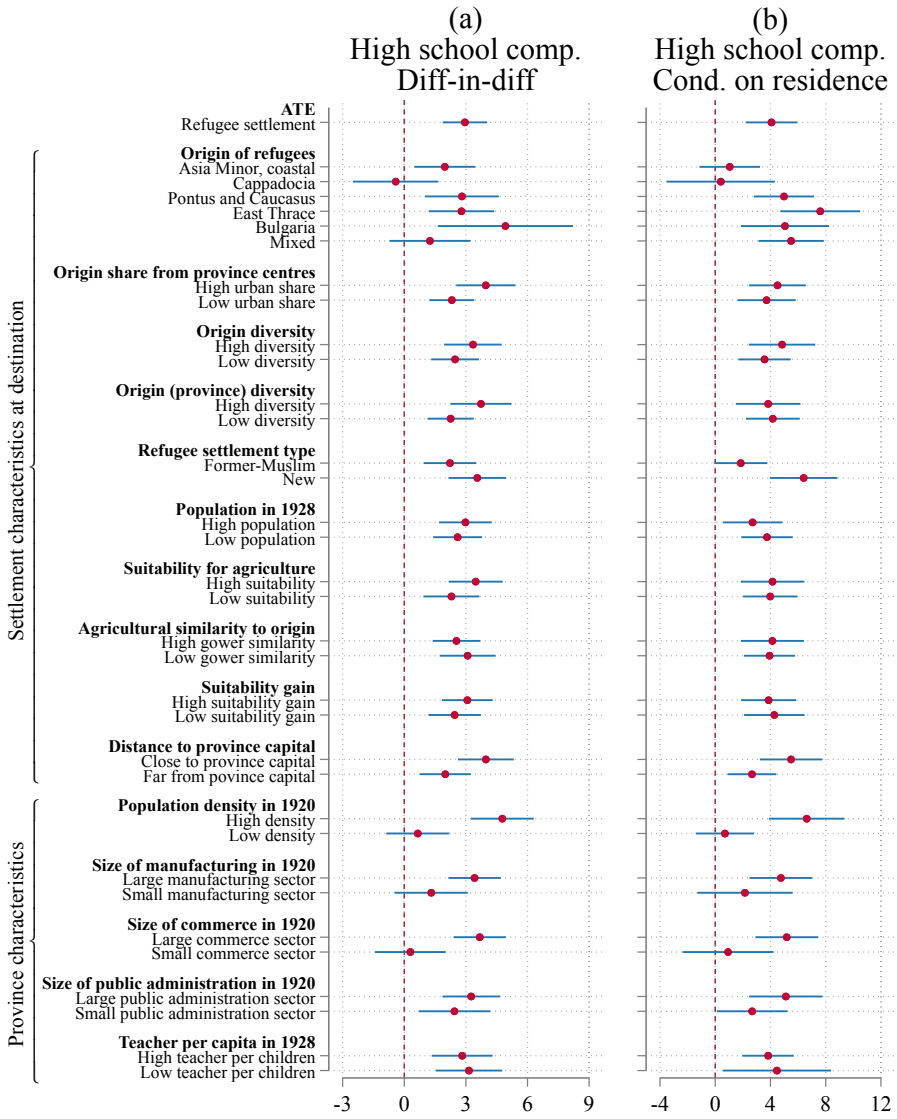
Notes. The panels present the heterogeneous treatment effects of difference-in-difference estimates (Panel *A*) and the estimates conditional on residence (Panel *B*) by settlement characteristics at destination and settlement province characteristics. For each characteristic we consider, we split the sample into two by the population-weighted median value of the characteristic at the settlement level in 1928; see Panel *B* of Appendix Table D1 to see the cutoff values. The confidence intervals in Panel *A* are based on standard errors clustered at the settlement-of-residence level, while in Panel *B* are based on standard errors double clustered at the birthplace-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure D2: Displacement and Junior High School Completion by Settlement Characteristics (1971–2011)



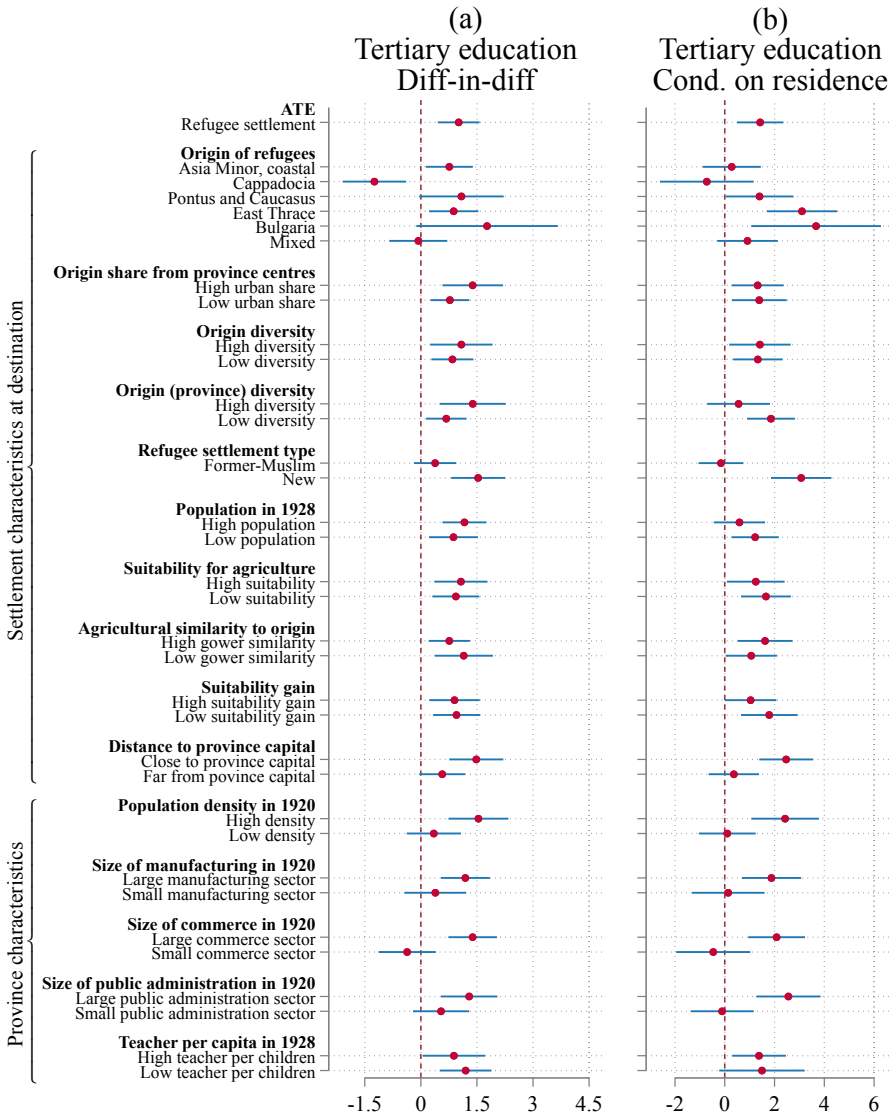
Notes. The panels present the heterogeneous treatment effects of difference-in-difference estimates (Panel *A*) and the estimates conditional on residence (Panel *B*) by settlement characteristics at destination and settlement province characteristics. For each characteristic we consider, we split the sample into two by the population-weighted median value of the characteristic at the settlement level in 1928; see Panel *B* of Appendix Table D1 to see the cutoff values. The confidence intervals in Panel *A* are based on standard errors clustered at the settlement-of-residence level, while in Panel *B* are based on standard errors double clustered at the birthplace-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure D3: Displacement and High School Completion by Settlement Characteristics (1971–2011)



Notes. The panels present the heterogeneous treatment effects of difference-in-difference estimates (Panel *A*) and the estimates conditional on residence (Panel *B*) by settlement characteristics at destination and settlement province characteristics. For each characteristic we consider, we split the sample into two by the population-weighted median value of the characteristic at the settlement level in 1928; see Panel *B* of Appendix Table D1 to see the cutoff values. The confidence intervals in Panel *A* are based on standard errors clustered at the settlement-of-residence level, while in Panel *B* are based on standard errors double clustered at the birthplace-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

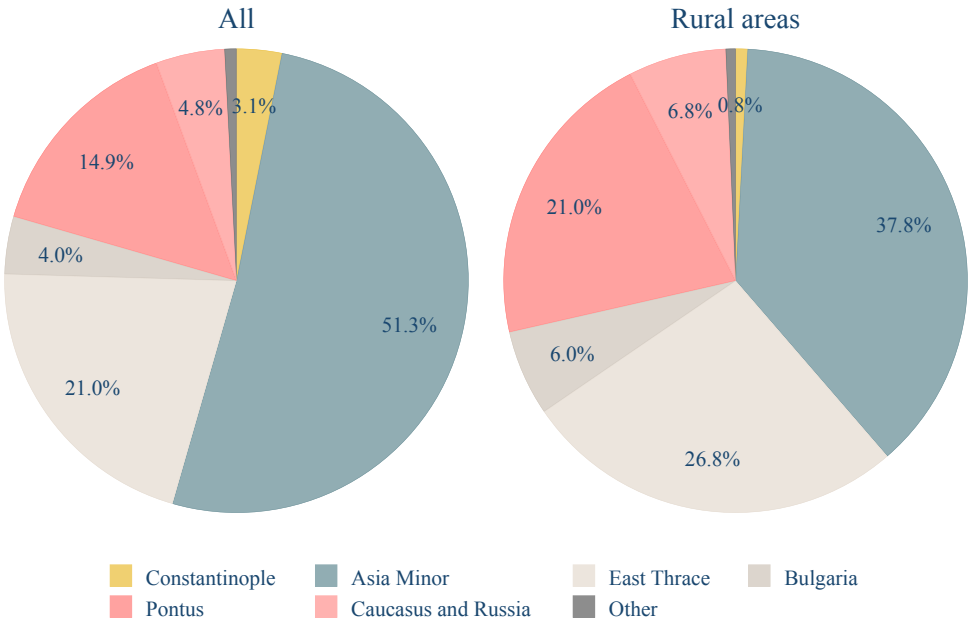
Figure D4: Displacement and Tertiary Education by Settlement Characteristics (1971–2011)



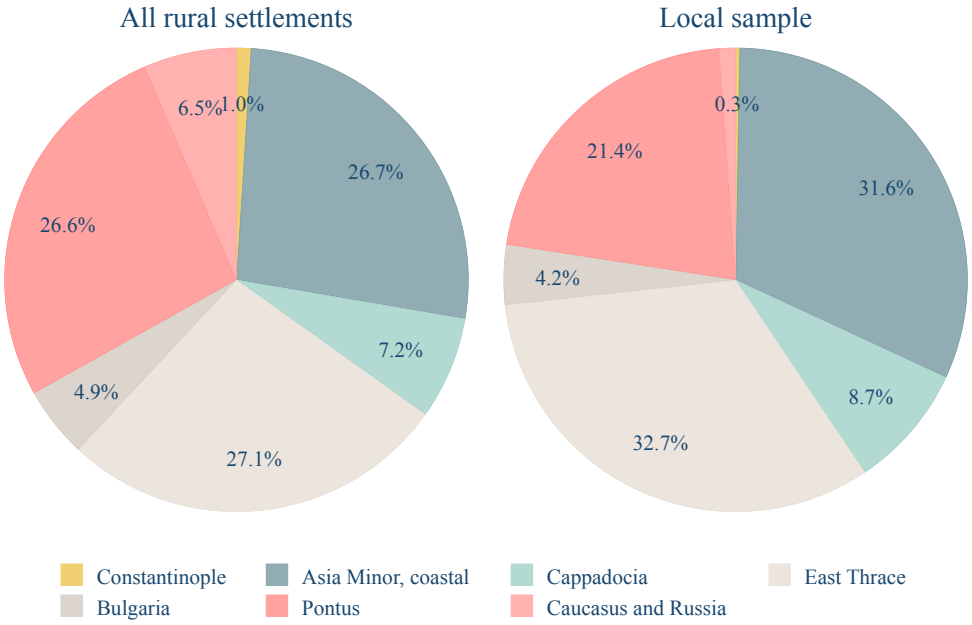
Notes. The panels present the heterogeneous treatment effects of difference-in-difference estimates (Panel *A*) and the estimates conditional on residence (Panel *B*) by settlement characteristics at destination and settlement province characteristics. For each characteristic we consider, we split the sample into two by the population-weighted median value of the characteristic at the settlement level in 1928; see Panel *B* of Appendix Table D1 to see the cutoff values. The confidence intervals in Panel *A* are based on standard errors clustered at the settlement-of-residence level, while in Panel *B* are based on standard errors double clustered at the birthplace-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure D5: Refugees' Origin

(a) 1928 General Population Census



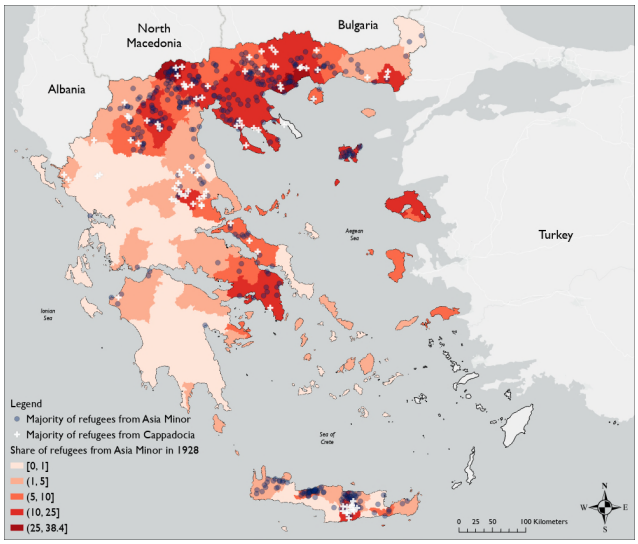
(b) Rural Catalog



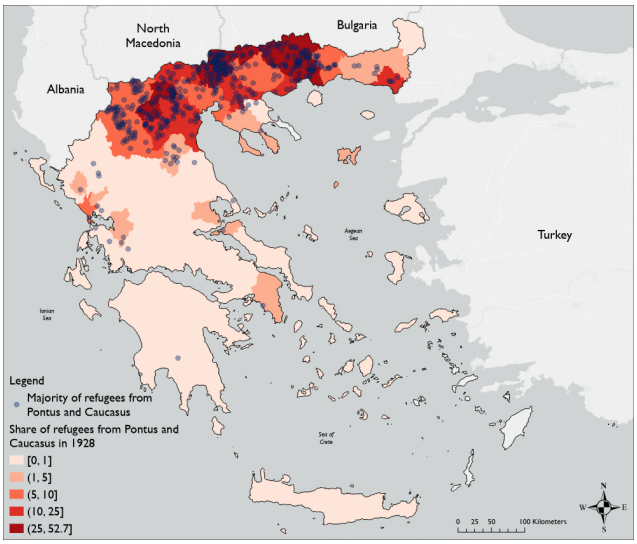
Notes. The Figures present the distribution of refugees based on region of origin. Panel *A* uses information on region of origin coming from the 1928 General Population Census and presents the share of refugees from different regions across all population (left) and rural population (right). Panel *B* uses information coming from the place of origin of refugees with land grant coming from the Rural Catalog and presents the share of refugees from different regions in the all rural settlements with a land grant (left) and among rural settlements with a land grant in the local sample (right), which focuses on a sample of rural predominantly-refugee and native settlements within a 25 km radius.

Figure D6: Refugee's Origin and Rural Settlement in Greece

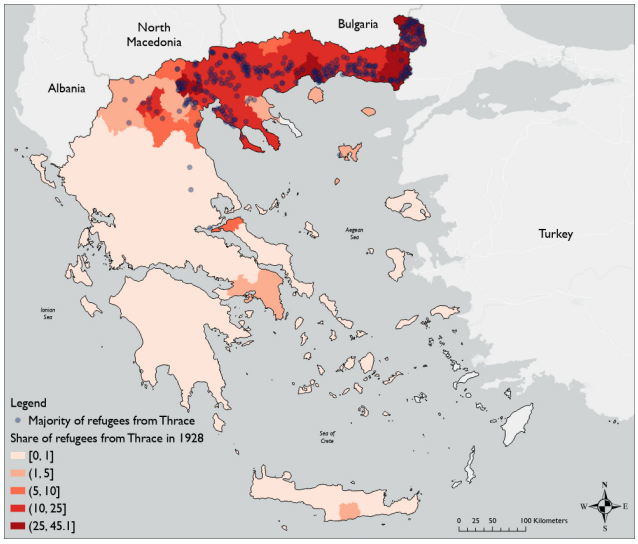
(a) Asia Minor and Cappadocia



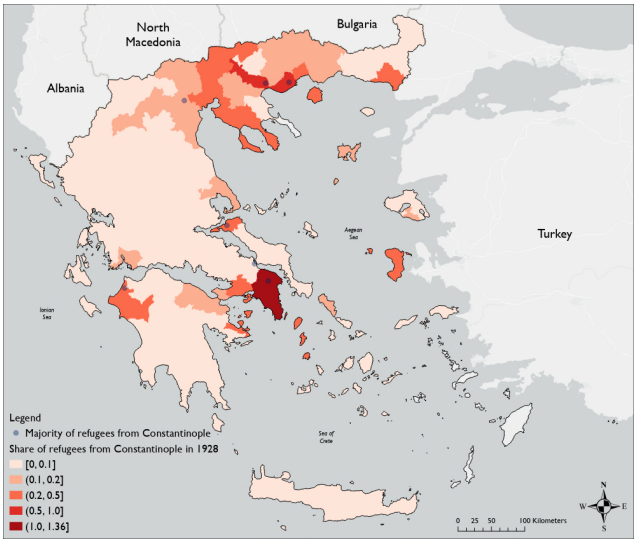
(b) Pontus and Caucasus



(c) Eastern Thrace



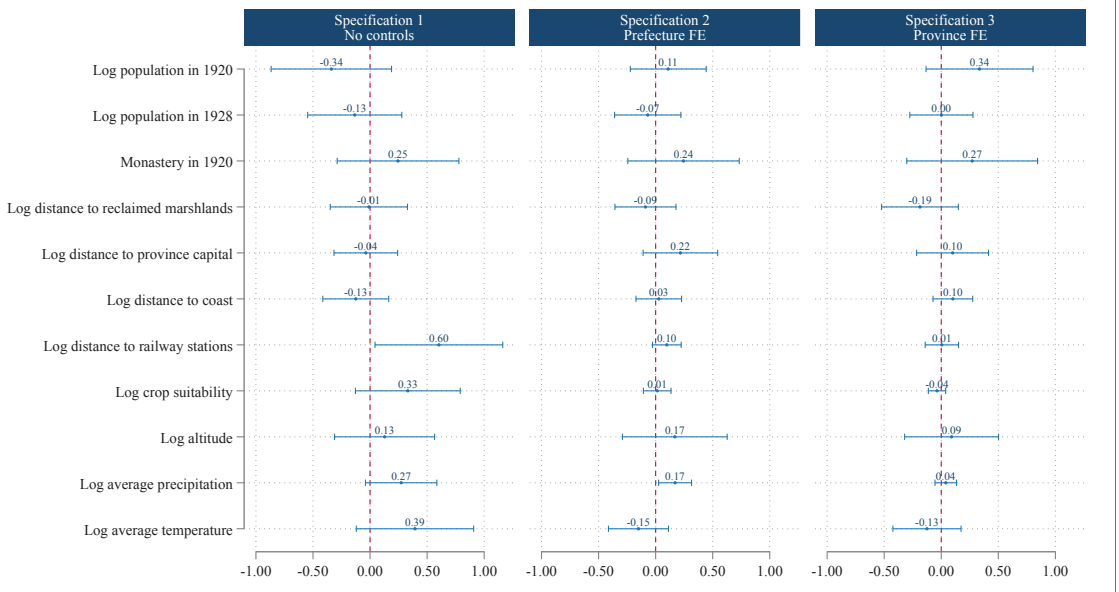
(d) Constantinople



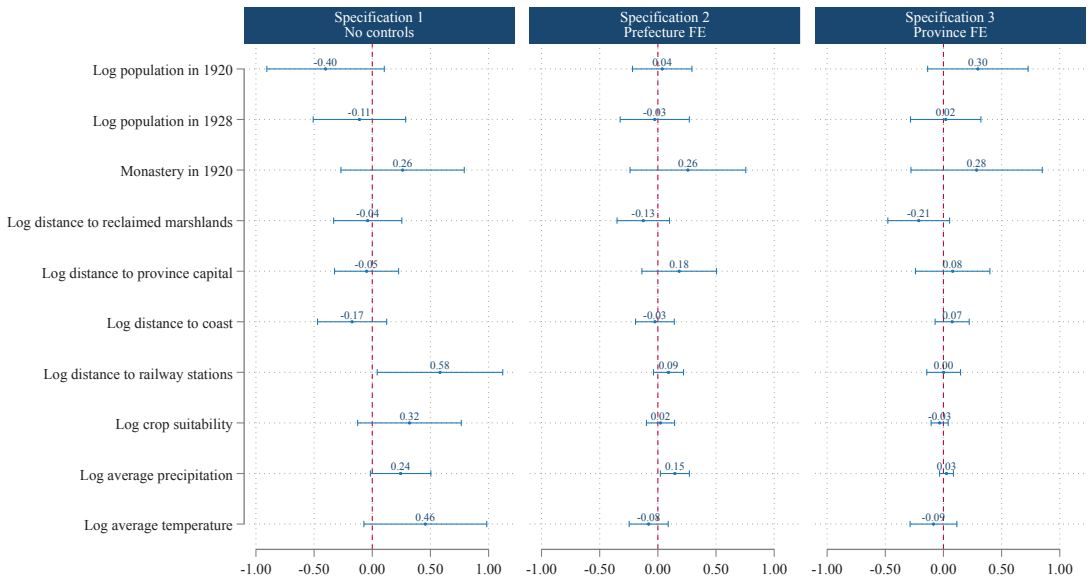
Notes. The Figure plots the share of refugees from Asia Minor and Cappadocia (Panel *A*), Pontus and Caucasus (Panel *B*), Eastern Thrace (Panel *C*), and Constantinople (Panel *D*) over the total population across 141 provinces (admin-3 level units) based on the breakdown provided by the 1928 Population Census. The 1928 Population Census tabulates refugees from Cappadocia as part of the refugees from Asia Minor. Dots (and plus signs in Panel *A*) portray settlements where refugees from each respective region in Anatolia received the majority of land grants, using data from the Rural Refugee Catalog compiled by the Refugee Settlement Committee in the mid-1920s.

Figure D7: Balance of observable characteristics across majority-Cappadocian and other refugee settlements

(a) Unconditional



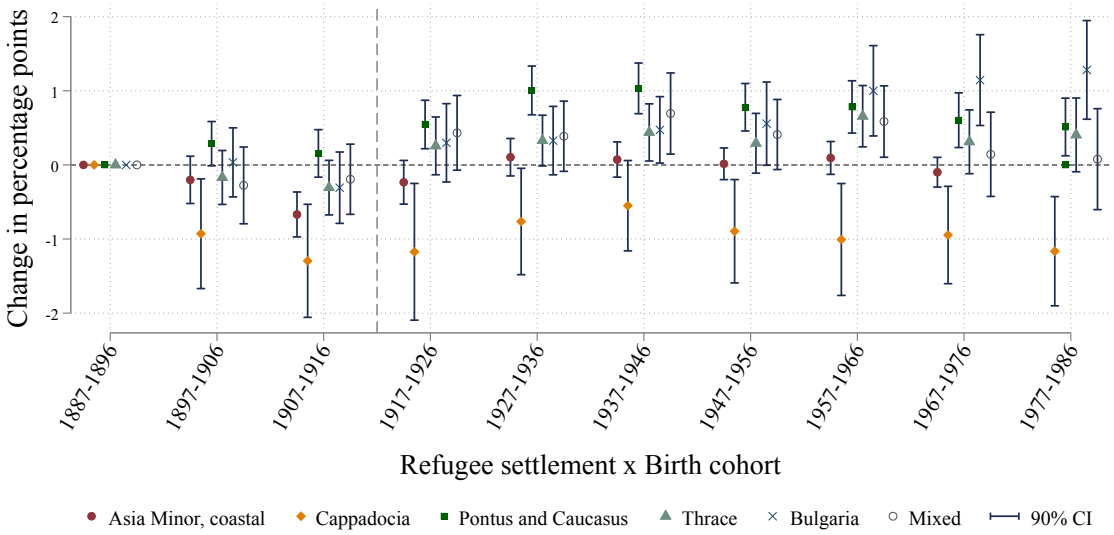
(b) Conditional on log altitude



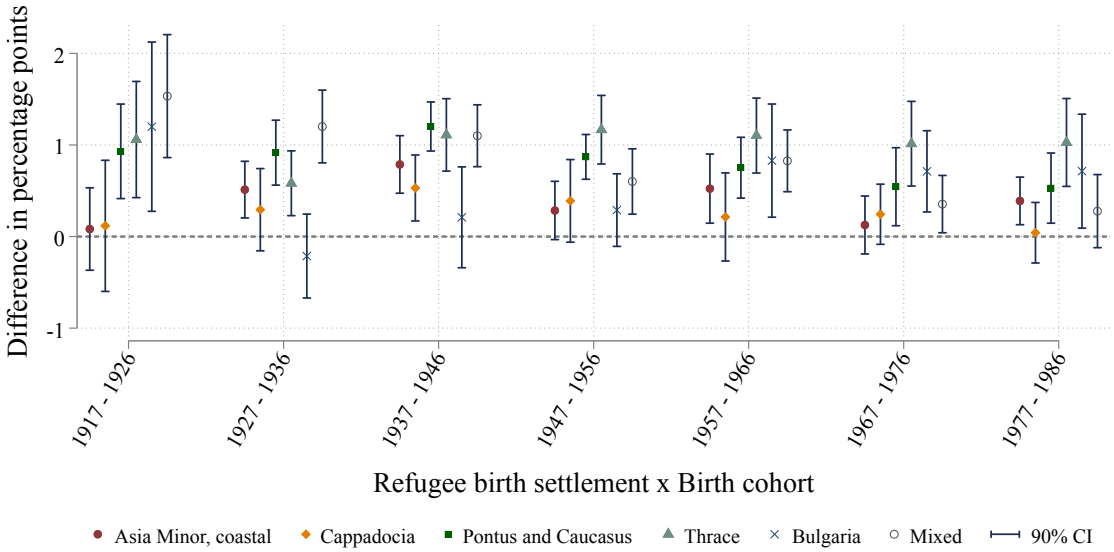
Notes. The panels explore whether various geographic, location, and population features differ between majority-Cappadocian refugee and other refugee settlements across all predominantly refugee settlements in the baseline sample. All specifications in both Panels give standardized coefficients and 90% confidence intervals obtained from regressing standardized settlement characteristics on a dummy variable indicating predominantly refugee settlements, where the share of refugees in 1928 exceeds 80%; the indicator equals zero for native settlements, where the share of refugees in 1928 is below 5%. Panel *A* presents unilateral unconditional specifications. Panel *B* gives the regression estimates conditioning on settlements log altitude. Specification 1 does not include any controls. Specification 2 conditions on prefecture fixed effects. Specification 3 conditions on province fixed-effects. Standard errors are clustered at the province level.

Figure D8: Forced Displacement and Schooling by Origin. Cohort-Level Estimates

(a) Difference-in-differences, 1971–2011



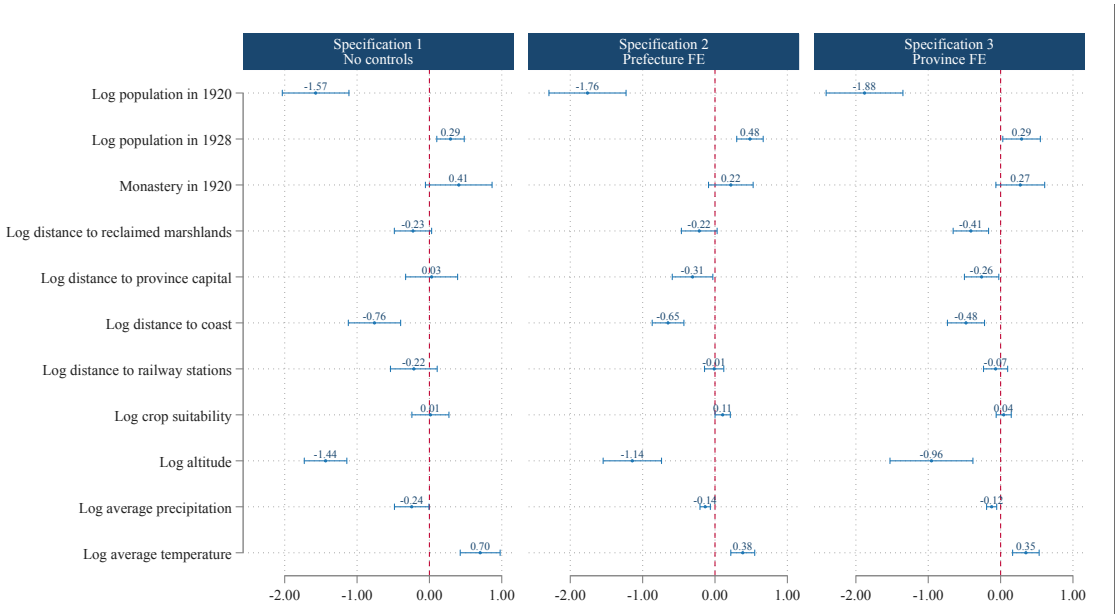
(b) Conditional on residence, 2011



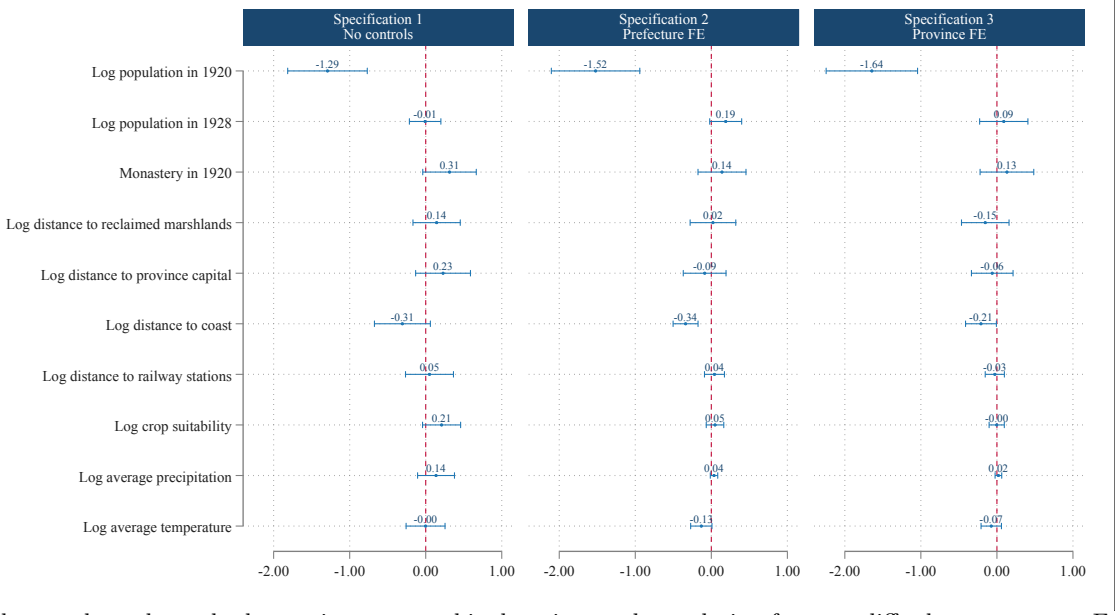
Notes. Panel *A* plots the event-study OLS estimates identifying the time-varying effect of residing in a predominantly refugee settlement interacted with the origin region of the majority of refugees with a land grant; see Figure 5 for the list of covariates. The 1887–96 birth cohort serves as the excluded cohort. Panel *B* plots the cohort-level estimates obtained from regressing years of schooling on a dummy variable indicating whether an individual was born in a predominantly refugee settlement interacted with the origin region of the majority of refugees with a land grant; see Appendix Figure B5. The confidence intervals in Panel *A* are based on standard errors clustered at the settlement-of-residence level, while in Panel *B* is based on standard errors double clustered at the birth-settlement and the municipality-of-residence levels. The difference-in-difference specifications use micro samples from the 1971, 1981, 1991, 2001, and 2011 General Population Censuses. The conditional-on-residence specifications use the micro-sample from the 2011 General Population Census.

Figure D9: Balance of observable characteristics across predominantly-Muslim and predominantly-non-Muslim refugee settlements

(a) Unconditional



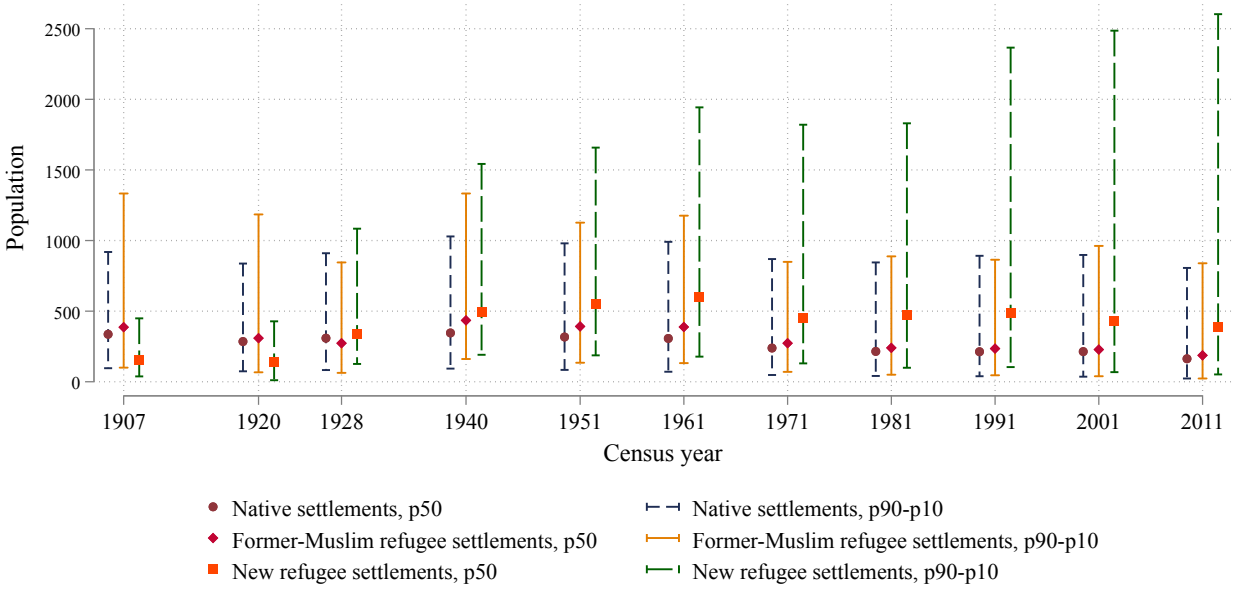
(b) Conditional on log altitude



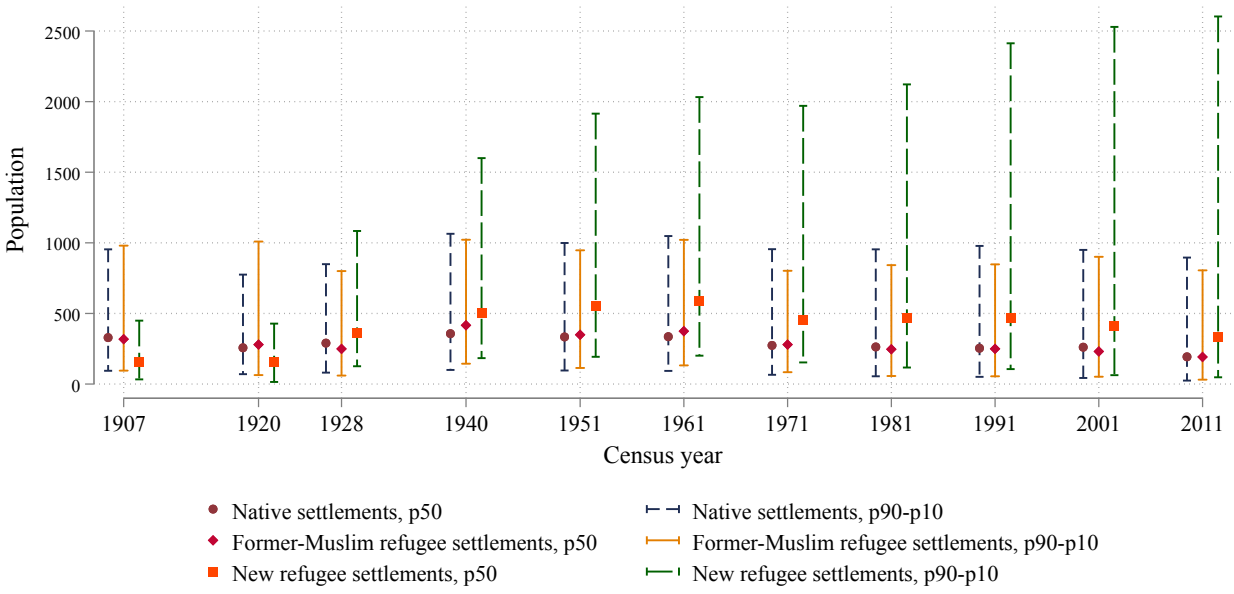
Notes. The panels explore whether various geographic, location, and population features differ between across Former-Muslim refugee settlements and new refugee settlements across all predominantly refugee settlements in the baseline sample. All specifications in both Panels give standardized coefficients and 90% confidence intervals obtained from regressing standardized settlement characteristics on a dummy variable indicating predominantly refugee settlements, where the share of refugees in 1928 exceeds 80%; the indicator equals zero for native settlements, where the share of refugees in 1928 is below 5%. Panel *A* presents unilateral unconditional specifications. Panel *B* gives the regression estimates conditioning on settlements log altitude. Specification 1 does not include any controls. Specification 2 conditions on prefecture fixed effects. Specification 3 conditions on province fixed-effects. Standard errors are clustered at the province level.

Figure D10: Population Dynamics. Rural Native and (Former-Muslim and New) Refugee Settlements

(a) All rural native and refugee settlements



(b) Local sample



Notes. The figures plot the evolution of the median and the 90th- and 10th-percentiles of the population in rural refugee and native settlements between 1907 and 2011, as recorded in the respective Censuses. Refugee villages are split into two categories: (i) Former-Muslim refugee settlements where the (imputed) share of Muslims in 1920 exceeds 80%; (ii) New refugee settlements with a corresponding share of the Muslim population in 1920 below 80%. Panel *A* plots the population statistics for all rural and native refugee settlements. Panel *B* focuses on the native and refugee settlements in the local analysis sample.

Appendix References

Anagnostopoulou, S. (1997): *Greek Orthodox Communities in Asia Minor - 19th century - 1919*. Ellenika Grammata, Athens.

Bester, C. A., T. G. Conley, and C. B. Hansen (2011): "Inference with Dependent Data Using Cluster Covariance Estimators," *Journal of econometrics*, 165(2), 137–151.

Colella, F., R. Lalive, S. O. Sakalli, and M. Thoenig (2019): "Inference with Arbitrary Clustering," *IZA DP No. 12584*.

——— (2023): "acreg: Arbitrary Correlation Regression," *The Stata Journal*, 23(1), 119–147.

Conley, T. G. (1999): "GMM estimation with cross-sectional dependence," *Journal of Econometrics*, 92(1), 1–45.

Güvenç, S. (2010): *Atlas of Old and New Toponyms of Northern Greece: Before and After the Population Exchange [Mübadele Öncesi ve Sonrası Eski ve Yeni Adları ile Kuzey Yunanistan Yer Adlari Atlası]*. Golden Medya Matbaacılık, Istanbul.

Minasidis, C. (2024): ""This Expulsion is Explained in Many Ways": Ottoman Greek Orthodox Internal Exiles during the Great War (1914–1918)," *New Perspectives on Turkey*, 70, 89–106.