Murard, Elie; Sakalli, Seyhun Orcan

Working Paper

Mass Refugee Inflow and Long-run Prosperity: Lessons from the Greek Population Resettlement

CReAM Discussion Paper Series, No. 05/20

Provided in Cooperation with:
Rockwool Foundation Berlin (RFBerlin)

Suggested Citation: Murard, Elie; Sakalli, Seyhun Orcan (2020) : Mass Refugee Inflow and Long-run Prosperity: Lessons from the Greek Population Resettlement, CReAM Discussion Paper Series, No. 05/20, Centre for Research & Analysis of Migration (CReAM), Department of Economics, University College London, London

This Version is available at:
https://hdl.handle.net/10419/295614

CReAM
Centre for Research & Analysis of Migration

Discussion Paper Series

CDP 05/20

- **Mass Refugee Inflow and Long-run Prosperity: Lessons from the Greek Population Resettlement**

- Elie Murard and Seyhun Orcan Sakalli

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

www.cream-migration.org

Mass Refugee Inflow and Long-run Prosperity:
Lessons from the Greek Population Resettlement*

Elie Murard† & Seyhun Orcan Sakalli‡

November 2019

Abstract

We investigate the long-term consequences of mass refugee inflow on economic development. After the Greco-Turkish war of 1919–1922, 1.2 million Greek Orthodox were forcibly resettled from Turkey to Greece, increasing the host population by more than 20% within a few months. To examine the long-term effects of this event, we build a novel geocoded dataset locating refugee settlements across the universe of more than four thousand Greek municipalities that existed in 1920. Using a battery of empirical strategies relying on different margins of spatial and temporal variation in the refugee inflow, we find that localities with a greater share of Greek refugees in 1923 display higher level of prosperity and industrialization sixty years after the event. These long-run benefits of refugees appear to be driven by the provision of new agricultural know-how and the transfer of technological knowledge in textile, which fostered growth through higher diversity in complementary skills. The economic gains of the resettlement were lower in places where refugees were clustered in separate enclaves and where their skills were less easily transferable due to local geographic conditions.

JEL Classification Codes: O10, O43, N34, N44

Keywords: Refugees, Immigration, historical persistence, economic development

*We thank Catia Batista, Samuel Bazzi, Simone Bertoli, Pierre-Philippe Combes, Mathieu Couttenier, Christian Dustmann, James Fenske, Stelios Michalopoulos, Çağlar Özden, Thomas Piketty, Hillel Rapoport, Claudia Senik, Steven Stillman, Mathias Thoenig, Sara Tonini, Ekaterina Zhuravskaya, and seminar participants at Aix-Marseille School of Economics, CERDI Clermont-Ferrand, CReAM UCL, GATE ENS Lyon, IZA, NOVA, Paris School of Economics, Stellenbosch University, University of Alicante, University of Humboldt, University of Warwick, and at the 11th Migration and Development Conference at Stanford University, Barcelona GSE Summer Forum, and Dial Development Conference for helpful suggestions. Elie Murard acknowledges financial support from IZA and Seyhun Orcan Sakalli acknowledges financial support from the ERC Starting grant GRIEVANCES-313327. We thank Zacharen Batziou for excellent research assistance.

†University of Alicante, IZA Institute of Labor Economics, LEAP, Murard@iza.org

‡King's Business School, King's College London, seyhun.sakalli@kcl.ac.uk

"The aspect of the country has entirely changed; Everywhere one sees the cheerful red roofs of the colonization settlements. Where formerly vast uncultivated plains stretched, there are now flourishing villages, full of bustling activity, and showing obvious signs of comfort, and in many cases prosperity. The results are due to the courage, energy, the capacity for work, the receptivity to new ideas, which characterize the mass of refugees."

Sir John Campbell, first Vice-Chairman of the League of Nations' Refugee Commission[1]

1 INTRODUCTION

Mass displacement of people is abound throughout history. The 20th century alone witnessed the uprooting of millions of Jews, ethnic Germans, Hindus and countless others. More recently, civil wars and ethno-religious conflicts displaced millions of Sudanese, Syrians, and Rohingyans, and many more refugees are to be expected as environmental change continues. While significant attention has been given to the costs that mass population inflows may cause for the hosting countries in the short-term, e.g., in terms of food scarcity or congestion in public services, much less is known about the long-term legacy of such events. In particular, little evidence exists on how large-scale population resettlements affect receiving communities in the long-run (Becker and Ferrara, 2019). This question is important because the short-term and long-term effects could be very different.

We aim to fill this gap by exploiting the 1923 Greek population resettlement to examine the long-run consequences of a mass refugee inflow on economic development and prosperity. In the aftermath of the Greco-Turkish war of 1919–1922, fearing persecutions, the Greek Orthodox population of the Ottoman Empire fled from their homeland in Turkey where they inhabited since antiquity.[2] In total 1.2 million Greek Orthodox were forcibly resettled from Turkey to Greece, increasing the host population by more than 20% within a few months. Although they did not necessarily identify themselves as Greeks, the refugees shared the same religion and often spoke the same language as the host native population. The inflow of refugees in Greece was shortly followed by the forced expulsion of all Muslim inhabitants of Greece (about 400,000 people), who resettled in Turkey by 1928. Yet we find no evidence that the displacement of Muslims played a significant role in the long-run economic development of Greece. The relevant consequence of this exchange of population was that it resulted in a mass population inflow, which we focus on in this paper.

We examine the effects of a mass refugee inflow in a setting in which substantial investments were made to resettle the newcomers and integrate them into the host country. Due to the limited resources of Greece, an agrarian country with low state capacity in the 1920s, the League of Nations (the precursor of the UN) intervened to provide immediate relief to the sudden mass inflow of refugees. The Refugee Settlement Commission (RSC) was formed to allocate the available funds and establish refugees in productive agricultural work. Arable land parcels, farm inputs, and cattle

[1] League of Nations, Official Journal, 11th Year, NO.6 (June 1930), Annex 1211, "Twenty-Fourth Quaterly Report of the Commission", No. C 559 M.210. 1929 II[F.727], p 712.

[2] In the Ottoman Empire, Orthodox Christians were grouped in the same confessional community, the Rum millet, under the spiritual jurisdiction of the Ecumenical Patriarch in Constantinople, a.k.a the Greek Orthodox Church. The Rum millet were mainly composed of ethnically Greek population whose customs and traditions dated back to the Byzantine civilisation and the ancient Hellenic world.

were provided, in addition to individual houses, often in the same villages as natives. Given that refugees could not return to their homelands and that their resettlement in Greece was permanent, they were also granted the Greek citizenship to foster their assimilation into the host society.

We built a novel geocoded dataset of refugee settlements, obtaining a full coverage of the universe of more than 4,000 municipalities that existed in Greece in 1920. We combined historical data from the 1920, 1923, and 1928 censuses with population censuses in 1971 and 1991. Exploiting the high granularity of our data, we study comparative economic development across municipalities within the same Greek province – an area of typically smaller than 900 square km. Consistent with the historical narrative of a great economic success, a "blessing in disguise" as the initial quote suggests, we find that places of refugees' resettlement in 1928 display higher levels of prosperity in 1991 (and 1971) relative to other localities without refugees. Resettlement localities have higher night light luminosity, dwellings with better amenities (e.g. electricity), a larger manufacturing sector, and higher average earnings, conditional on their geographical and pre-resettlement economic characteristics. However, the inflow of refugees is not always associated with higher benefits. We uncover a hump-shaped relationship between the share of refugees and long-run prosperity: The resettlement produced smaller economic gains in places with more than 50 percent of refugees in the local population.

We address endogeneity concerns due to self-selection of refugees into localities with better economics prospects in four different ways. First, we exploit the distribution of refugees reported in the census of April 1923, instead of using that in 1928. At that time, refugees had been in Greece for less than 6 month. Upon arrival, their principal worries were arguably to find relief and temporary shelter where they could and less so to seek places with better economic opportunities for permanent settlement.[3] Second, we restrict the analysis to clusters of spatially contiguous municipalities, where a municipality hosted a sizeable refugee population and its neighbors did not. Then, we use a propensity-score matching estimator to obtain a fully balanced sample in terms of geographical and pre-resettlement characteristics.[4] Third, we use province-level panel data and estimate a fixed-effects model to control for time-invariant unobservables at the province level, in addition to accounting for time-varying unobservables at a greater administrative level (region or department). Fourth, we focus on the effects of refugees in Northern Greece, where three fourth of all refugees were resettled. The RSC assigned refugees to agricultural colonies in Northern Greece, mainly on the basis of available vacant land suitable for cultivation. This process prevented agricultural refugees from choosing their settlement location.

The estimates obtained from these different empirical strategies all point to beneficial effects of refugees on long-run economic development. The magnitude of the estimates suggests that, for example, moving a municipality's share of refugees in 1928 from zero to 10 percent results in a 15 percent larger manufacturing sector in 1991. Using province-level panel data we find that the effect

[3] In April 1923, the refugees still hoped that they could return to their homeland in Turkey. It is only in July 1923 that Greece and Turkey agreed on a definitive exchange of population, which made their settlement in Greece permanent (see section 2).

[4] We also compute Oster's delta that quantifies how important selection of refugees in 1928 on unobservables should be compared to that on observables to explain the estimated effects (Oster, 2019).

of refugees on the expansion of the non-primary sector was felt as early as 1928. The magnitude of this effect more than doubles after the 1960s, when the overall process of structural transformation gained pace in Greece.

Importantly, we examine whether the economic gains enjoyed by localities that received a greater number of refugees came at the expense of other nearby localities with fewer or no refugees. We find no evidence of such negative spillover effects in nearby localities. Instead, we document evidence of positive spillover effects. This suggests that our findings are *not* driven by reallocation of economic activity into places of resettlement, but rather by creation of new activity in them.

We also document significant heterogeneous effects of refugees across regions. 80% of the refugees who were resettled in Northern Greece were farmers, generally less educated than natives. In contrast, those who settled in Central Greece were mostly skilled non-agricultural workers, with more education than natives on average. We find that refugees had significant positive effects within both regions. In line with the large literature on the benefits of high-skilled immigration, we find higher economic gains of refugees within Central Greece than within Northern Greece.[5] Interestingly, refugees also had beneficial (although smaller) effects within Northern Greece, despite having lower levels of human capital relative to natives.

We then examine the mechanisms explaining this nontrivial result within Northern Greece. The economic gains of the resettlement in Northern Greece appear to arise from the new complementary skills brought by the refugees. We find evidence for two channels through which skill complementarities fostered long-term growth: (i) the transfer of new technologies, in particular in the textile sector, which was conducive to industrialization; (ii) the provision of new agricultural know-how that increased the diversity of skills when combined with the ones of natives. Consistent with (i), we find that the resettlement of refugees led to the expansion of the textile industry and the manufacturing sector, especially in localities with a greater suitability for cotton cultivation. In these places refugees' knowledge in textile could be more effectively transferred and employed. Consistent with (ii), we find that the resettlement produced smaller economic gains in localities where interactions between refugees and natives were less likely. These were localities with more than 60% of refugees in the population (and thus with a less diverse set of skills), or localities where natives and refugees were segregated into separate villages within the same municipality. These interactions – and the ensuing knowledge sharing – were certainly facilitated by the fact that refugees and natives often spoke the same (or similar) language. We also find evidence of agglomeration economies due the increase in population density. Yet, agglomeration effects can only account for a small fraction of the refugees' impacts.

Other contextual factors could have played a role in shaping the beneficial effects of refugees. In particular, the inflow of refugees was shortly followed by the forced expulsion of all Muslim inhabitants of Greece. The RSC redistributed the large tracts of land left vacant by the Muslims to the refugees. Also, the RSC undertook drainage and reclamation works to expand the cultivable

[5] See Kerr et al. (2016) for an overview of the literature on the productivity gains generated by present-day high-skilled immigration. For the contributions of high-skilled migrants in historical settings, see Hornung (2014), Rocha, Ferraz and Soares (2017), Droller (2018), Moser, Voena and Waldinger (2014) among others.

land. These facilitated the establishment of refugees in agricultural work and reduced potential conflicts with natives due to land scarcity. However, we find no evidence that the positive impacts of refugees are driven by localities formerly inhabited by the Muslim population (where more land was available), or by places that received greater levels of RSC's investment in public infrastructure in terms of reclamation works or health facilities. Nevertheless, the allocation of land titles to the refugees was a part of a broader nation-building policy aimed at the assimilation of the newcomers into the Greek society, which began with their naturalization. Such integration policies, in addition to the relative cultural similarity between refugees and natives, are likely to have mitigated the costs generally associated with population diversity, notably in terms of coordination, trust or public goods (Alesina and Ferrara, 2005).[6]

Our paper primarily contributes to the literature examining the consequences of forced migration on receiving communities, reviewed in Becker and Ferrara (2019) and Ruiz and Vargas-Silva (2013). As recent work has highlighted, forced displacement has effects distinct from voluntary migration, notably due to the unexpected, massive, and often permanent nature of the population inflow. Refugee inflows can cause significant short-term costs for receiving communities, in particular in developing countries, either in terms of disease outbreaks (Baez, 2011; Montalvo and Reynal-Querol, 2007), surge in food prices (Alix-Garcia and Saah, 2010), or increased wage competition (Calderón-Mejía and Ibáñez, 2015; Morales, 2018; Ruiz and Vargas-Silva, 2015). In contrast, using historical contexts, several studies find evidence of beneficial effects of refugees, especially in terms of agglomeration economies. Focusing on the influx of millions of ethnic Germans to West-Germany after World War II, Braun and Kvasnicka (2014) and Peters (2017) find that expellees fostered structural transformation away from agriculture and led to higher rates of employment in manufacturing and higher incomes per capita. Schumann (2014) shows that this population shock had persistent effect on population size. Sarvimäki (2011) finds similar agglomeration effects in the case of the forced resettlement of Finnish population after WII. While these historical studies investigate medium-run effects, typically over a 20-year period, we uncover the long-run legacy of a large-scale population resettlement, more than 60 years after the event. Similarly, Bharadwaj and Mirza (2019) explore the effects of the 1947 population exchange between India and Pakistan and find higher increase in agricultural productivity in refugee-receiving regions after the Green Revolution until 2009. Becker et al. (2019) study the long-run effects of the resettlement of Poles after WWII and find that descendants of forced migrants invested more in education as a result of a shift in their preferences towards mobile assets, such as human capital. In our context, we find no evidence that the second-generation of Greek refugees have higher levels of education relative to natives.

Our findings also complement the studies on the long-run consequences of European immigration to the U.S. during the Age of Mass Migration (Abramitzky, Boustan and Eriksson, 2014;

[6]It is yet often the case that refugees are hosted in neighboring countries that are culturally similar. For example, Bangladesh hosts current Muslim Rohingya refugees who fled religious persecution in Myanmar; most of the Syrian refugees are hosted in neighboring Muslim countries some of which are Arabic speaking; and Colombia hosts Venezuelan refugees who migrated after the economic collapse of the country.

Sequeira, Nunn and Qian, 2019) and of high-skilled European settlers in South America in the mid-nineteenth century (Rocha, Ferraz and Soares, 2017; Droller, 2018), which identify positive effects on host communities in the long-run. Differently from the contexts of these studies, Greek refugees were not more skilled than the host population on average.[7] Our results suggest that refugees contributed to long-term growth trough skill complementarities, in line with the literature documenting the economic benefits of diversity (Ager and Brückner, 2013; Peri, 2012; Ortega and Peri, 2014; Alesina, Harnoss and Rapoport, 2016; Docquier et al., 2019). Our findings are also consistent with the studies highlighting the role of technology diffusion and knowledge transfers caused by population movements (Bahar and Rapoport, 2018). In particular, Hornung (2014) analyzes the effects of the technological transfers brought by the resettlement of French Hugenots in seventeenth-century Prussia and finds evidence of persistent productivity gains in textile manufactories in the long-run.

Relative to the existing literature, we make several contributions. First, we analyze the hitherto unstudied mass population resettlement in Greece in 1923. This unique setting allows us to examine the long-run effects of a permanent refugee resettlement in a poor agrarian country, under the assistance of an international organisation – the League of Nations. This case study is of high policy relevance as most of the currently displaced people are hosted in developing countries, often in rural areas.[8] Second, the heterogeneous effects of refugees that we uncover provide several insights for the design of resettlement policies aimed at maximizing the contributions of refugees to the host economy.[9] We identify greater economic benefits of the resettlement in localities where native and refugee populations were more mixed and less segregated, as well as in localities where refugees' skills were more easily transferable. These findings highlight the crucial role of social interactions, knowledge sharing, and skill transferability. Finally, our results also suggest that, by providing complementary knowledge, refugees can generate significant benefits in the long-run, even when they do not bring higher human capital.

The remainder of the paper is organized as follows. Section 2 provides a description of the historical background. Sections 3 and 4 describe our data and empirical strategy. Section 5 presents the main results and robustness checks. Section 6 explores the mechanisms, and section 7 concludes.

[7] Other research examines the role played by high-skilled ethnic or religious groups in economic development, such as the contributions of Jews to economic growth in Europe (Pascali, 2016) and in the U.S. (Moser, Voena and Waldinger, 2014), the economic legacy of Protestantism in Germany (Becker and Woessmann, 2009), or the role of Armenians and Greeks in the Ottoman Empire (Arbatli and Gokmen, 2016).

[8] 85% of the 70 million currently displaced people in the world are hosted in developing countries (UNHCR, 2018). Sub-Saharan Africa currently hosts the largest number of forcibly displaced people (26 millions), which are mostly resettled in rural areas (Huang and Graham, 2018).

[9] See Clemens et al. (2018) for a thorough discussion of the policies aiming at facilitating the integration of current refugees in the host country labor market and enhancing their potential contribution to the economy.

2 Historical Background

2.1 The Population Exchange

The Kingdom of Greece gained its independence from the Ottoman Empire in 1832. Since then, it pursued the goal to unite ethnic Greek population scattered across Ottoman territories. This irredentist aspiration is known as the "Megali Idea". Greece annexed Thessaly in 1881, Epirus, Macedonia, Crete, and Ionian islands in 1913. After the First World War, backed by the Allies, the Greek Army occupied territories in western Turkey that were heavily populated by ethnic Greeks since antiquity, such as western Asia Minor, historically known as Ionia, and Thrace (see Figure A.2 in Supplementary Appendix).[10] However, the Greek advance stalled as the military and diplomatic support of the Allies faded. The Turkish Army's counter-offensive at the end of August 1922 was succesful. The persecution of Greek Orthodox communities by the advancing Turkish Army, in particular the fire in the Christian neighborhoods of Smyrna in September 13, 1922, triggered a mass exodus of population to Greek islands and mainland. Almost one million of Greek Orthodox, whose ancestors have been living in Asia Minor for millennia, left Turkey within a few months time.

The Lausanne Peacy Treaty, signed in July 1923, ended the Greco-Turkish War of 1919–1922. As a part of the treaty, the Kingdom of Greece and Turkey signed the "Convention for the Exchange of Populations", which devised and implemented the principle of a compulsory exchange of population for the first time in modern history. The convention dictated that all Greek Orthodox of Turkey and all Muslims of Greece were to be denaturalized, leave the country they reside in, and acquire the citizenship of their destination country. It affected all remaining Greek Orthodox in Turkey and all Muslims living in Greece.[11] It retroactively affected the refugees who had already fled to Greece.

As a result of the population exchange, Greece, a country with less than five million inhabitants, received more than one million refugees from Turkey. In return, about 400,000 Muslims left Greece.[12] Many of the refugees did not identify themselves as Greeks. Their expulsion from Turkey was only based on their adherence to the Greek Orthodox Christian church. In addition to sharing common religious bonds, Greek natives and refugees also often spoke the same or similar language.[13]

[10] Cappadocia and Pontus, the other two regions that were populated by Greeks, were under the control of the Turkish forces during the Greco-Turkish War. Table A.1 in Supplementary Appendix presents the timeline of events.

[11] Only Greek Orthodox residing in Constantinople and Muslims residing in Western Thrace, which was a part of Bulgaria before WWI, were exempted. With the Treaty of Lausanne, Greece and Turkey also decided on exchanging the wealth of the exchanged populations. Therefore, Greece seized control of the properties abandoned by the Turks who left.

[12] A population exchange of a much smaller scale took place between Greece and Bulgaria in 1920. After the Balkan Wars (1912–1913) that reshuffled the Greek and Bulgarian territories in Macedonia, the two countries agreed on a reciprocal and voluntary exchange of their respective ethnoreligious minorities at the 1919 Convention in Neuilly-sur-Seine. According to the estimates provided by the commission supervising the migration, about 90,000 Bulgarians left Greece and 50,000 Greeks left Bulgaria (R., 1944).

[13] See section A in Supplementary Appendix for a short description of cultural similarities between refugees and natives.

2.2 Resettlement of Refugees in Greece

The relief and resettlement of such a large scale refugee inflow were far beyond the limited resources of the Greek State, which had almost gone bankrupt after a long and costly war. As a result, the League of Nations intervened and founded the Refugee Settlement Commission (RSC) in September 1923. The RSC had full authority over the distribution of funds and resettlement of refugees.

The Mandate of the RSC. The mandate of the RSC was to establish refugees in agricultural work and assist them to become self-supporting. To this end, the RSC provided them with land suitable for cultivation and with various farm inputs, such as livestock, seeds, and some machinery (Kontogiorgi, 2006, pp. 78, 86, 90). The houses occupied by refugees were often the ones left behind by the expelled Muslims. By contrast, the refugees who settled in urban centers – about 50% of them – were given no means of providing for themselves, apart from accommodation (Kontogiorgi, 2006, p. 103). The urban settlement was not or little supervised by the RSC, and almost no funds were spent on industrial or urban projects (Pentzopoulos, 1962, p. 114).

Agricultural Resettlement. The RSC concentrated its operation in Northern Greece where large tracts of arable land were left vacant by the expelled Muslim population. The RSC split the large estates of Muslim landlords (*chiftliks*) into smallholdings and distributed them to the Greek refugees.[14] The RSC also expanded the cultivable land area by undertaking drainage and reclamation works in the Northern provinces, mainly in Macedonia. As a result, the area of cultivated land doubled in Northern Greece (Kontogiorgi, 2006, pp. 277–290). In total, the Greek regions of Macedonia, Thrace, Epirus, and Thessaly hosted more than 90% of the refugees' agricultural settlements.[15]

Resettlement Process. The first stage of the agricultural settlement consisted in assembling refugees into legally constituted groups based on their community of origin. The RSC assigned a place of settlement to each group through its designated representatives. Then, refugees were transported to the settlement location, where a government representative delivered the land (Pentzopoulos, 1962, p. 108). The RSC planners determined the settlement locations by assessing the availability of suitable land for cultivation, which primarily depended on geographical features.[16] Cases of individual settlement were very rare and 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).

[14] The Greek State also expropriated some of the native population to free additional land for the refugees. This included people who had purchased or occupied the land left over by expatriated Muslims. Out of the 850,000 hectares that the Greek State transferred to the RSC by 1929, only less than 10% were acquired by expropriation of Greek citizens (Kritikos, 2005).

[15] Authors's calculation using data from the 1930 Statistical Yearbook of Greece.

[16] As C. B. Eddy, the last president of the RSC remarked: *"This distribution was determined by several quite patent factors, of which the geographical was the controlling one. The capacity of the country in the absorption of refugees was limited by its physical features*" (Kontogiorgi, 2006, p. 255).

2.3 The Economic Role of Refugees

According to historians, notably to D. Pentzopoulos, the inflow of refugees "proved to be a motivating force of tremendous effectiveness in the reorganization and invigoration of the economy" and transformed Greece from a "backward and parochial country to a dynamic and economicallly fully developed state" (Pentzopoulos, 1962, pp. 115, 254). We list the potential economic contributions of refugees below.

Provision of Agricultural Know-how. About 65% of the refugees were farmers, who often introduced new varieties of crops and agricultural techniques in Greece. Refugees were notably experienced cultivators of tobacco. They adapted to Greek soil and climate so successfully that tobacco became one of the most important source of export revenue in the 1920s (Kritikos, 2005; Pentzopoulos, 1962).[17] They also brought new knowledge in vine-growing (especially in the sultana variety), in silkworm breeding, and in the cultivation of new varieties of cotton, millet, hemp, and other crops. Moreover, they are credited for having superior know-how in farming practices, such as the rotation of crops, the polyculture (instead of monoculture), and the use of forage plants, previously unknown in Greece (Pentzopoulos, 1962, pp. 151, 155).

Provision of Industrial Skills and Know-how. The boom of the Greek textile industry in the 1920s is largely credited to the entrepreneur refugees. They provided expertise in the manufacture of woollen materials and silk fabrics and developed the carpet-marking industry in Greece (Kritikos, 2005; Pentzopoulos, 1962).[18] More generally, refugees brought their traditional know-how in various arts and crafts, such as the manufacture of copper articles for Pontic refugees, or pottery, silver objects, and leather for refugees from Asia Minor.

Provision of Commercial Skills. Some refugees originated from Constantinople and Smyrna – two of the most important commercial centers in the East Mediterranean at that time. A large proportion of them were merchants with trade connections in Western Europe and vast experience in business and international finance (Pentzopoulos, 1962, pp. 115, 165).[19]

Increase in Labor Force and Market Demand. The refugees provided a very large and cheap labour force to Greek employers in many urban centres (Kritikos, 2000). Also, the dramatic population increase due the refugee inflow generated a corresponding increase in market demand in manufactured goods, potentially leading to agglomeration effects.[20]

[17] According to Henry Morgenthau, the chairman of the RSC, refugees produced two-thirds of the total tobacco production in Greece in 1926. Tobacco production doubled in Greece between 1922 and 1927 and an important share was exported to the United States and used in popular brands such as Lucky Strikes, Camels, and Chesterfields (Kritikos, 2005).

[18] "Various types of carpets were manufactured [...] competing successfully in the international market and bringing into Greece 400,000 pounds annually. It was really the birth of a new industry" (Pentzopoulos, 1962, p. 163).

[19] According to the survey conducted by the League of Nations in 1926: "Of the 7,000 merchants and industrialists enrolled in the Athens Chamber of Commerce, 1,000 were refugees, while the proportion was even higher in Piraeus. In 1961, 20 percent of Greek industrialists had been born in Asia Minor and Eastern Thrace" (Hirschon, 2003).

[20] "From the industrial view-point, the very fact of the sudden increase of the population to the extent of 25 per cent has created a demand for every kind of manufactured goods which can only be met by increased industrial output" (Sir Eric Hambros, Financial News, 23 Apr. 1929, cited by Kontogiorgi (2006, p. 284)).

3 Data and Location of Settlements

3.1 Data

In this paper, we examine whether the localities that hosted a greater inflow of refugees have different levels of development in the long-run. To this end, we assemble a novel dataset that combines historical data on refugee settlement with contemporary socio-economic outcomes at a very disaggregated administrative level.

3.1.1 Historical Data

Our primary data source to identify the distribution of refugees is the published results of the 1928 population census.[21] This dataset provides information on the number of refugees both at the province (*eparchia*) level, i.e., third sub-national administrative units, and at the municipality level. There were 140 provinces and 5,042 municipalities in Greece in 1928. We consider the population share of refugees in 1928 as our main variable of interest. We also draw on the published results of the 1923 refugee census at the municipality level in order to check the robustness of our analysis to using the distribution of refugees in 1923.[22]

The 1928 census provides a rich set of information on the socio-economic characteristics, e.g., literacy rate by age and professions, of the total Greek population and of the refugees separately. It also tabulates the characteristics of the population in 1920 for comparison. However, these information are available only at the province level and city level.[23] We complement these data by drawing on the published results of the 1920 census, which provides information on the population size at the municipality level. Unfortunately, the 1920 census does not provide information on the distribution of population by ethnicity or religion. Therefore, it is not possible to identify the distribution of Muslims and their socio-economic characteristics in 1920 using census data. The 1928 census nonetheless tabulates the number of Muslims and Bulgarians who left Greece between 1920 and 1928, but only at the province and city level.[24]

To identify the municipalities where Muslim/Turkish minority lived, we use an ethnographic map from Güvenç (2010). This map presents the distribution of different ethoreligious groups in Macedonia as of August 1915 (see Figure A.3 in Supplementary Appendix). According to the 1928 census, 86.4 percent of the Muslims and Bulgarians who left between 1920 and 1928 resided in Macedonia. In this region, we match 721 of the 780 municipalities tabulated in the 1920 census

[21] See Figure A.1 in Supplementary Appendix for the questionnaire (bulletin) and an example of historical table.

[22] This census was taken in April 1923, after the mass inflow of refugees but before the Lausanne Peace Treaty and the agreement on the forced population exchange between the Kingdom of Greece and Turkey. Given that we are interested in the effects of the refugee settlement on long-term development, the distribution of refugees in 1928 serves as a better proxy for the final location of refugees following the efforts of the RSC. Table A.1 in Supplementary Appendix presents the timeline of events that are relevant for the settlement and enumeration of refugees.

[23] In the 1920 and 1928 censuses cities are defined either as municipalities that served as the administrative center of a department or as municipalities with more than 10,000 population. There were 59 cities in Greece in 1928 according to this statistical definition.

[24] About 90,000 Bulgarians left Greece, relative to 400,000 expelled Muslims. See footnote 12 for details about the Greek-Bulgarian exchange of populations.

to the localities presented on the map. We code two binary variables: mostly-Muslim locality and some-Muslim locality. The former indicates whether the Muslim population represented about 75 percent or more of the population of a locality, while the latter indicates whether Muslims represented about about 25 percent or more of its population.[25] Figure A.4 in Supplementary Appendix presents the distribution of Muslims and Bulgarians who left between 1920 and 1928 at the province level and the location of mostly-Muslim and some-Muslim municipalities.

In addition, we collect province-level data from the agricultural censuses of 1914 and 1929 in order to investigate the effects of refugees on the level and diversity of agricultural production by crops.[26]

According to the refugee census, there were already 778 thousand refugees in Greece in April 1923. After the population exchange, their number rose up to 1.07 million in 1928, accounting for 16.6 percent of the total population (out of 6.08 million). 61.7 percent of refugees settled in urban areas in 1923 and 53.9 percent in 1928, while only 30 percent of Greek natives lived in cities in 1928. The three biggest cities in Greece, Athens, Piraeus, and Thessaloniki, hosted about half of the urban refugees. We exclude these three cities from the analysis as they may have a different development trajectory and may thus introduce endogeneity. Figure 1 visualizes the distribution of refugees in 1923 and 1928 at the municipality level. The average share of refugees was 4.2 percent in 1923 and 7 percent in 1928 (Panel A of Table 1).[27] More than 60 percent of municipalities in 1923 and almost 50 percent of municipalities in 1928 did not host any refugees (Figure A.5 in Supplementary Appendix).

3.1.2 Contemporary Outcomes

Our primary source of data on contemporary outcomes are the 1971 and 1991 population censuses provided by the Hellenic Statistical Authority (ELSTAT).[28] These censuses provide information on population size, occupational choices of the working age population, and dwelling characteristics at the municipality level. Unfortunately, no information on country of birth is available in the censuses. As a result we cannot distinguish refugees from natives, so these data are an aggregate

[25] We validate the information coming from the ethnographic map by examining the changes in the population between 1920 and 1928. We expect former Muslim localities to experience a substantial inflow of refugees, as free land and housing were abundant due to the departure of Muslims, and a decline in their non-refugee population between 1920 and 1928. 130 of the 133 mostly-Muslim localities experienced a loss in their population – a total of 136,574 inhabitants. Similarly, 178 of the 188 some-Muslim localities experienced a loss in their population – a total of 191,907 in habitants. (These figures do not take into account natural population growth rates and therefore are lowerbound estimates of the outflow of Muslims.) Moreover, the share of refugees in 1928 is 80.1 percent and 68 percent of the population living in mostly-Muslim and some-Muslim localities, respectively.

[26] The 1914 agricultural census covers only Crete and the territories annexed by Greece in 1913, namely Macedonia, Western Thrace, Crete, and Aegean Islands. 93.4 percent of the agricultural refugee settlements were in these regions.

[27] We compute the share of refugees in 1928 as the ratio of the number of refugees in 1928 over the total population in 1928. At the time of the 1923 refugee census, Muslim population of Greece hasn't yet been affected by the population exchange agreement. We thus compute the share of refugees in 1923 as the ratio of the number of refugees in 1923 over the sum of the total population in 1920 and the number of refugees in 1923.

[28] We were provided by ELSTAT a 25 percent extract of the 1971 census and the 100 percent extract of the 1991 census, both aggregated at the municipality level.

of both. We focus on the level of development before 1997 because a major administrative reform, known as the Kapodistrias reform, took place and reorganized the country's administrative divisions in this year.[29] We complement these with data on top-corrected night lights (Bluhm and Krause, 2017). The resolution of the top-corrected night lights data is 30 by 30-arc seconds, which equals about half a kilometer in Greece.

We use population density, luminosity per square kilometer, and dwelling quality as proxies for contemporary local economic development. We compute the share of dwellings with amenities such as electricity, sewage, bathing facilities, and tap water inside the house. We construct a dwelling quality index as the first component of a Principal Component Analysis (PCA) of these shares. To investigate the effects of refugee inflow on the structure of the economy and the labor market, we compute the share of employment in non-agricultural sector and in manufacturing. In the absence of income data in the censuses, we impute labor income by following the common practice of occupation-based income scores.[30] We match the occupation-gender pairs (45 x 2) that exist in the 1991 census to the Greek Labor Force Surveys (GLFS) of 1999–2002. We assign to each pair the average earnings of all employed persons with that particular occupation in the Labor Force Surveys. For simplicity, we often refer to this occupation-based income measure as "earnings". Panel B of Table 1 summarizes the contemporary outcomes considered in the analysis at the municipality level.[31]

3.1.3 Panel Data Over the Period 1920–1991

We match 5,035 of the 5,042 municipalities that existed in 1928 to the 4,852 municipalities that existed in 1920. Also, we match 4,845 (96.1 percent) of the 5,042 municipalities that existed in 1928 to the 5,847 municipalities that existed in 1991 and was a part of Greece as of 1928.[32] We collapse our dataset at the administrative borders of 1920 and treat observations that merged after 1920 as a single unit. As a result, we obtain a panel dataset with 4,388 observations (cluster of municipalities) covering the period between 1920 and 1991, achieving a match of 97.7 percent of the 1928 population to the rest of the dataset.[33]

We also construct a province-level panel dataset covering the period between 1920 and 1991. We split provinces into urban and rural parts to obtain a final dataset of 176 spatial units that are either urban areas (i.e. cities) or rural areas. We complement this dataset with the 1951 population census to obtain occupational data comparable over time, namely in 1920, 1928, 1951, 1971, and 1991.[34]

[29] Between 1928 and 1997, the administrative organization of Greece was very stable and municipality borders did not change drastically. With the Kapodistrias reform, many municipalities were merged to ensure better administration. As a result, the number of municipalities fell down sharply, from 5,775 to 1,033.

[30] See for example Abramitzky, Boustan and Eriksson (2012); Olivetti and Paserman (2015); Bleakley (2007).

[31] Table A.2 in Supplementary Appendix summarizes the outcomes in 1971.

[32] Dodecanese islands joined Greece in 1947 and are therefore excluded from the sample.

[33] See section B.1 in Supplementary Appendix for details of the matching procedure. 385 municipalities that existed in 1920 merged into 171 municipalities between 1920 and 1928, and 59 municipalities that existed in 1928 merged into 16 municipalities between 1928 and 1991.

[34] See section B.3 in Supplementary Appendix for details of the province- and city-level data sources and construction

3.1.4 Greek Economy Over the Period 1920–1991

At the beginning of the twentieth century, Greece was primarily a rural agrarian country. Before the arrival of refugees, in 1920, the urbanization rate was only of 22.9 percent, and more than half of the population was employed in agriculture (Table A.3 in Supplementary Appendix). After the population exchange, the urbanization rate went up to 31.1 percent and the share of employment in primary sector fell from 53.5 percent to 49.3 percent between 1920 and 1928. The urbanization process gained pace after WWII. Between 1941 and 1991, the urbanization rate steadily increased from 32.8 percent to 58.8 percent, with an exceptional 10 percentage points increase between 1961 and 1971. However, the share of employment in agriculture stayed stable around 50 percent until 1961. The Greek economy witnessed a rapid industrialization between 1961 and 1991. During this period, the employment share of the primary sector fell from 53.8 percent to 18.5 percent.

3.2 Determinants of Settlement Locations

What were the economic and geographical factors that affected the settlement location of refugees? Table A.4 in Supplementary Appendix presents the determinants of settlement locations in 1923 and 1928 across municipalities within the same province.[35] In 1923, refugees were more numerous in urban areas, islands, and municipalities close to the shoreline. This is consistent with historical accounts that refugees left Turkey by boats and were initially hosted in public buildings that were converted to shelters. They were more numerous in areas more suitable for agriculture, potentially because these areas were able to feed and host a greater number of people. In 1928, refugees were also more likely to settle in urban areas and islands but to a lesser extent compared to 1923. Also, refugees were resettled in areas more suitable for agriculture, consistent with the RSC's mandate to establish agricultural colonies in which refugees can be self-sufficient.

An substantial relocation of refugees took place between 1923 and 1928. Out of the 4,388 municipalities in our sample, 1,081 experienced a net decrease in the number of refugees between 1923 and 1928.[36] The share of refugees in 1923 explains only 26 percent of the variation in the share of refugees in 1928 across municipalities. Figure A.6 in Supplementary Appendix visualizes the change in the refugee share between 1923 and 1928 at the municipality level. Refugees moved from coastal areas, islands, and Thrace to Macedonia (especially close to the northern national border) and to interiors of Crete. Column 3 of Table A.4 confirms that, in this period, the share of refugees decreased in cities, islands, and municipalities near the shoreline. Another factor that explains the relocation of refugees is the expulsion of Muslims. At the time of the refugee census in April 1923, the population exchange agreement has not yet been signed and Muslims were still present in Macedonia and Crete.[37] The land and houses that Muslims left behind were used for

of our province-level panel dataset.

[35] Panel C of Table 1 summarizes the historical population and geographical characteristics at the municipality level.

[36] Among these localities, a total of at least 201,036 refugees relocated in a different settlement, out of a total of 552,996 refugees present in Greece in April 1923.

[37] Thrace was exempted from the population exchange as it was a part of Bulgaria before WWI. Muslim population in Thrace was thus allowed to stay in Greece.

the resettlement of refugees by the RSC. Between 1923 and 1928, the share of refugees increased by 59 percentage points in mostly-Muslim localities and by 49 percentage points in some-Muslim localities. At the province level, a ten percentage-points increase in the share of Muslims and Bulgarians left is associated with a 3.6 percentage points increase in the share of refugees between 1923 and 1928 (Figure A.7 in Supplementary Appendix). The share of expelled Muslims and Bulgarians also explains 65.6 percent of the variation in the share of refugees in 1928 (and 61.5 percent of the variation in the change in the share of refugees between 1923 and 1928).

4 Empirical Strategy

4.1 Baseline Specification

We estimate the long-term effects of refugee settlements on economic development using the following regression at the municipality level:

$$Y_{mp} = f(Ref_{mp}) + \alpha X'_{mp} + \delta_p + \varepsilon_{mp} \tag{1}$$

where m indexes municipalities and p indexes provinces. Y_{mp} is a measure of contemporary prosperity, such as night light luminosity or earnings, generally measured in 1991.[38] $f(Ref_{mp})$ is a flexible function of the share of refugees in the 1928 population of municipality m. Importantly, the specification includes a set of province fixed-effects δ_p that controls for unobserved heterogeneity at the province level. The vector of covariates at the municipality level, X_{mp}, comprises population density in 1920; a dummy variable indicating whether the locality was a city in 1920; and the geographical variables used as covariates in Table A.4: A dummy variable indicating whether the municipality is on an island; distances to the railway network in 1920, to the shoreline, and to the national border; altitude; average annual temperature and precipitation; and suitability of crops for rain-fed agriculture. We correct the standard errors for spatial correlation across municipalities following Conley (1999), assuming that the spatial correlation is null between municipalities that are more than 100 kilometers apart.[39]

2,300 municipalities – out of 4,388 – hosted some refugees in 1928. There was a substantial variation in the size of the population shock that the municipalities received. The share of refugees was less than 2 percent in half of them while it was more than 40 percent in 300 municipalities (see Figure A.5 in Supplementary Appendix). Given the heterogeneity in the relative size of the shock, we expect the inflow of refugees to have non-linear impacts along the distribution of the refugees' share. We estimate the shape of the function f by using non-parametric regression techniques (discussed in detail in section 5.1).

[38] We focus on outcomes before 1997 because an administrative reform significantly changed administrative divisions of Greece after 1998. See footnote 29 for details and implications of the reform. We report robustness of our estimates using outcomes observed in 1971.

[39] We discuss why we decide on distance threshold of 100 kilometers for the Conley error correction and show the robustness of our results to varying distance thresholds in section 5.3.3.

The inflow of refugees is correlated with the outflow of Muslim population. Ideally, we would like to control for the historical presence of Muslims in equation (1). However, this information is available at the municipality level only for the Macedonia region, where most of the Muslim population of Greece resided in. In section 6.3, we provide evidence that the effects of refugees are not driven by the outflow of Muslims.

4.2 Identification

An important concern with the estimation equation (1) is the self-selection of refugees into their final places of settlement. Refugees could have been attracted to places with better economic opportunities and greater potential for long-run growth. This would cause OLS estimates to be biased upwards. We address this concern in four different ways:

1. Pre-settlement differences. We first document that places of resettlement were not more prosperous before the population exchange. As presented in Table A.4, we find no statically significant difference in the pre-settlement population density and urbanization between municipalities with varying shares of refugees in 1928, conditional on geographic features. While we cannot explore this question any further at the municipality level due to lack of data, we can leverage richer data at the province-level. Table A.5 in Supplementary Appendix shows that provinces that hosted more refugees in 1928 were less developed in 1920, conditional on region fixed-effects and on the covariates displayed in Table A.4. The share of refugees in 1928 is correlated with a lower share of employment in the non-agricultural sector and in manufacturing and with a lower literacy rate in 1920. However, these correlations become insignificant and almost zero when we control for the share of Muslims in 1920, which is a strong predictor of the settlement locations of refugees in 1928 (columns 4 to 6). This finding is consistent with the historical anecdotal evidence that Muslims were less literate and more agriculturalist relative to the Greek population (Glavinas, 2009, pp. 294–310). This suggests that Greek natives had similar occupations and literacy rates in 1920 across provinces with varying shares of refugees.[40]

2. Selection on unobservables. In our baseline municipal-level analysis, we follow the approach of Oster (2019) to quantify how important the selection of refugees on unobservables should be relative to the one on observables in order to explain the estimated effects. As a robustness check, we restrict the analysis to clusters of spatially contiguous neighbours, where a municipality hosted a sizeable refugee population and its neighbors did not, and use a propensity-score matching estimator to deal with the potential selection of refugees on unobservables. We report the results obtained from this estimation strategy in section 5.3.1. Both of these methodologies assume that the selection on unobservables goes in the same direction as the selection on observables. Another way to control for the selection on time-invariant unobservable factors is to use panel data and

[40] We also examine pre-settlement trends in the literacy of the 1920 population. Figure A.8 shows that the average increase in literacy across cohorts born between 1861 and 1905 follows a very similar pattern between high- and low-refugee provinces. Reassuringly, a statistical test of parallel pre-trends cannot be rejected, which suggests that places of settlement were not growing faster prior to refugees' arrival.

estimate a fixed-effects model. Unfortunately, we can't perform such analysis at the municipality level due to the lack of data prior to 1923 (pre-settlement period). Therefore, we use province-level panel data to estimate a fixed-effects model to control for time-invariant unobservables at the province level, as well as for time-varying unobservables at a greater administrative level (region or department) in section 5.1.2.

3. Refugee distribution in 1923. Self-selection on unobservables is arguably less problematic for the distribution of refugees reported in the census of April 1923. At that time, refugees had been in Greece for less than 6 months (since the fire of Smyrna in September 1922), and still hoped to return to their homeland in Turkey as the agreement on a population exchange, which made their settlement in Greece permanent, was not yet signed. Upon arrival, the main worry of refugees, who were in a state of "utter destitution", was to find relief and shelter where they could and less so to seek places with better economic opportunities (Kontogiorgi, 2006, pp. 88). The first place of settlement in 1923 was often temporary, with more than one third of refugees relocating to a different settlement between 1923 and 1928. We report the robustness of our estimates to using the distribution of refugees in 1923 instead of that in 1928 in section 5.3.2.

4. RSC-assigned settlement in Northern Greece. According to the historical accounts, the RSC assigned refugees to agricultural colonies mainly on the basis of vacant land suitable for cultivation. As a result, agricultural refugees couldn't choose their settlement location. More than 80 percent of refugees in Northern Greece were resettled by the RSC in agricultural settlements, while it was the case only for 11 percent of refugees in the rest of Greece. Therefore, focusing on a sample of municipalities in Northern Greece may help to address the self-selection problem. One concern is that the RSC planners might have settled refugees in land parcels that were more fertile for agriculture. Although controlling for the suitability of land for agriculture partly addresses this concern, unobserved characteristics of the agricultural settlements can still bias the OLS estimates upwards if the RSC planners allocated more refugees to places where they foresaw a greater potential for growth. Were such selection effects be driving the results, we would *a priori* expect a linear (or a convex) relationship between the inflow of refugees and long-run prosperity. Instead, as we report in section 5.4, long-run prosperity first increases with the share of refugees in 1928 and then declines when refugees make up more than 40 to 50 percent of the local population. Finding a nonlinear inverted-U shaped relationship between the share of refugees and long-run prosperity supports a causal interpretation of the OLS estimates.[41]

5 Results

5.1 Main Findings

We start the empirical analysis by examining the shape of the function $f(Ref_{mp})$ in specification (1). Figure 2 presents the non-linear effects of refugees on local development, industrialization,

[41] We also find an inverted-U shaped relationship in all Greece, and not only within Northern Greece

and earnings. We report the estimates obtained from semiparametric regressions following Robinson (1988)'s double residual estimator controlling for province fixed effects and the vector of municipality-level covariates described in specification (1). We find, consistently across outcomes, that long-run prosperity increases with the inflow of refugees until their share reaches 40 to 50 percent of the municipal population in 1928, then it declines.[42] This hump-shaped pattern suggests that the non-linear effects of the refugee inflow can be approximated with a quadratic function, which we adopt for the rest of the empirical analysis.

5.1.1 Effects on Local Development

Table 2 presents the effects of refugees on local development in 1991.[43] In the absence of municipal-level data on GDP per capita, we consider population density and luminosity per square kilometer as measures of prosperity. We complement these measures with a standardized index (PCA) of dwelling quality in terms of the following household amenities: the presence of electricity, sewage, tap water, and bathing facilities inside the house.[44] The inflow of refugees has a positive effect at a decreasing rate on all of these proxies of development. The effect of refugees reaches its maximum when the share of refugees in 1928 is between 40.7 and 44.8 percent depending on the outcome. The magnitude of the coefficients suggests that moving a municipality's refugee share from zero to 10 percent (the 87th percentile of the sample) results in an increase in population density by 19.4 percent and in luminosity per square kilometer and dwelling quality by 11 percent of a standard deviation.

Next, we assess whether these estimated effects can be explained by the selection of refugees on unobservables. We compute Oster's delta (Oster, 2019), which quantifies the degree of selection on unobservables relative to observables that should prevail to fully account for the estimated effects. The range of Oster's delta we obtain indicate that selection on unobservables should be 1.2 to 1.7 times greater than the one on observables to fully explain the positive association between the refugees and long-term development.[45]

[42] We find a similar hump-shaped relationship when using outcomes measured in 1971 (instead of 1991). See Figure A.9 in Supplementary Appendix.

[43] Columns 1 and 2 of Table A.6 in Supplementary Appendix present the effects on local development in 1971 for the outcomes that are available for that census year.

[44] Table A.7 and Table A.8 in Supplementary Appendix present the effects for each of the household amenity separately, in 1991 and 1971 respectively. One concern with using household amenities as a proxy for prosperity in this context is that the effects of refugees might be driven by the quality of houses especially built to resettle refugees upon their arrival. We obtain similar results if we look at the effects of refugees on quality of dwellings built between 1945 and 1991, i.e. 20 years later than the arrival of refugees, in Table A.9 in Supplementary Appendix. This suggests that the estimated positive effect of refugees on dwelling quality does not simply arise from the houses built to accommodate refugees.

[45] We focus on the linear effect of the share of refugees on a sample of municipalities with a share of refugees below the turning point of the bell curve, that is, when the marginal effect is still positive. We do so because Oster's procedure does not apply to non-linear (quadratic) treatment effects. The coefficient obtained from a linear model estimated below the maximum point is smaller than the one obtained using a quadratic function in the full sample, consistent with decreasing marginal effects of the refugees' share. Following Oster (2019)'s guidelines, we compute Oster's delta by assuming that the share of the outcome variance (R^2) explained by a hypothetical regression including both observed and unobserved controls is 1.3 times the one already explained by the vector of municipal-level covariates included in the model.

5.1.2 Effects on Structural Transformation

Table 3 presents the effects of refugees on labor market outcomes in 1991, with a particular focus on structural transformation.[46] We focus on a sample of individuals aged between 20 and 65 (who were born after the refugee resettlement).[47] The inflow of refugees has a positive effect at a decreasing rate on the share of employment in non-agriculture and manufacturing (Columns 1 and 2). Similarly, it is positively associated with higher earnings in 1991 (Column 3). The effect of refugees is the greatest when their share in the municipal population is around 40 percent, as it is the case with local development outcomes. The magnitude of the coefficients suggests that moving a municipality's refugee share in 1928 from zero to 10 percent results in an increase in the share of employment in non-agriculture by 3.3 percentage points and in manufacturing by 0.8 percentage points. These effects are substantial: they represent an increase of 6.2% and 14.8% relative to the average in places with no refugees, respectively. Moving a municipality's refugee share from zero to 10 percent translates into an increase in earnings by 1.6 percent. These estimated effects are unlikely to be driven by selection of refugees on unobservables. The estimated Oster's deltas imply that the degree of selection on unobservables should be 6.2 to 46.4 times as important as the one on observed covariates to fully explain the effects.

Timing of the Effects. We turn to province-level panel data to investigate the time-varying effects of the refugee inflow. We estimate the following equation, where each β_t is interpreted as the average effect of refugees in year t relative to the base year of 1920:[48]

$$Y_{pt} = \beta_t Ref_p + \alpha_t X'_p + \delta_p + \mu_t + \mu_t \times \lambda_{r(p)} + \varepsilon_{pt} \tag{2}$$

where p indexes the urban or rural area of a province and t indicates the census year, i.e., either 1920, 1928, 1951, 1971 or 1991.[49] Y_{pt} is the employment share of the non-agricultural sector in locality p in census year t, which is the only outcome that is comparable across different censuses. Ref_p is the share of refugees in locality p in 1928. δ_p and μ_t represent province and census-year fixed effects, respectively. X_p is a vector of geographic and population controls measured in 1920 (the same as in Table A.5), which are interacted with census-year dummies to allow for differential trends across provinces. We also include a set of region/department fixed effects, $\lambda_{r(p)}$, interacted with census-year dummies to control for time-varying regional heterogeneity (there are 10 regions and 35 departments in Greece). This estimation strategy helps to address the concern that refugees self-selected into localities based on unobserved characteristics that are correlated with higher potential for economic growth.

[46] Columns 3 and 4 of Table A.6 in Supplementary Appendix present the effects on labor market outcomes in 1971.

[47] Due to the lack of information on country of birth in the Greek population censuses, we cannot unfortunately analyse the differences in labor market outcomes between refugees and natives. Using survey data, we find in section 6.3 that, in the 2000s, the second-generation of refugees have similar education, occupation, and income to the offspring of Greek-natives.

[48] This approach is similar to that adopted by Hornbeck and Naidu (2014) and Rocha, Ferraz and Soares (2017).

[49] We have a balanced panel of 176 spatial units that are either the province's urban or rural area over the years 1920, 1928, 1971 and 1991. Given that 9 observations are missing for the year 1951, we have $5 \times 176 - 9 = 871$ province-by-year-level observations.

Table 4 present the results. We find positive short-term effects of refugees on the expansion of the non-agricultural sector between 1920 and 1928. The coefficients displayed in column 1 suggest that a 10 percentage point increase in the share of refugees is associated with a 1.37 percentage-points increase in the share of employment in non-agriculture in 1928, relative to the base year of 1920.[50] The size of the effect remains similar in 1951. Then, it doubles to reach 2.7 percentage points in 1971 and further increases to 3.9 percentage points in 1991. Importantly, the effects of refugees are significantly larger in 1971 and 1991 than in 1928 (see p-values of the test of equality of coefficients at the bottom of the table). This implies that resettlement localities experienced a higher *growth* in the non-agricultural sector between 1928 and 1971/1991. In contrast, there is no significant effect on sectoral change between 1928 and 1951. This suggests that it is only when the overall process of structural transformation started in Greece in the 1960s (as documented in section 3.1.4) that provinces with a higher share of refugees experienced greater industrial growth relative to other provinces. The results remain similar when we include region-by-year fixed effects (column 2) or department-by-year fixed effects (column 3), respectively.[51] Overall, these results are consistent with the cross-sectional municipality-level estimates presented in Table 3, indicating that refugees fostered structural transformation away from agriculture in places of resettlement.

5.2 Spatial spillovers

Did the refugee inflow lead to the reallocation of economic activity into places of resettlement? Or, did it lead to the creation of new economic activity? Our municipal-level estimates can a priori capture both of these effects. It is important to disentangle them to better understand the implications of our results for aggregate economic growth. To assess whether our estimates are driven by reallocation effects, we test whether being close to a municipality with more refugees resulted in lower or higher levels of prosperity in the long-run. Were reallocation effects important, we would expect economic activity to relocate from nearby localities to places with more refugees, and therefore long-run prosperity to be lower in localities close to places of resettlement. This approach is similar to the one employed by Sequeira, Nunn and Qian (2019).

For each municipality, we compute the average share of refugees in its spatially contiguous neighbors of first and second degree (i.e neighbors of neighbors) in 1928. On average, municipalities have 18 neighbors defined as such, at an average distance of 9 km (and a maximum distance of 17 km). Table 5 presents the results. The estimated spillover effects are positive for all prosperity outcomes in 1991 and statistically significant at conventional levels for 4 out of 6 of them. Being close to municipalities with more refugees in 1928 was economically beneficial in the long-

[50] The short-term positive impact is not driven by a composition effect that arises due to differences in occupations between refugees and non-refugees (natives). In Table A.10 in Supplementary Appendix, we examine the effect of refugees on the employment of natives in 1928, instead of that of the total population. We find a significant positive effect on natives' employment share in the non-agricultural sector in 1928, relative to 1920.

[51] A concern about our panel analysis is that the employment data in 1920 include Muslim population, which had possibly a higher tendency to work in agriculture than their Greek counterparts. Given that 1920 serves as the base period for the analysis, the outflow of Muslims between 1923 and 1928 could bias the estimates upwards due to a composition effect. We address this concern by checking that the results are robust to including as controls the share of Muslims left between 1920 and 1928 interacted with census-year dummies in Table A.11 in Supplementary Appendix.

term. Importantly, the effects of refugees on their municipality of settlement remain similar in size after controlling for the spillover effects. We also find positive spillover effects when we allow for a quadratic specification in the average share of refugees in neighboring municipalities (Table A.12 in Supplementary Appendix).[52] In addition, we test the sensitivity of the estimated spillover effects to different definitions of neighbors based on varying euclidean distances (from 10 to 40 kms) between centroids of municipalities. As shown in Figure A.10 in Supplementary Appendix, we continue to find positive spillovers. These results suggest that the refugee inflow fostered local development in places of resettlement and the surrounding area instead of displacing economic activity from nearby localities.

5.3 Robustness Checks

In this section, we address concerns of self-selection of refugees into settlement localities with greater potential for long-run growth. We also check the sensitivity of our statistical inference to using different distance bandwidths for Conley error correction.

5.3.1 Spatial Matching Estimator: Contiguous Neighbors

We develop a matching approach that relies on comparisons of spatially contiguous neighbors. A typical municipality in Greece is 30 square kilometers and the distance between centers of two contiguous neighboring municipalities is only 5 kms (and maximum 8 kms). Contiguous municipalities are arguably exposed to similar economic conditions, natural environments, shocks, and other unobservable characteristics. We begin by defining a binary treatment that takes the value of 1 if the share of refugees in 1928 in a municipality is higher than 5 percent and 0 otherwise.[53] This threshold characterizes a treatment group of 642 municipalities, about 15 percent of the sample. To construct the control group, we only consider the untreated municipalities that are spatially contiguous to a treated municipality. Using a propensity score approach, for each treated municipality we further select untreated neighbors that are the most similar in terms of observed geographical features and 1920 population density. Some untreated municipalities serve as a control for different treated municipalities and thus are used more than once in order to construct the counterfactual. To account for this, we reweigh the control sample.[54]

[52] For some outcomes, the quadratic term of the average share of refugees in neighboring localities is negative and statistically significant, implying a negative marginal spillover effects once the average share of refugees in neighboring localities passes 40%. Even in these cases, the spillover effects is positive because no municipality in the sample has an average refugee share among its neighbors higher than 80%. We also find similar results for the prosperity outcomes in 1971 in Table A.13 in Supplementary Appendix.

[53] Given that localities on islands often have no spatial neighbors, we exclude the municipalities in Cyclades, Aegean and Ionian islands from the analysis. We further drop 59 cities from the sample to avoid matching urban municipalities to rural ones. This leaves us with a sample 3,874 municipalities (out of the universe of 4,388 municipalities). We check that the results are robust to using the alternative thresholds of 2% and 10% in terms of share of refugees in 1928 to define the treatment group.

[54] We exclude treated-control municipality pairs in which the difference in the likelihood of being treated (i.e. the pscore) between the treated and the untreated neighbors (due to difference in their geographical characteristics) is higher than a given threshold. We decide on a threshold such that we exclude 20 percent of the treated localities. See Figure A.11 in Supplementary Appendix for the map of treated and control municipalities. The matching estimator of

Table A.14 in Supplementary Appendix presents a balancing test showing the average differences in observables between treated and untreated neighbors. These differences become insignificant in size and in statistical terms when we use propensity-score matching (Panel B). We obtain a sample of neighboring localities with comparable observables, which are thus potentially similar in terms of unobservable characteristics. Panel A of Table A.15 in Supplementary Appendix presents the average treatment effects obtained from our matching estimator. Treated municipalities have on average greater levels of prosperity, larger non-primary and manufacturing sectors, and higher earnings relative to their untreated neighbors. Moreover, the estimated treatment effects are likely to be lower bound given the presence of positive spillover effects of refugees on neighboring municipalities.

5.3.2 Refugee Distribution in 1923

Another way to address the potential self-selection of refugees is to use the distribution of refugees in April 1923. At that time, refugees had been in Greece for less than 6 months. Their principal worries was to find relied and shelter. It is unlikely that they had time to self-select into locations with better economic opportunities and greater potential for long-term growth. Table A.16 in Supplementary Appendix presents the effects of refugees according to their distribution in 1923. We find positive and statistically significant impacts on all prosperity outcomes when using the OLS estimates btained from re-estimating equation (1) (Panel A) or the spatial matching estimator (Panel B).

Furthermore, we also find significant beneficial effects of refugees when we exploit only the intensive margin in the share of refugees, i.e. restricting the analysis to municipalities with more than 0.1% of refugees in 1923 (Panel C), or, alternatively, to municipalities with more than 0.1% of refugees in 1928 (Panel B of Table A.15).

5.3.3 Inference: Conley Error Correction

The inference of our estimates depends on how we treat the spatial autocorrelation in the residuals of our regression. Given that there is no metric to identify the optimal distance threshold to define the spatially correlated clusters, we follow the guidelines of Colella et al. (2019) and present the robustness of our estimates to using various distance thresholds for the Conley error correction in Table A.17 in Supplementary Appendix. For comparison, we also present the standard errors clustered at the province level. Lines (1) to (6) present the standard errors and their significance levels obtained from a different regression and clustering structure. Our results are robust to different ways of treating the standard errors. We obtain estimates that are statistically significant at 5 percent when we use distance thresholds of 20 kms, 50k kms, 100 kms, 200 kms, and 250 kms

the average treatment effect is obtained as: $\frac{1}{N_1}\sum_{T_i=1} Y_i - \sum_{T_j=0}\frac{1}{N_1}\sum_{T_i=1}\frac{d(i,j)}{\sum_{T_j=0} d(i,j)} Y_j$, where Y_i is the outcome of municipality i, T_i the treatment status and N_1 is the number of treated. The regression weight of each control municipality j is given by $\frac{1}{N_1}\sum_{T_i=1}\frac{d(i,j)}{\sum_{T_j=0} d(i,j)}$, where $d(i,j)=1$ if the municipalities i and j are spatially contiguous neighbors and zero otherwise.

for the Conley error correction and also when we cluster the standard errors at the province level. We use 100 kms as the baseline distance threshold to define the spatial clusters as it gives us the largest and most conservative standard errors for most of the outcomes.

5.4 Regional Heterogeneity

Next, we investigate the heterogeneous effects of refugees by regions. We begin by documenting the differences in the characteristics of refugees and natives (i.e. non-refugees) across different regions. Table 6 shows that, overall, the literacy rate and occupations of refugees were quite similar to those of natives in 1928 on average (column 1). These similarities mask a considerable level of heterogeneity across regions (columns 2 and 3). In Northern Greece, which hosted three fourth of the newcomers and where the RSC operated, refugees were mostly farmers (79 percent) – slightly more than natives were (74 percent). Also, they were less literate compared to natives (46 vs. 50 percent).[55] In contrast, refugees who settled in the rest of Greece (i.e. Central Greece and Islands) were mostly non-farmers and twice more likely to work in the non-agricultural sector relative to natives (67 percent vs 30 percent). In particular, they were over-represented in the manufacturing compared to natives (35 percent vs. 13 percent).[56] They were also over-represented in high-skilled occupations, such as banking and finance or liberal occupations (e.g. engineers, scientists, doctors, lawyers).

Table 7 examines the heterogenous impacts of refugees by regions. Panels A and B report the quadratic effects of refugees for a sample of municipalities in Northern Greece and in the rest of Greece, i.e. Central Greece and Islands, respectively.[57] Figure A.13 in Supplementary Appendix visualizes the estimated marginal effects.[58] We find that refugees had significant positive effects in both regions. The benefits brought by refugees are higher in Central Greece than in Northern Greece for all prosperity outcomes in 1991 (expect for the share of employment in manufacturing).[59] For example, the coefficients in column 6 suggest that moving a municipality's share of refugees from zero to 10 percent leads to a 4.6 percent increase in earnings in Central Greece compared to a 1.1 percent increase in Northern Greece. These findings are consistent with the fact that the refugees who settled in Central Greece were more educated and skilled than those who settled in Northern Greece, both in absolute terms and relative to natives.

[55] This difference is not driven by the fact that some refugees did not speak Greek as literacy was reported in respondent's own language in the census.

[56] In Greece, refugees in the manufacturing sector were mostly working in the textile and leather industry (32 percent), in the wood industry (7 percent) and metallurgy (8 percent), as well in the tobacco industry (15 percent).

[57] See Figure A.12 in Supplementary Appendix for the distribution of refugees by regions.

[58] Note that we do not display the marginal effects for municipalities where refugees made up more than 60 percent of the population because such municipalities do not exist in Central Greece.

[59] We also find larger beneficial effects of refugees in Central Greece in 1971 than we find in Northern Greece – see Figure A.14 in Supplementary Appendix.

6 Mechanisms

In this section, we consider several mechanisms underpinning our results. Given the abundant literature on the benefits of high-skilled immigration, the positive impacts of refugees in Central Greece are largely expected.[60] In contrast, the fact that refugees brought long-run benefits in Northern Greece despite having lower levels of human capital relative to natives requires further explanations. We focus on shedding light on the mechanisms behind the positive effects of refugees within Northern Greece. We first present the mechanisms that are consistent with the data and then discuss several alternative explanations for which we find no empirical evidence.

6.1 Skill Complementarity in Northern Greece

The historical literature documents that refugees introduced new technological knowledge in textile and brought new agricultural know-how in Greece.

Transfer of Technological Knowledge in Textile. Historians largely credit the boom of the textile industry in Greece in the 1920s, and in particular the development of the carpet industry, to the inflow of refugees. The refugees introduced new weaving techniques and provided expertise in the cultivation and manufacture of new woolen materials and silk fabrics (Kritikos, 2005; Pentzopoulos, 1962).[61] Did this transfer of technological knowledge lead to higher rates of industrialization in places of resettlement? Did this provision of new expertise in textile produce higher benefits in places with higher agro-climatic suitability for the cultivation of crops used in the textile sector, such as cotton?

We test these hypotheses in Table 8. The share of refugees in 1928 is significantly and positively associated with a higher employment share in the textile sector in 1991 (column 1). Importantly, the positive effect of refugees on the size of textile sector is larger in municipalities that are more suitable for the cultivation of cotton (column 2).[62] The estimated coefficients imply that moving the share of refugees in 1928 from zero to 10 percent increases the share of employment in textile in 1991 by 0.86 percentage points among municipalities highly suitable for cotton – which equals 22 percent of the sample average – and by 0.66 percentage points (17 percent) among other municipalities. Furthermore, the beneficial effects of refugees on the expansion of the non-primary sector and manufacturing are significantly more pronounced in places with a higher suitability for cotton (columns 3 and 4). The impact of refugees on the employment share in manufacturing is 50 percent larger in these municipalities relative to other municipalities. Moreover, greater suitability

[60] See Kerr et al. (2016) for an overview of the literature on the productivity gains generated by present-day high-skilled immigration. For the contributions of high-skilled migrants in historical settings, see Hornung (2014), Rocha, Ferraz and Soares (2017), Droller (2018), Moser, Voena and Waldinger (2014) among others.

[61] Refugees accounted for half of the workforce in the carpet industry in Northern Greece despite representing only one fourth of region's total population in 1928.

[62] High suitability for cotton is a dummy variable indicating whether a municipality is in the top tercile of the sample distribution of the (continous) suitability index for cotton cultivation. Greece is a cotton exporting country and has become the largest cotton producer in Europe since the 1960s (Vouras, 1963). Also, it seems that cotton in Greece is more likely to be spun or woven where it is grown, as the positive correlation between the (continuous) suitability index of cotton cultivation and the 1991 share of employment in textile sector suggests (not reported).

for cotton does not capture the positive effects of greater suitability for agriculture in general. The results are robust to controlling for the interaction term between the share of refugees in 1928 and the suitability index for the cultivation of commonly grown cereals in Greece, i.e., wheat, oat, and barley (Table A.18 in Supplementary Appendix).

Overall, these results suggest that the transfer of textile-specific knowledge was conducive to greater levels of industrialization. This was the case especially in municipalities with a greater agro-climatic suitability for the cultivation of textile crops, where refugees' skills could be more effectively transferred and employed.[63] Our findings are consistent with the evidence provided by Bazzi et al. (2016) that the transferability of migrants' skills depend on the similarity in agro-climatic conditions between places of origin and destination.

Provision of Agricultural Know-how. Refugees of Northern Greece were mostly farmers living in agricultural settlements. According to historical accounts, refugees introduced new varieties of crops, such as cotton, tobacco, and cereals (e.g. millet). They are also credited for having superior know-how in farming practices, such as polyculture (instead of monoculture) and the use of forage plants. In Table 9, we examine the effects of refugees on agricultural production (gross output in quintal) drawing on the agricultural censuses of 1914 and 1929.[64] Controlling for province and year fixed-effects, we find that the refugee inflow led to a significant increase in the production of cash crops between 1914 and 1929, in particular in tobacco and cotton (columns 1, 2, and 3). The estimated coefficients imply that a ten percentage-point increase in the share of refugees in 1928 leads to a 23.9 percent increase in the production of cash crops. To assess whether refugees effectively introduced new crop varieties, we examine their impacts on the diversity of cultivated varieties within three types of crops: cash crops, cereals, and leguminous staple crops. Following Fiszbein (2017), we measure diversity as 1 minus Hirschman-Herfindahl index of the shares of each variety in the total agricultural output of the crop type.[65] We find that the inflow of refugees significantly increased agricultural diversity for all three types of crops (columns 4, 5, and 6). A ten percentage-point increase in the share of refugees in 1928 translates into a 4.2 percentage-points increase in the diversity of cash crops, which equals 16 percent of the sample average.

Diversity, Skill Complementarity, and Segregation. Although refugees' know-how might be beneficial on their own, they are more likely to be valuable when, combined with those of natives, they increase the diversity of potentially complementary skills.[66] This implies that interactions between

[63] The transfer of textile-specific knowledge and the associated development of the textile industry may have fostered higher economic growth through positive input-output spillovers along the production chain. The textile sector represented 16.1% of the total industrial GDP in Greece in 1991 (Gouzi, 2017, p. 71).

[64] We focus on a sample of provinces covered by both censuses, mostly within Northern Greece. The 1914 agricultural census covers only the territories annexed by Greece in 1913, namely Macedonia, Western Thrace, Crete, and Aegean Islands. The data is available for 63 different provinces, for a total of 126 province-year observation in the panel dataset. 36 provinces out of 63 are located in Northern Greece. A quintal is a unit of weight that equals to 100 kg. See section B.5 in Supplementary Appendix for more details.

[65] For example, the census reports 11 different varieties of cereals. We compute the diversity of cereals as: $Diversity_{cereals} = 1 - \sum_{i=1}^{11} \left(\frac{prod_i}{\sum_{k=1}^{11} prod_k} \right)^2$, where $prod_i$ is production level of crop i, for i = 1, 2, ..., 11. See section B.5 in Supplementary Appendix for the list of varieties for each crop type.

[66] A large literature has documented the economic benefits of ethnic, cultural, and birthplace diversity (Ottaviano and

refugees and natives should play an important role in the realization of the economic benefits of refugees. The first testable hypothesis of this mechanism is that we should observe larger benefits in places with more heterogenous population (and thus more diverse skills) than in more homogenous localities. The second testable hypothesis is that, for a given size of the refugee population, we should observe smaller economic benefits when refugees and natives are segregated into separate settlements within a municipality. As spatial segregation is likely to erect barriers to interactions, it may impede knowledge transfers and thus lower the productivity gains arising from skill diversity.[67]

We begin by exploring the role of population diversity. Figure A.15 and Figure A.16 in Supplementary Appendix present the non-linear effects of refugees across municipalities in Northern Greece using the same semiparametric regressions as presented in Figure 2 for whole Greece. We find a hump-shaped relationship between the share refugees in 1928 and long-run prosperity, consistently across outcomes in 1971 and 1991. Prosperity increases with the share of refugees until it reaches its maximum in municipalities with around 50 percent refugees, where the population is the most diverse. Past this point, prosperity declines as the share of refugees increases and diversity decreases. OLS estimates in Table A.19 in Supplementary Appendix confirm that municipalities with more than 60 percent refugees display significantly lower prosperity relative to municipalities that hosted between 20 percent and 60 percent of refugees.[68] We interpret this observed pattern as supportive evidence that the increase in the diversity of know-how contributed to shape the beneficial impacts of the refugees. Had the positive effects of refugees been driven by their more advanced know-how, we would have observed greater prosperity where they made up the majority of the population.

Next, we turn to the effects of spatial segregation between refugees and natives. Drawing on the 1928 census data at the settlement level (i.e. infra-municipal level), we measure spatial segregation within each municipality using the standard Dissimilarity Index, as proposed by Reardon and Firebaugh (2002).[69] Table 10 presents the long-run effects of segregation on a sample of 555

Peri, 2006; Ager and Brückner, 2013; Alesina, Harnoss and Rapoport, 2016; Docquier et al., 2019). With respect to the benefits of agricultural diversity in particular, Fiszbein (2017) shows that pre-industrial diversity in crop varieties (induced by climatic features) led to greater long-run economic growth in the U.S. This is mainly because the diversity of skills required by the cultivation and processing of different crops fostered technological progress and industrialization through cross-product knowledge spillovers and recombining innovation – that is, the production of new ideas by combining multiple old ideas. Agricultural diversity can foster development also through reduced exposure to product-specific shocks (which decreases volatility and risk).

[67] An important strand of literature has provided evidence on the importance of social learning in agriculture, i.e., the fact that farmers learn how to adopt new crop and new technologies from their neighbors' experiences (Conley and Udry, 2010; Foster and Rosenzweig, 1995; Munshi, 2004).

[68] A concern with this empirical test is that the municipalities with share of refugees above 60 percent may have unobserved attributes that are disadvantageous for long-run economic growth. We gauge the extent of this problem by examining selection on observables. Column 1 of Table A.20 reveals that settlements with more than 60 percent refugees are significantly more elevated relative to settlements that hosted between 20 percent and 60 percent refugees. When we control for altitude in a quadratic form as we do in Table A.19, we find no statistically significant differences in other observables between the two types of resettlement localities, i.e., those with more than 60 percent refugees and the ones with between 20 percent and 60 percent refugees.

[69] The Dissimilarity Index for a given municipality is equal to $\frac{1}{2p(1-p)} \sum_{j=1}^{J} \frac{t_j}{T} |p_j - p|$ where p_j is the share of refugees in settlement j 's population, p is the share of refugees in the entire municipality, $\frac{t_j}{T}$ is the proportion of village j's population in the total municipality's population T, and J is the number of settlements. The index varies between 0 and 1, where zero corresponds to perfect integration and 1 to perfect segregation. In our sample, the index varies from

municipalities in Northern Greece that were composed of more than one settlement and received a positive number of refugees.[70] We find that, conditional on the share of refugees, a greater spatial segregation is significantly associated with lower prosperity in 1991.[71] A one standard-deviation increase in the dissimilarity index results in a reduction in earnings by 0.85 percent, which equals 10 percent of a standard-deviation. This effect is sizeable relative the impact of refugees on earnings: it reduces the impact of a one-standard deviation increase in the share of refugees (from 1 percent to 32 percent) by 40 percent. This evidence is suggestive of the important role of knowledge transfers in shaping the long-term benefits of refugees.

6.2 Agglomeration Economies

The refugee inflow represented an important increase in the population density and market size. This could lead to positive long-run effects on development through agglomeration effects and increasing returns to scale in the production, as suggested by the economic geography literature.[72] Following Combes, Démurger and Li (2015), we test this mechanism by controlling for the post-resettlement population density, namely in 1928, in the regressions. Theoretically, once the population density is controlled for, the coefficients on the share of refugees in 1928 only capture the effects of skill complementarity purged of agglomeration effects.

Table 11 presents the estimates.[73] The effects of refugees remain significantly positive across all prosperity outcomes in 1991, both in the entire sample (Panel A) and within Northern Greece (Panel B). Importantly, the impacts of refugees remain comparable in magnitude, although slightly smaller in size, relative to the baseline estimates without the controls for population density in 1928. The reduction in the point estimates implies that the agglomeration effects accounts for between one sixth and one third of the total impact of refugees depending on the outcome. For example, agglomeration economies explains one fifth of the total impact of the refugees on the share of employment in manufacturing and earnings (column 5 and 6).[74] Overall, we find evidence that some of the effects of refugees operated through agglomeration economies. However, agglomeration economies explain only a small fraction of the effect of refugees on sectoral transformation.

0.002 to 1, with an average of 0.43 and a standard deviation of 0.26. Figure A.17 in Supplementary Appendix shows that there is no statistical relationship between segregation and the share of refugees in the municipal population. The correlation coefficient between the two variables is -0.04.

[70] For details on the sample restrictions for the segregation analysis, see section B.4 in Supplementary Appendix.

[71] The only outcome for which segregation does not have a statistically significant impact is the share of employment in manufacturing in 1991. Table A.21 in Supplementary Appendix presents the effects of spatial segregation on development and labor market outcomes in 1971. We find a statistically significant negative effect of segregation on share of manufacturing in 1971.

[72] For example, see Ciccone and Hall (1996); Combes et al. (2010); Baldwin and Martin (2004).

[73] We control for the 1928 population density in a flexible way by including ten dummy variables corresponding to the deciles of the population density distribution in 1928.

[74] In all Greece (Panel A), the quadratic coefficients are virtually identical and the linear coefficient falls from 0.93 to 0.73 for manufacturing (column 5) and from 0.179 to 0.143 for earnings (column 6), that is, a 20% reduction for both outcomes. The reduction in point estimates in similar in Northern Greece.

6.3 Alternative mechanisms

Economic Legacy of Muslims. Many refugees were resettled in villages left vacant by the Muslims who were all expelled from Greece after 1923. Therefore, our estimates could also capture the economic legacy of Muslim minorities in Greece or the effect of the ethno-religious homogenization resulting from the population exchange. In Table A.22, we reestimate the effects of refugees controlling for whether a municipality was mostly inhabited by Muslims in 1915.[75] The estimated effects are significantly positive across all prosperity outcomes and remain very similar in magnitude to the baseline estimates obtained without controlling for the presence of Muslims, as shown at the bottom of the table. This holds for both the entire sample (Panel A) and within Northern Greece (Panel B). Moreover, conditional on a given share of refugees in 1928, former Muslim villages display lower level of long-run prosperity relative to other localities.[76] We obtain similar findings when using province-level data on the share of Muslims who left Greece between 1920 and 1928, as presented in Table A.23 in Supplementary Appendix. These results suggest that the impacts of refugees are not driven by the expulsion of Muslims.

Infrastructure Investments. Although most of the RSC's funds were used to provide refugees with cattle, farm inputs, and houses in the agricultural settlements, some were also spent on infrastructure works. These works mainly consisted of the provision of health facilities and the reclamation of marshlands. We examine whether such public investments could partly explain our findings. These public utility programs were mostly carried out in the Macedonia region, which received about two-third of the RSC's expenditures (Pentzopoulos, 1962, p. 110). Would such infrastructure works drive our results, we would expect higher benefits of refugees' presence in Macedonia relative to other regions of Northern Greece. Table A.24 in Supplementary Appendix shows that it is the opposite: the effects of refugees are lower in Macedonia than in other regions of Northern Greece. Nevertheless, we test whether, within Macedonia, the municipalities that received these public investments benefited more from the presence of refugees. First, we leverage geocoded data on the 63 dispensaries that the RSC built in rural Macedonia to curb the spread of malaria and tuberculosis, which were rife among refugees.[77] Panel A of Table A.25 in Supplementary Appendix presents the effects of the RSC's dispensaries. Including in the regression controls a dummy variable indicating the presence of a dispensary within a 10-km radius from the municipality in 1929 and its interaction with the share of refugees in 1928 doesn't affect the estimated effects of refugees within

[75] We obtain information on the distribution of Muslim localities from an ethnographic map published in Güvenç (2010). We code a dummy variable "Mostly-Muslim locality in 1915" taking one if Muslims represented about 75 percent or more of the local population. This map covers the municipalities in the Macedonia region, where 86.4 percent of all Muslims of Greece resided in. For municipalities that are not covered by the map (3,683 out of 4,388), we code this dummy variable either as zero if the share of Muslims in the municipality's province is below 0.1% of the 1920 population (2,643 cases), or as missing otherwise (1,040 cases). This leaves us with a sample of 3,348 municipalities in all Greece and with a sample of 994 municipalities in Northern Greece.

[76] We also find very similar results at the municipal level when we use the dummy variable "Some-Muslim locality" taking one if a municipality's population was composed of more than 25 percent of Muslims in 1915.

[77] See Kontogiorgi (2006, p. 275) for the location of the RSC-operated dispensaries in rural settlements of Macedonia, in April 1929. Each dispensary was staffed by a circuit doctor who provided his services to all villages within a fixed radius and at regular intervals. Besides the provision of health care and medicines, these facilities educated the refugees on the rudiments of preventive medicine and hygiene.

Macedonia. Moreover, we find no systematic evidence of higher benefits of refugees in places closer to a dispensary. Second, we examine the role of the reclamation works that the RSC undertook to provide arable land for refugees.[78] Panel B of Table A.25 presents the results. We find no systematic evidence of positive effects of reclamation works or of greater long-run benefits in resettlement localities within a 10-km radius from reclaimed marshlands.[79]

Second-Generation Refugees. Although our main estimates only focus on a sample of people born after the refugee inflow, if the second-generation refugees became more economically successful than the descendants of natives, this could partly explain our results. For instance, Becker et al. (2019) find that descendants of displaced Poles invested more in education as a result of a shift in preferences towards human capital. We address this question by using the European Social Survey (ESS) that provides information on parents' birthplace for a sample of individuals living in Greece in the 2000s. We identify second-generation refugees as individuals born in Greece with a mother or a father born in Turkey. We focus on the cohorts born between 1923 and 1950, who are more likely to be the children of the first-generation refugees compared to other cohorts.[80] Table A.27 presents the mean differences in educational attainment, last (or current) occupation and deciles of household income between individuals with a parent born in Turkey and those with parents born in Greece.[81] We find no evidence that the second-generation refugees fare better than the offsprings of Greek natives, irrespective of the empirical specification. If anything, the second-generation refugees were less likely to have a high-skilled profession relative to natives. This suggests that our main results are unlikely to be driven by differences in human capital levels and labor market outcomes between refugees' and natives' offsprings.

7 Conclusion

The UNHCR estimates that 70 million people are currently displaced from their home as a result of conflicts or natural disasters, with 85% of them being hosted in developing countries (UNHCR, 2018). While the short-term effects of refugee inflows are well documented, little is known about their long-run impacts on the receiving countries, especially in the developing world. To fill this gap, we examine the long-term consequences of the 1923 mass refugee inflow in Greece, a poor agrarian country with low state capacity at that time. We estimate the effects of refugees on eco-

[78] According to (Kontogiorgi, 2006, p. 277–296), the RSC's undertook reclamation works in the valleys of the Strymon river and the Axios (Vardar) river, as well as the drainage (or damming) of the following six lakes: Lake Achinos, Lake Doirani, Lake Kopais, Lake Artzan, Lake Amatovon, and Lake Yannitsa.

[79] We obtain similar results when use a 5-km or a 20-km radius.

[80] Table A.26 in Supplementary Appendix shows that the proportion of ESS respondents with parents in Turkey is the highest for people born in the 1920s, the 1930s, and the 1940s. It dramatically drops for people born after 1950. In contrast, the share of individuals with parents born outside of Turkey remains stable across birth cohorts, with a peak in the 1990s. This provides empirical support that we are able to identify second-generation refugees.

[81] Given that we exclude Athens, the Piraeus and Thessaloniki from the main analysis, we also exclude individuals living in big cities in the ESS sample. Among the final sample of 1715 individuals, 11% have parents born in Turkey. We also systematically include in the set of controls a dummy variable indicating whether the individual's parents were born in a foreign country other than Turkey in order to compare second-generation refugees with children of Greek-natives. Results are similar if we do not include this control.

nomic prosperity 60 years after this event by using a battery of empirical strategies that rely on different margins of spatial and temporal variation in the refugee presence. All estimates point to the same finding: localities that received more refugees in 1923 display higher levels of prosperity and industrialization in 1971 and 1991.

Importantly, we find beneficial effects of the resettlement in Northern Greece, where most of the refugees were farmers with lower literacy than natives. Furthermore, we uncover a hump-shaped relationship between the share of refugees and long-run prosperity, with the resettlement producing smaller gains in places with more than 50 percent of refugees in the local population. The economic gains brought by the refugees in Northern Greece appear to arise from two channels: (i) the transfer of new technological knowledge, in particular in textile, which was conducive to industrialization; (ii) the provision of new agricultural know-how that increased the diversity of skills in the population. These findings indicate that, by providing complementary – and not necessarily superior – knowledge, refugees can generate significant benefits in the long-run, even when they do not bring higher human capital relative to the host native population.

The League of Nations intervened for the first time in history to organize the resettlement, provide funds and establish refugees in agricultural work. In addition to land titles, refugees were given the citizenship. This was part of a nation-building policy that aimed at their full assimilation into the Greek society. Such resettlement efforts and integration policy are likely to have fostered refugees' economic contributions to the receiving communities. This would suggest that initial investments in refugees can have positive returns in the long-run.

Our findings may have broader implications for the design of resettlement policies aimed at maximizing refugees' contributions to the host economy in the long-run. We uncover that the resettlement had smaller economic gains in places where the local post-settlement population was less diverse, i.e., predominantly refugees, and in municipalities where refugees and natives were segregated into separate villages. These results highlight the crucial role of social interactions and the ensuing knowledge sharing, which in our context were likely to facilitated by the fact that refugees and natives shared the same religion and often spoke the same language. Also, refugees led to higher benefits in localities where their knowledge could be more effectively transferred and employed, due to favorable agroclimatic conditions. Taken together, our findings suggest that resettlement policies should avoid clustering refugees in separate enclaves while trying to match them to locations where their skills are more easily transferable.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *American Economic Review*, 102(5): 1832–56.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration." *Journal of Political Economy*, 122(3): 467–506.

Ager, Philipp, and Markus Brückner. 2013. "Cultural diversity and economic growth: Evidence from the US during the age of mass migration." *European Economic Review*, 64(C): 76–97.

Alesina, Alberto, and Eliana La Ferrara. 2005. "Ethnic diversity and economic performance." *Journal of economic literature*, 43(3): 762–800.

Alesina, Alberto, Johann Harnoss, and Hillel Rapoport. 2016. "Birthplace diversity and economic prosperity." *Journal of Economic Growth*, 21(2): 101–138.

Alix-Garcia, Jennifer, and David Saah. 2010. "The Effect of Refugee Inflows on Host Communities." *World Bank Economic Review*, 24(1): 148–170.

Arbatli, Cemal Eren, and Gunes Gokmen. 2016. "Minorities, Human Capital and Long-Run Development: Persistence of Armenian and Greek Influence in Turkey." CESifo Working Paper Series.

Baez, Javier E. 2011. "Civil wars beyond their borders: The human capital and health consequences of hosting refugees." *Journal of Development Economics*, 9(2): 391–408.

Bahar, Dany, and Hillel Rapoport. 2018. "Migration, knowledge diffusion and the comparative advantage of nations." *The Economic Journal*, 128(612): F273–F305.

Baldwin, Richard E, and Philippe Martin. 2004. "Agglomeration and regional growth." In *Handbook of regional and urban economics*. Vol. 4, 2671–2711. Elsevier.

Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong. 2016. "Skill transferability, migration, and development: Evidence from population resettlement in Indonesia." *American Economic Review*, 106(9): 2658–98.

Becker, Sascha O, and Andreas Ferrara. 2019. "Consequences of forced migration: A survey of recent findings." *Labour Economics*, 59: 1–16.

Becker, Sascha O, and Ludger Woessmann. 2009. "Was Weber wrong? A human capital theory of Protestant economic history." *The Quarterly Journal of Economics*, 124(2): 531–596.

Becker, Sascha O., Irena Grosfeld, Pauline A. Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya. 2019. "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers." National Bureau of Economic Research Working Paper 24704.

Bharadwaj, Prashant, and Rinchan Ali Mirza. 2019. "Displacement and development: Long term impacts of population transfer in India." *Explorations in Economic History*, 73.

Bleakley, Hoyt. 2007. "Disease and development: evidence from hookworm eradication in the American South." *The Quarterly Journal of Economics*, 122(1): 73–117.

Bluhm, Richard, and Melani Krause. 2017. "Top Lights: Bright cities and their contribution to economic development." *mimeo*.

Braun, Sebastian, and Michael Kvasnicka. 2014. "Immigration and structural change: Evidence from post-war Germany." *Journal of International Economics*, 93(2): 253–269.

Calderón-Mejía, Valentina, and Ana María Ibáñez. 2015. "Labour market effects of migration-related supply shocks: evidence from internal refugees in Colombia." *Journal of Economic Geography*, 16(3): 695–713.

Ciccone, Antonio, and Robert E Hall. 1996. "Productivity and the Density of Economic Activity." *The American Economic Review*, 86(1): 54–70.

Clemens, Michael, Cindy Huang, Jimmy Graham, et al. 2018. "The Economic and Fiscal Effects of Granting Refugees Formal Labor Market Access." CGD Report No. 496. Washington DC: Center for Global Development.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2019. "Inference with Arbitrary Clustering." *IZA Discussion Paper No 12584*.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2010. "Estimating Agglomeration Economies with History, Geology, and Worker Effects." In *Agglomeration Economics*. 15–66. University of Chicago Press.

Combes, Pierre-Philippe, Sylvie Démurger, and Shi Li. 2015. "Migration externalities in Chinese cities." *European Economic Review*, 76: 152–167.

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

Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *The American Economic Review*, 100(1): 35–69.

Docquier, Frédéric, Riccardo Turati, Jérôme Valette, and Chrysovalantis Vasilakis. 2019. "Birthplace diversity and economic growth: Evidence from the US states in the Post-World War II period." *Journal of Economic Geography*, forthcoming.

Droller, Federico. 2018. "Migration, population composition and long run economic development: Evidence from settlements in the pampas." *The Economic Journal*, 128(614): 2321–2352.

Fiszbein, Martin. 2017. "Agricultural diversity, structural change and long-run development: Evidence from the us." National Bureau of Economic Research Working Paper 23183.

Foster, Andrew D, and Mark R Rosenzweig. 1995. "Learning by doing and learning from others: Human capital and technical change in agriculture." *Journal of Political Economy*, 103(6): 1176–1209.

Glavinas, Giannis N. 2009. "Οι μουσουλμανικοί πληθυσμοί στην Ελλάδα (1912-1923). Από την ενσωμάτωση στην ανταλλαγή [Muslim Populations in Greece (1912-1923)]." , ed. Thessaloniki Ant. Stamoulis Publications. Giannis N. Glavinas.

Gouzi, Vincent. 2017. "L'Industrie Grecque de la reconstruction à la crise 1950-2014: transformation du produit et permanence des structures d'emploi." PhD diss. Sorbonne Paris Cité.

Güvenç, Sefer. 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.

Hirschon, Renée. 2003. *Crossing the Aegean: An Appraisal of the 1923 Compulsory Population Exchange Between Greece and Turkey*. Vol. 12 of *Studies in Forced Migration*, Berghahn Books.

Hornbeck, Richard, and Suresh Naidu. 2014. "When the levee breaks: black migration and economic development in the American South." *American Economic Review*, 104(3): 963–90.

Hornung, Erik. 2014. "Immigration and the diffusion of technology: The Huguenot diaspora in Prussia." *American Economic Review*, 104(1): 84–122.

Huang, C, and J Graham. 2018. "Are refugees located near urban job opportunities." CGD Report. Washington DC: Center for Global Development.

Kerr, Sari Pekkala, William Kerr, Çağlar Özden, and Christopher Parsons. 2016. "Global Talent Flows." *The Journal of Economic Perspectives*, 30(4): 83–106.

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

Kritikos, George. 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, Georgios. 2005. "The agricultural settlement of refugees: A source of productive work and stability in Greece, 1923-1930." *Agricultural history*, 321–346.

Montalvo, Jose G., and Marta Reynal-Querol. 2007. "Fighting against Malaria: Prevent Wars while Waiting for the "Miraculous" Vaccine." *The Review of Economics and Statistics*, 89(1): 165–177.

Morales, Juan S. 2018. "The impact of internal displacement on destination communities: Evidence from the colombian conflict." *Journal of Development Economics*, 131: 132–150.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish émigrés and US invention." *American Economic Review*, 104(10): 3222–55.

Munshi, Kaivan. 2004. "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution." *Journal of Development Economics*, 73(1): 185–213.

Olivetti, Claudia, and M Daniele Paserman. 2015. "In the name of the son (and the daughter): Intergenerational mobility in the united states, 1850-1940." *American Economic Review*, 105(8): 2695–2724.

Ortega, Francesc, and Giovanni Peri. 2014. "Openness and income: The roles of trade and migration." *Journal of International Economics*, 92(2): 231–251.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics*, 37(2): 187–204.

Ottaviano, Gianmarco IP, and Giovanni Peri. 2006. "The economic value of cultural diversity: evidence from US cities." *Journal of Economic Geography*, 6(1): 9–44.

Pascali, Luigi. 2016. "Banks and development: Jewish communities in the Italian Renaissance and current economic performance." *Review of Economics and Statistics*, 98(1): 140–158.

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

Peri, Giovanni. 2012. "The effect of immigration on productivity: Evidence from US states." *Review of Economics and Statistics*, 94(1): 348–358.

Peters, Michael. 2017. "Refugees and Local Agglomeration-Evidence from Germany's Post-War Population Expulsions." mimeo.

Reardon, Sean F, and Glenn Firebaugh. 2002. "Measures of multigroup segregation." *Sociological methodology*, 32(1): 33–67.

R., J. 1944. "The Exchange of Minorities and Transfers of Population in Europe since 1919." *Bulletin of International News*, 21(15): 579–588.

Robinson, P. M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica*, 56(4): 931–954.

Rocha, Rudi, Claudio Ferraz, and Rodrigo R Soares. 2017. "Human capital persistence and development." *American Economic Journal: Applied Economics*, 9(4): 105–36.

Ruiz, Isabel, and Carlos Vargas-Silva. 2013. "The Economics of Forced Migration." *The Journal of Development Studies*, 49(6): 772–784.

Ruiz, Isabel, and Carlos Vargas-Silva. 2015. "The labour market consequences of hosting refugees." *Journal of Economic Geography*, 16(3): 667–694.

Sarvimäki, Matti. 2011. "Agglomeration in the periphery." mimeo.

Schumann, Abel. 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.

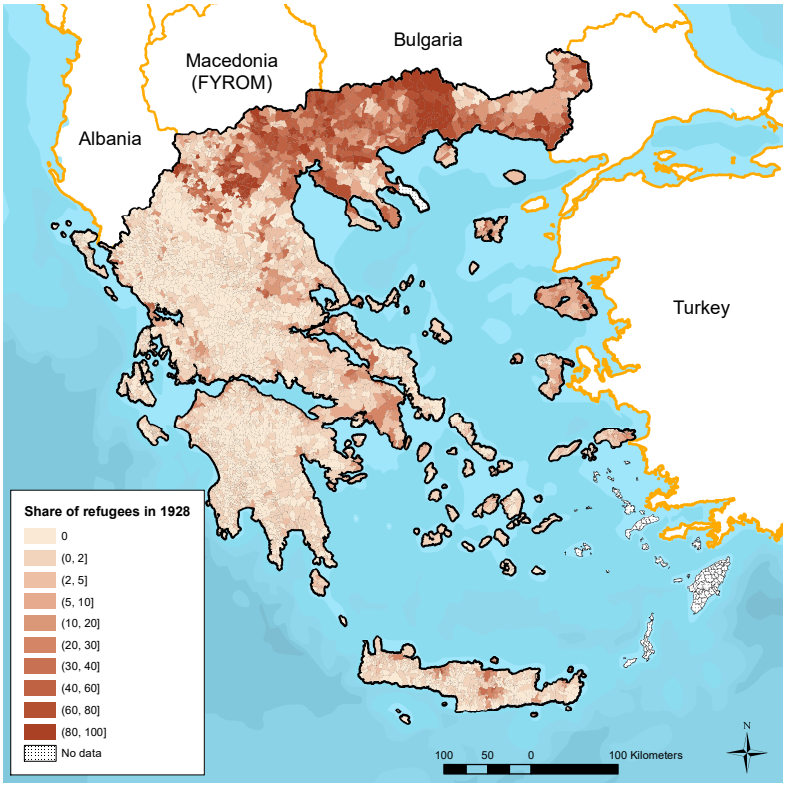
Sequeira, Sandra, Nathan Nunn, and Nancy Qian. 2019. "Immigrants and the Making of America." *Review of Economic Studies*, forthcoming.

UNHCR. 2018. "UNHCR Global Trends Forced Displacement in 2018." United Nations High Commission for Refugees, Geneva.

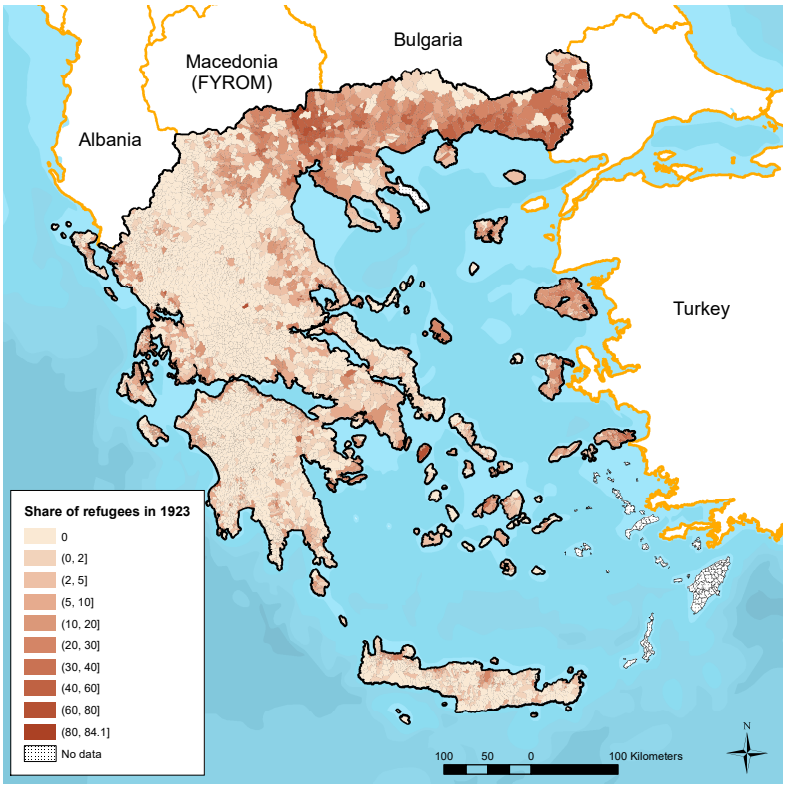
Vouras, Paul P. 1963. "The Significance of Cotton Growing in Greece." *The Professional Geographer*, 15(3): 11–13.

Figure 1: Distribution of refugees in 1923 and 1928 at the municipality level

(a) Share of refugees in 1928

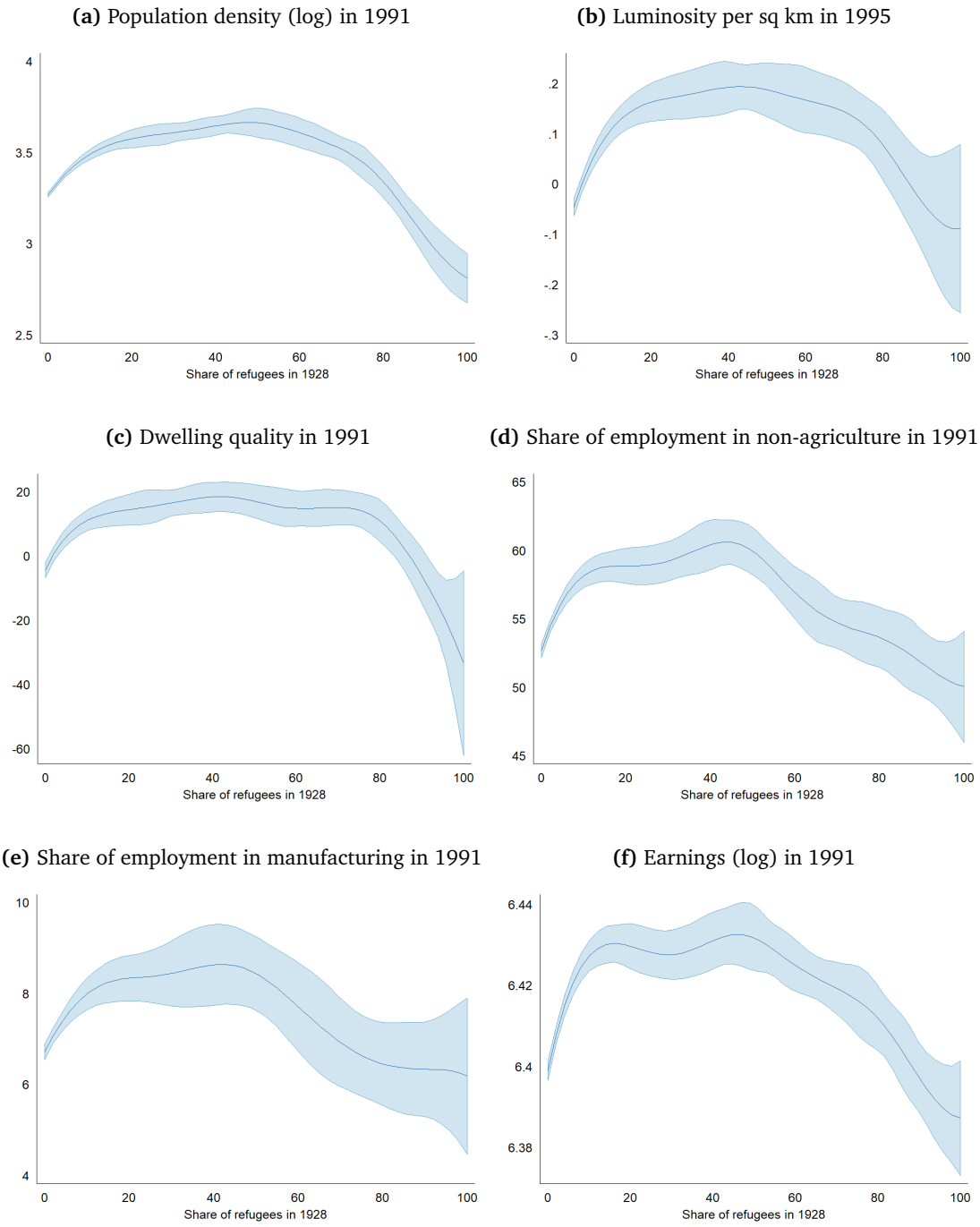


(b) Share of refugees in 1923



Notes. This figure visualizes the share of refugees in 1923 and 1928 at the municipality level. The maps approximate the municipal boundaries as defined by the 1920 population census of the Kingdom of Greece. Data sources: The 1923 refugee census and the 1928 population census of the Kingdom of Greece.

Figure 2: Semiparametric effects of refugees on contemporary development levels



Notes. This figure presents the kernel-weighted local polynomial fits showing the relationship between the share of refugees in 1928 and development levels in 1991/1995 at the municipality level, controlling for province fixed effects; log population in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude; mean annual precipitation and temperature; and crop suitability. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. We perform these estimations using the `semipar` command provided by the statistical software Stata 15, which is based on Robinson (1988)'s double residual estimator. Standard errors are clustered at the province level and 90% confidence interval are reported.

Table 1: Summary statistics at the municipality level

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Distribution of refugees					
Share of refugees in 1928 (%)	7.0	18.9	0	100	4388
Share of refugees in 1923 (%)	4.2	9.7	0	84.1	4388
Panel B: Contemporary outcomes					
Population density (log) in 1991	3.3	1	-1.5	8.2	4388
Luminosity per sq km in 1995 (Z-score)	0	1	-1.8	2.3	4388
Dwelling quality in 1991 (Z-score)	0	1	-7.6	1.3	4388
- Electricity in 1991 (%)	93.1	8.9	0	100	4388
- Sewage inside in 1991 (%)	94.4	9.5	0	100	4388
- Bathing facilities in 1991 (%)	56.4	21.5	0	100	4388
- Tap water inside in 1991 (%)	82.4	17.7	0	100	4388
Share of employment in non-agriculture in 1991 (%)	54	22.7	0	100	4388
Share of employment in manufacturing in 1991 (%)	7	8.3	0	100	4388
Share of employment in textile sector in 1991 (%)	2.1	5.2	0	81.8	4388
Earnings (log) in 1991	6.4	0.1	6.1	6.8	4388
Panel C: Historical and geographical controls					
Log population density in 1920	3.5	0.9	-1.9	8.6	4388
Dummy: city in 1920	0.012	0.1	0	1	4388
Dummy: island	0.214	0.4	0	1	4388
Distance to railway network in 1920 (km)	65.3	85	0.4	353.4	4388
Distance to shoreline (km)	25.6	27.8	0	130.6	4388
Distance to border (km)	203.3	142.8	0.9	565.6	4388
Altitude (m)	460.5	353.6	-0.7	1731.4	4388
Mean annual precipitation (mm)	696.5	178.5	96.6	1217.4	4388
Mean annual temperature (C°)	13.8	2.3	3.1	18	4388
Suitability of crops (Z-score)	0	1	-1.6	4.1	4388
- Suitability index for cotton	8.4	7.8	0	34	4388
- Suitability index for barley	21.3	13.3	1	86	4388
- Suitability index for oat	21.9	13.9	0.1	88	4388
- Suitability index for wheat	21.7	13.6	1	88	4388
- Suitability index for potato	13.3	6.9	0	47	4388
- Suitability index for olive	17.9	12.2	0	64	4388

Notes. This table presents the summary statistics for the distribution of refugees in 1923 and 1928, the main outcomes of interest in 1991/1995, and the historical and geographical controls used in the analysis. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Data sources: The 1923 refugee census and the 1928 population census of the Kingdom of Greece for the distribution of refugees in 1923 and 1928; Greek Statistical Office for contemporary outcomes; the 1920 census of the Kingdom of Greece for 1920 population; GAEZ v3.0 for altitude, precipitation, temperature, and suitability of crops. Distance measures are authors' own calculations.

Table 2: Refugee resettlement and long-term development

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991
	(1)	(2)	(3)
Share of refugees in 1928	2.217***	1.246***	1.284***
	(0.602)	(0.319)	(0.343)
Share of refugees in 1928, squared	-0.027***	-0.014***	-0.015***
	(0.007)	(0.005)	(0.004)
N	4,388	4,388	4,388
R2	.755	.638	.329
Province FE	Y	Y	Y
Controls	Y	Y	Y
Share of refugees maximizing outcome	40.7	44.8	43.0
Linear effect when marginal effect is positive	1.356	0.783	0.742
Oster's delta	1.7	1.2	1.7
Average outcome	332.6	0	0

Notes. This table presents the effects of refugees in 1928 on development outcomes in 1991/1995. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees in 1928 varies between 0 and 100. The outcome variable in column 1 is expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. To compute Oster's delta, we focus on the linear effects of refugees on a sample of municipalities where the marginal effect is positive in the quadratic specification. We report the estimated linear effects on this sample. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Refugee resettlement and structural transformation

Outcome:	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)
Share of refugees in 1928	0.377**	0.093***	0.179***
	(0.154)	(0.020)	(0.057)
Share of refugees in 1928, squared	-0.005***	-0.001***	-0.002***
	(0.002)	(0.000)	(0.001)
N	4,388	4,388	4,388
R2	.422	.406	.341
Province FE	Y	Y	Y
Controls	Y	Y	Y
Share of refugees maximizing outcome	41.6	40.8	42.6
Linear effect when marginal effect is positive	0.271	0.077	0.109
Oster's delta	46.4	10.5	6.2
Average outcome	53.9	7.0	640.5

Notes. This table presents the effects of refugees in 1928 on labor market outcomes in 1991. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees in 1928 varies between 0 and 100. The outcome variables in columns 1 and 2 are expressed in shares, varying between 0 and 100. The outcome variable in column 3 is expressed in logs ($\times$ 100). All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. To compute Oster's delta, we focus on the linear effects of refugees on a sample of municipalities where the marginal effect is positive in the quadratic specification. We report the estimated linear effects on this sample. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Time-varying effects of refugees on structural transformation: Province-level panel-data specification

Outcome:	Share of employment in non-agriculture		
	(1)	(2)	(3)
Share of refugees in 1928 $\times$ 𝟙{Year = 1928}	0.137***	0.153***	0.129**
	(0.035)	(0.048)	(0.050)
Share of refugees in 1928 $\times$ 𝟙{Year = 1951}	0.125***	0.120**	0.121*
	(0.042)	(0.050)	(0.068)
Share of refugees in 1928 $\times$ 𝟙{Year = 1971}	0.273***	0.363***	0.394***
	(0.064)	(0.079)	(0.082)
Share of refugees in 1928 $\times$ 𝟙{Year = 1991}	0.392***	0.439***	0.483***
	(0.088)	(0.106)	(0.105)
N	871	871	871
R2	0.96	0.96	0.97
Year FE	Y	Y	Y
Province FE	Y	Y	Y
Controls $\times$ Year FE	Y	Y	Y
Region $\times$ Year FE		Y	Y
Department $\times$ Year FE			Y
Average outcome	44.7	44.7	44.7
P-value for equality of coefficients between *Share of refugees* $\times$ 𝟙{*Year = 1928*} and:			
Share of refugees $\times$ 𝟙{Year = 1951}	.743	.406	.856
Share of refugees $\times$ 𝟙{Year = 1971}	.025	.004	0.001
Share of refugees $\times$ 𝟙{Year = 1991}	.003	.007	0.001

Notes. This table presents the time-varying effects of refugees in 1928 on the employment share in non-agriculture between 1920 and 1991. The unit of analysis is the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). There are 176 units of observation that are stable across census years; 1920 serves as the baseline year. The share of refugees varies between 0 and 100. The outcome variable is expressed in shares, varying between 0 and 100. All specifications control for province and census-year fixed effects and interactions of census-year dummies with a set of province-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the observation unit was a city in 1920 and whether it is an island; the distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the province. In addition, in columns 2 and 3, we include region $\times$ year fixed effects and department $\times$ year fixed effects. Standard errors clustered at the level of the unit of analysis are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Spillover effects of refugees on neighboring municipalities

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)*
Share of refugees in 1928	2.178***	0.982***	1.172***	0.346**	0.082***	0.163***
	(0.589)	(0.301)	(0.333)	(0.149)	(0.020)	(0.042)
Share of refugees in 1928, squared	-0.027***	-0.012**	-0.014***	-0.004***	-0.001***	-0.002***
	(0.006)	(0.005)	(0.004)	(0.001)	(0.000)	(0.000)
Share of refugees in neighbouring municipalities	0.120	1.418***	0.624**	0.113***	0.081**	0.060
	(0.282)	(0.491)	(0.309)	(0.029)	(0.040)	(0.038)
N	4,348	4,348	4,348	4,348	4,348	4,348
R2	.757	.659	.331	.429	.417	.346
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	40.8	42.1	41.6	41.1	39.0	41.9

Notes. This table presents the effects of refugees in neighboring municipalities in 1928 on development levels and labour market outcomes in 1991/1995. Neighbouring municipalities are defined as spatially contiguous municipalities of first and second degree (i.e. neighbors of neighbors). The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The sample excludes 50 municipalities that correspond to an entire island and thus that have no spatially contiguous neighbours. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 6, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Characteristics of refugees and natives in 1928 across regions

	Greece	Northern Greece	Central Greece and Islands
	(1)	(2)	(3)
Number of refugees	741,282	559,012	182,270
Population share of refugees	14.5	25.1	6.3
Literacy rate of			
Natives	54.2	50.2	56.6
Refugees	49.6	46.1	60.1
Employment share of non-agricultural occupation			
Natives	29.2	26.0	31.3
Refugees	29.8	21.9	67.9
Employment share of manufacturing			
Natives	13.6	13.6	13.6
Refugees	17.2	13.5	35.4
Employment share of skilled professionals			
Natives	3.3	2.7	3.6
Refugees	2.3	1.6	6.0
Employment share of banking and Insurance			
Natives	0.43	0.37	0.48
Refugees	0.37	0.18	1.32

Notes. This table presents the characteristics of natives and refugees in 1928 for all Greece, for Northern Greece, and for Central Greece and Islands. For each region, it tabulates the total number of refugees and the population share of refugees. Moreover, it tabulates the literacy rate and employment shares in different sectors separately among natives and refugees in each region. Data source: The 1928 population census of the Kingdom of Greece.

Table 7: Heterogenous effects of refugees across regions

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)*	(3)	(4)	(5)	(6)
Panel A: Northern Greece						
Share of refugees in 1928	1.420***	0.794***	0.672***	0.206*	0.090***	0.110***
	(0.433)	(0.252)	(0.222)	(0.120)	(0.017)	(0.031)
Share of refugees in 1928, squared	-0.019***	-0.009***	-0.008***	-0.003**	-0.001***	-0.001***
	(0.005)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)
N	1,650	1,650	1,650	1,650	1,650	1,650
R2	.709	.593	.322	.424	.387	.283
Share of refugees maximizing outcome	37.0	43.4	40.0	38.1	39.7	41.7
Average outcome	318.4	-6	13.6	53.6	9.2	641.2
Panel B: Central Greece and Islands						
Share of refugees in 1928	4.385***	2.675***	3.443***	1.122***	0.119*	0.564***
	(0.997)	(0.639)	(0.540)	(0.320)	(0.072)	(0.156)
Share of refugees in 1928, squared	-0.060***	-0.036***	-0.052***	-0.018***	-0.002*	-0.010***
	(0.023)	(0.014)	(0.017)	(0.006)	(0.001)	(0.003)
N	2,738	2,738	2,738	2,738	2,738	2,738
R2	.785	.675	.339	.444	.382	.391
Share of refugees maximizing outcome	36.2	37.1	33.3	30.8	25.1	28.2
Average outcome	341.2	3.6	-8.21	54.1	5.6	640.1
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y

Notes. This table presents the heterogeneous effects of refugees in 1928 on development levels and labour market outcomes in 1991/1995 by regions. Panel A and B present the estimated effects of refugees in Northern Greece and in Central Greece and Islands, respectively. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 2, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Mechanism: Transfer of technological knowledge in textile in Northern Greece

Outcome:	Share of textile sector in 1991		Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991
	(1)	(2)	(3)	(4)
Share of refugees in 1928	0.088***	0.076***	0.131	0.067***
	(0.012)	(0.015)	(0.114)	(0.016)
Share of refugees in 1928, squared	-0.001***	-0.001***	-0.002**	-0.001***
	(0.000)	(0.000)	(0.001)	(0.000)
High suitability for cotton	0.989	0.555	-4.237	0.149
	(0.706)	(0.791)	(3.619)	(0.944)
Share of refugees in 1928 × High suitability for cotton		0.020**	0.133***	0.036***
		(0.008)	(0.034)	(0.013)
N	1,650	1,650	1,650	1,650
R2	.505	.506	.429	.389
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Average outcome	3.8	3.8	53.6	9.2

Notes. This table presents the heterogeneous effects of refugees in 1928 on industrialization levels in 1991 by suitability of municipalities for cotton cultivation. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The analysis is restricted to municipalities in Northern Greece. "High suitability for cotton" is a dummy variable indicating whether a municipality is in the top tercile of the sample in terms of suitability for cotton cultivation. The share of refugees varies between 0 and 100. All outcome variables are expressed in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: Mechanism: Agricultural production diversity at the province level (1914–1929)

Outcomes:	Production cash crops (log)	Production tobacco (log)	Production cotton (log)	Diversity cash crops	Diversity cereals	Diversity leguminous staple crops
	(1)	(2)	(3)	(4)	(5)	(6)
Share of refugees in 1928 × $\mathbb{1}\{Year > 1923\}$	2.394***	13.096***	5.693**	0.417***	0.178***	0.527***
	(0.771)	(2.222)	(2.444)	(0.113)	(0.051)	(0.117)
N	126	126	126	121	126	126
R2	.911	.805	.541	.725	.808	.688
Province FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Average outcome	696.2	333.8	125.5	26.1	62.7	55.7

Notes. This table presents the effects of refugees in 1928 on the changes in the agricultural production levels and diversity of crop varieties produced between 1914 and 1929. The unit of analysis is provinces. The analysis is restricted to provinces that were enumerated in the 1914 agricultural census: Provinces in Macedonia, Western Thrace, Crete, and Aegean Islands regions. The share of refugees varies between 0 and 100. $\mathbb{1}\{Year > 1923\}$ is a dummy variable indicating whether production levels are observed after the arrival of refugees in 1923. The outcome variables in columns 1 to 3 are in logs (× 100). The outcome variables in columns 4 to 6 are in levels and vary between 0 and 100. All specifications include province fixed effects and year fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Mechanism: Spatial segregation between refugees and natives within municipalities in Northern Greece

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)*	(2)	(3)	(4)	(5)	(6)
Share of refugees in 1928	1.505***	1.064***	1.229***	0.370***	0.096*	0.146***
	(0.533)	(0.369)	(0.318)	(0.138)	(0.050)	(0.048)
Share of refugees in 1928, squared	-0.021***	-0.014***	-0.014***	-0.005***	-0.001*	-0.002***
	(0.005)	(0.004)	(0.003)	(0.001)	(0.001)	(0.000)
Dissimilary index (Z-score)	-5.146	-4.623***	-4.617***	-1.851**	-0.053	-0.779***
	(3.260)	(0.546)	(1.052)	(0.788)	(0.299)	(0.090)
N	555	555	555	555	555	555
R2	.749	.638	.476	.596	.549	.527
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Average outcome	339.2	28.8	40.1	50.9	11.3	641.5

Notes. This table presents the effects of spatial segregation between refugees and natives across settlements within a municipality in 1928 on development levels and labour market outcomes in 1991/1995. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The analysis is restricted to municipalities in Northern Greece with at least two settlements (or villages) and hosting some refugees in 1928. The spatial "Dissimilarity index" is standardized (z-score). The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects, the number of settlements within the municipality, and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 1, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: Mechanism: Agglomeration economies

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)*	(3)	(4)	(5)	(6)
Panel A: All municipalities						
Share of refugees in 1928	1.650***	0.873***	1.050***	0.290**	0.073***	0.143***
	(0.446)	(0.248)	(0.332)	(0.139)	(0.021)	(0.048)
Share of refugees in 1928, squared	-0.019***	-0.009**	-0.012***	-0.003**	-0.001***	-0.002***
	(0.005)	(0.004)	(0.003)	(0.001)	(0.000)	(0.001)
N	4,388	4,388	4,388	4,388	4,388	4,388
R2	.784	.655	.337	.441	.414	.368
F-stats for the joint-significance of population density controls	167.988	74.664	118.09	102.911	20.557	79.373
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Coefficients without 1928 population density in controls						
Share of refugees in 1928	2.217	1.246	1.284	.377	.093	.179
Share of refugees in 1928, squared	-.027	-.014	-.015	-.005	-.001	-.002
Panel B: Municipalities in Northern Greece						
Share of refugees in 1928	1.002***	0.478**	0.604**	0.160	0.072***	0.098***
	(0.236)	(0.233)	(0.263)	(0.116)	(0.027)	(0.029)
Share of refugees in 1928, squared	-0.012***	-0.004	-0.007***	-0.002*	-0.001***	-0.001***
	(0.003)	(0.003)	(0.002)	(0.001)	(0.000)	(0.000)
N	1,650	1,650	1,650	1,650	1,650	1,650
R2	.762	.623	.326	.447	.403	.305
F-stats for the joint-significance of population density controls	2185.913	318.926	15.316	38.806	30.071	51.526
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Coefficients without 1928 population density in controls						
Share of refugees in 1928	1.42	.794	.672	.206	.09	.11
Share of refugees in 1928, squared	-.019	-.009	-.008	-.003	-.001	-.001

Notes. This table presents the effects of refugees in 1928 on development levels and labour market outcomes in 1991/1995, controlling for the effects of agglomeration economies. Panel A presents the estimated effects of refugees across all municipalities, whereas panel B presents those across municipalities in Northern Greece. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The reported regressions control for ten dummy variables corresponding to the deciles of the population density distribution in 1928 to capture the effects of agglomeration economies. F-statistics showing the joint-significance of these dummy variables are reported. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects, the number of settlements within the municipality, and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 2, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For online publication

SUPPLEMENTARY APPENDIX

Mass Refugee Inflow and Long-run Prosperity: Lessons from the Greek Population Resettlement

Elie Murard* & Seyhun Orcan Sakalli†

December 22, 2019

*University of Alicante, IZA Institute of Labor Economics, LEAP, Murard@iza.org
†King's Business School, King's College London, seyhun.sakalli@kcl.ac.uk

Contents

A Cultural Similarities Between Refugees and Natives

At the turn of the twentieth century, ethnic Greeks were split between the Ottoman Empire (1299–1923) and the Kingdom of Greece, which secured its independence from the Ottoman Empire in 1832. They lived in territories that constitute today's Greece and in some parts of Turkey.[1] The Ottoman Empire acquired the territories inhabited by ethnic Greeks with the conquest of the Byzantine Empire, which ended with the Fall of Constantinople in 1453. Facing the problem of governing large non-Muslim populations, Ottomans established the *millet* (people, nation) system, which grouped non-Muslim communities according to their religious confession. Each monotheistic non-Muslim community, e.g., Greek Orthodox, Armenian Apostolic Christians, Roman Catholics, and Jews, had autonomy in the administration of their schools and regulation of certain branches of civil law, such as marriage, divorce, and inheritance (Masters, 2009). Consequently, ethnoreligious groups could preserve their identity while being subject to the Ottoman economic, social, and political system.

Greek Orthodox were under the jurisdiction of the Ecumenical Patriarch in Constantinople, and were classified by the Ottoman authorities as the *Rum* (Greek) millet, as the Ecumenical Patriarch was invariably of Greek origin. The Rum millet consisted not only of Greeks but also of other Orthodox Christian groups, such as Serbians, Vlachs, and Bulgarians. However, the Greeks were a distinct ethnocultural group within the larger Orthodox Christian millet. They distinguished themselves from other Orthodox Christians by retaining their Greek culture, i.e., customs, language, and traditions, which dated back to the Byzantine civilization and originated from the ancient Hellenic world. The Greek Orthodox Church played a key role in keeping alive a sense of Greek identity through the use of Greek in the liturgy and educational efforts during the long centuries of Ottoman suzerainty.

The Greek refugees were therefore similar in culture, religion, and ethnicity to the natives of the Greek mainland.[2] The majority of refugees also shared the language with Greek natives, although the refugees from the Black Sea (Pontics) could not communicate with local people very easily (Kontogiorgi, 2006, p. 165). There were also groups of refugees that had been linguistically Turkified during the Ottoman Empire, such as the Cappadocian Greeks, known as the Karamanlides. Once in Greece though, they adopted the modern Greek language. However, after the independence of the Kingdom of Greece, its constituents developed a national consciousness. In contrast, the ethnic Greeks from Asia Minor, Thrace and Pontus under Ottoman rule, tended to identify themselves in regional or religious terms, and most of them lacked any national consciousness.

B Data Appendix

B.1 Matching municipalities over time

Many municipalities in Greece were affected by the arrival of refugees. Some new localities were founded to settle the refugees; others were split into more than one settlement as they grew in terms of population due to the inflow of refugees. A smaller number of municipalities were merged to form a new one because of the outflow of population. Also, many municipalities changed name after the arrival of refugees as their population structure were affected to a great extent.

We matched the municipalities that existed in 1928 to both the municipal-level data in 1920 and in 1971–

[1] In Turkey, they were concentrated in eastern Thrace and around Constantinople, western Asia Minor (especially around Smyrna), central Asia Minor (Cappadocia), and northeastern Black Sea cost (Pontus) (see panel (a) of Figure A.2).

[2] Most of the refugees were settled in the former territories of the Ottoman Empire that were annexed by the Kingdom of Greece recently, such as Thessaly (1881), and Epirus, Macedonia, Crete, and Ionian islands (1913), and Western Thrace (1919).

1991. We used the following procedure to match the municipalities that existed in 1928 to those existed in 1920. First, we matched them by municipality and province names or by municipality and department names (considering only the municipalities with the same name but in different provinces that are adjacent to each other within a department). If such match does not exist, this means either of the following: (1) the name of the municipality changed between 1920 and 1928; (2) the municipality ceased to exist or was merged to another existing municipality between 1920 and 1928; (3) the municipality in 1928 was founded between 1920 and 1928. We relied on information provided by the database of Hellenic Agency for Local Development and Local Government (EETAA) on "Administrative Changes of Municipalities and Communities" to find which of these scenarios is true and to identify the correct match. EETAA provides detailed information on the date of foundation of each municipality, the changes in its borders and name if there was any, and information on whether the municipality was ever transferred from one province or department to another.

To match the municipalities that existed in 1928 to those existed in 1971–1991, we followed the same procedure. In cases when two communities with the same name or very similar names existed in a given province, we also made use of the Atlas of Greece showing the administrative boundaries of municipalities as of 1948 (Ministry of National Economy, 1950). In order to obtain a map showing the boundaries of municipalities as of 1920, we start with a shapefile showing the boundaries of municipalities in 1991.[3] We merge the 1991 boundaries of municipalities that were a part of the same municipality in 1920 to approximate the boundaries of municipalities in 1920. We then use these boundaries to compute the geophysical features of municipalities and to identify neighbouring municipalities in 1920.

B.2 Geographical features

We use data coming from the GAEZ v3.0 dataset (Fischer, 2008) to identify geographical characteristics of municipalities. We compute the average latitude, mean annual temperature, mean annual precipitation, and suitability indices for cultivation of cotton, barley, oat, wheat, potato, and olives at the level of 1920 municipality borders. We construct an overall crop suitability index as the first component of a Principal Component Analysis (PCA) of the suitability index for these crops. We consider crop suitability indices under rain-fed conditions with low-level inputs.

B.3 Data at the province level

The 1928 census provides information on literacy and occupation of the population. It also provides the same information for 1920 for a retrospective comparison. These data are reported by province and by city. For 137 provinces, occupational and literacy data are reported in a consistent way for both 1920 and 1928.[4] For 39 cities out of the 59 cities that existed in Greece in 1920, the same data are available for both 1920 and 1928.[5] These 39 cities are located in 37 different provinces, which we split into rural areas and urban areas (into one city and one rural area for 35 provinces and into two different cities and one rural area for 2 provinces). 85 provinces do not contain any cities in 1920, and are therefore considered as rural areas. 13 provinces include a city for which we do not have historical data for both 1920 and 1928. In that case,

[3]This shapefile can be accessed through the following link: http://www.geodata.gov.gr/en/dataset/oria-ota-pro-kapodistria. Last accessed on October 31, 2019.

[4]Data for 1920 are missing for three provinces: Sapes, Arnaia, and Mount Athos.

[5]Data for 1920 are missing for 17 cities: Kastoria, Livadeia, Megara, Polygyros, Didymoteicho, Lavreotiki, Kallithea, Plomari, Giannitsa, Filiatra, Salamina, Aegion, Agios Nikolaos, Soufli, Katerini, Argos, and Vrontados. The three biggest cities of Greece (Athens, Piraeus, and Thessaloniki) hosted half of the urban refugees. We exclude them from the sample as they may introduce endogeneity to the analysis given that they may have a different development trajectory compared to the rest of the country.

we aggregate the data at the province level and treat the entire province at the unit of observation.[6] In total, our final dataset includes 176 unit of observations: 124 rural areas, 39 urban areas (i.e. cities) and 13 aggregate provinces.

We complement these data with the 1951 Greek population census that provides information on occupational structure at the same level of aggregation, i.e., by province and by city. The 1951 census does not report any information for 7 cities and the 6 provinces that include these cities.[7] Consequently, 11 observations (out of 176) have missing occupational data for 1951 in our final panel dataset.

B.4 Data at the settlement level: Residential segregation of refugees and natives

Greek municipalities often comprise several settlements (or villages). We use the 1928 census data at the settlement level to examine the residential segregation between refugees and natives across settlements within municipalities in Northern Greece. We can compute an index of residential segregation between refugees and natives only for municipalities that received some refugees and composed of multiple settlements. 914 out of the 1650 municipalities in Northern Greece were composed of only one settlement in 1928. Among the 736 municipalities with at least two settlements, 10% received no refugees in 1928.[8] As a result we are left with a sample of 558 municipalities. On average these, these municipalities had 3.8 settlements and 30% of their population were refugees in 1928. In our empirical specification, we measure residential segregation between refugees and natives using the dissimilarity index proposed by Reardon and Firebaugh (2002):

$$Dissimilarity = \frac{1}{2p(1-p)} \sum_{j=1}^{J} \frac{t_j}{T} |p_j - p|$$

where p_j is the share of refugees in settlement j's population, p is the share of refugees in the entire municipality, $\frac{t_j}{T}$ is the proportion of settlement j's population in the total municipal population T, and J is the number of different settlements. Conceptually, this dissimilarity index can be interpreted as the fraction of the refugee population that should move from one settlement to another in order to achieve an even distribution in the municipality, divided by the proportion of refugee population that should move if the municipality were perfectly segregated. The index varies between 0 and 1, where zero corresponds to perfect integration and 1 to perfect segregation.[9] In our sample, the index varies between 0.002 and 1, with an average of 0.43 and a standard deviation of 0.26. Figure A.17 shows that there is no statistical relationship between the degree of refugees' segregation (as measured by the dissimilarity index) and the share of refugees in the municipality. The correlation coefficient between the two variables is -0.04.

B.5 Agricultural censuses of 1914 and 1929

We collect province-level data from the agricultural censuses of 1914 and 1929. The 1914 agricultural census covers only Crete and the territories annexed by Greece in 1913, namely Macedonia, Western Thrace, Crete, and Aegean Islands. The data are available for 63 different provinces in 1914, for a total of 126 province-year observation in the panel dataset. We examine the effects of refugees on the total production

[6] This is the case for the provinces of Aigialeia, Argos, Chalkidiki, Didymoteicho, Kastoria, Livadeia, Megaris, Mirabello, Pieria, Plomari, Soufli, Trifylia, and Giannitsa. The province of Chios comprises the city of Chios and the small town of Vrontados. Since we have no data for Vrontados, we split the province of Chios into one urban area (the city of Chios) and one rural area that also includes the small town of Vrontados.

[7] These cities are Athens, Piraeus, Argostoli, Sparta, Nafplio, Vathy, and Trikala.

[8] Among these 736 municipalities, none has a population of 100% refugees.

[9] The dissimilarity index is highly correlated with other measures of segregation, and in particular with the index used by Alesina and Zhuravskaya (2011) for which we obtain a correlation coefficient of 0.73.

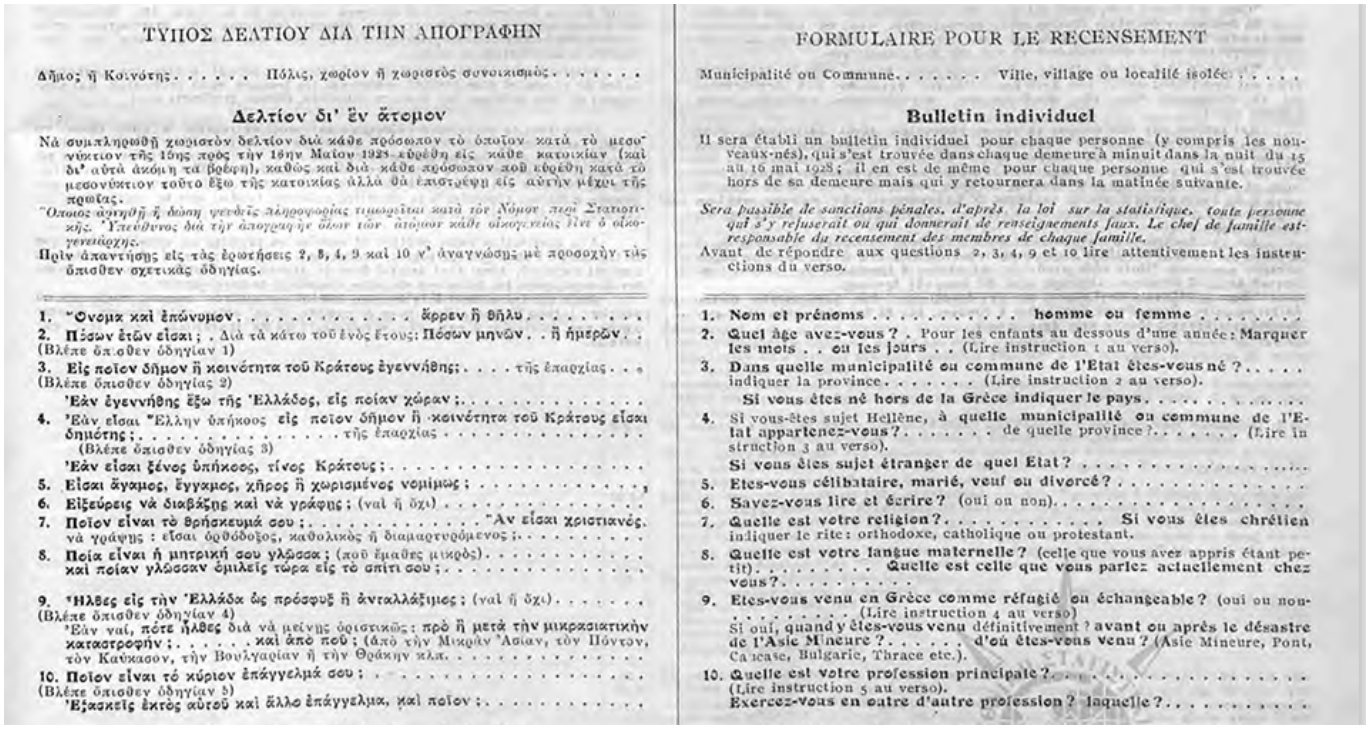
of cash crops, namely tobacco and cotton, measured in quintal (100 kg). We also examine those on the diversity in the cash crop, cereal, and leguminous staple crop varieties that were produced. The cereals comprises the following 11 variety of crops, as reported in the 1914 and 1929 censuses: wheat, oat, barley, rye, millet, white millet, corn, maize, rice, meslin, and sorghum. The leguminous staple crops comprises the following 8 variety of crops: pea, lentil, beans, chick peas, Indian pea, lupine, vetch, and bitter vetch. The cash crops comprises the following 12 variety of crops: tobacco, cotton, linen, hemp, anise, Arabic pistachio, sesame, crocus (saffron), sunflower, opium, Cayenne pepper, and mastic. For a given type of crop, we compute agricultural diversity as 1 minus a Hirschman-Herfindahl index of the shares of each product in the gross agricultural output of the crop type. For example, we compute the diversity of cereals as:

$$Diversity_{cereals} = 1 - \sum_{i=1}^{11} \left(\frac{prod_i}{\sum_{k=1}^{11} prod_k} \right)^2$$

where $prod_i$ is the production level of crop i (for i = 1, 2, ..., 11) in the list of all 11 possible crop varieties among cereals.

Figure A.1: Description of the 1928 census: examples

(a) Questionnaire



ΤΥΠΟΣ ΔΕΛΤΙΟΥ ΔΙΑ ΤΗΝ ΑΠΟΓΡΑΦΗΝ

Δῆμος ἢ Κοινότης. . . . . Πόλις, χωρίον ἢ χωριστὸς συνοικισμός. . . . . .

Δελτίον δι' ἓν ἄτομον

Νὰ συμπληρωθῇ χωριστὸν δελτίον διὰ κάθε πρόσωπον τὸ ὁποῖον κατὰ τὸ μεσονύκτιον τῆς 15ης πρὸς τὴν 16ην Μαΐου 1928 εὑρέθη εἰς κάθε κατοικίαν (καὶ δι' αὐτὰ ἀκόμη τὰ βρέφη), καθὼς καὶ διὰ κάθε πρόσωπον τοῦ εὑρέθη κατὰ τὸ μεσονύκτιον τοῦτο ἔξω τῆς κατοικίας ἀλλὰ θὰ ἐπιστρέψῃ εἰς αὐτὴν μέχρι τῆς πρωΐας.

Ὁποιος ἀρνηθῇ ἢ δώσῃ ψευδεῖς πληροφορίας τιμωρεῖται κατὰ τὸν Νόμον περὶ Στατιστικῆς. Ὑπεύθυνος διὰ τὴν ἀπογραφὴν ὅλων τῶν ἀτόμων κάθε οἰκογενείας εἶναι ὁ οἰκογενειάρχης.

Πρὶν ἀπαντήσῃς εἰς τὰς ἐρωτήσεις 2, 3, 4, 9 καὶ 10 ν' ἀναγνώσῃς μὲ προσοχὴν τὰς ὀπίσθεν σχετικὰς ὁδηγίας.

1. Ὄνομα καὶ ἐπώνυμον. . . . . ἄρρεν ἢ θῆλυ. . . . . .
2. Πόσων ἐτῶν εἶσαι; . Διὰ τὰ κάτω τοῦ ἑνὸς ἔτους: Πόσων μηνῶν. . ἢ ἡμερῶν. . (Βλέπε ὄπισθεν ὁδηγίαν 1)
3. Εἰς ποῖον δῆμον ἢ κοινότητα τοῦ Κράτους ἐγεννήθης; . . . . τῆς ἐπαρχίας. . . (Βλέπε ὄπισθεν ὁδηγίας 2)
   Ἐὰν ἐγεννήθης ἔξω τῆς Ἑλλάδος, εἰς ποίαν χώραν; . . . . . .
4. Ἐὰν εἶσαι Ἕλλην ὑπήκοος εἰς ποῖον δῆμον ἢ κοινότητα τοῦ Κράτους εἶσαι δημότης; . . . . . τῆς ἐπαρχίας . . . . . (Βλέπε ὄπισθεν ὁδηγίας 3)
   Ἐὰν εἶσαι ξένος ὑπήκοος, τίνος Κράτους; . . . . . .
5. Εἶσαι ἄγαμος, ἔγγαμος, χήρος ἢ χωρισμένος νομίμως; . . . . . ,
6. Εἰξεύρεις νὰ διαβάζῃς καὶ νὰ γράφῃς; (ναὶ ἢ ὄχι) . . . . . .
7. Ποῖον εἶναι τὸ θρήσκευμά σου; . . . . . Ἂν εἶσαι χριστιανός, νὰ γράψῃς: εἶσαι ὀρθόδοξος, καθολικὸς ἢ διαμαρτυρόμενος; . . . . . .
8. Ποία εἶναι ἡ μητρική σου γλῶσσα; (ποῦ ἔμαθες μικρός). . . . . καὶ ποίαν γλῶσσαν ὁμιλεῖς τώρα εἰς τὸ σπίτι σου; . . . . . .
9. Ἦλθες εἰς τὴν Ἑλλάδα ὡς πρόσφυξ ἢ ἀνταλλάξιμος; (ναὶ ἢ ὄχι). . . . . (Βλέπε ὄπισθεν ὁδηγίαν 4)
   Ἐὰν ναί, πότε ἦλθες διὰ νὰ μείνῃς ὁριστικῶς: πρὸ ἢ μετὰ τὴν μικρασιατικὴν καταστροφήν; . . . . . καὶ ἀπὸ ποῦ; (ἀπὸ τὴν Μικρὰν Ἀσίαν, τὸν Πόντον, τὸν Καύκασον, τὴν Βουλγαρίαν ἢ τὴν Θράκην κλπ. . . . . .
10. Ποῖον εἶναι τὸ κύριον ἐπάγγελμά σου; . . . . . (Βλέπε ὄπισθεν ὁδηγίαν 5)
    Ἐξασκεῖς ἐκτὸς αὐτοῦ καὶ ἄλλο ἐπάγγελμα, καὶ ποῖον; . . . . . .

FORMULAIRE POUR LE RECENSEMENT

Municipalité ou Commune. . . . . Ville, village ou localité isolée. . . . . .

Bulletin individuel

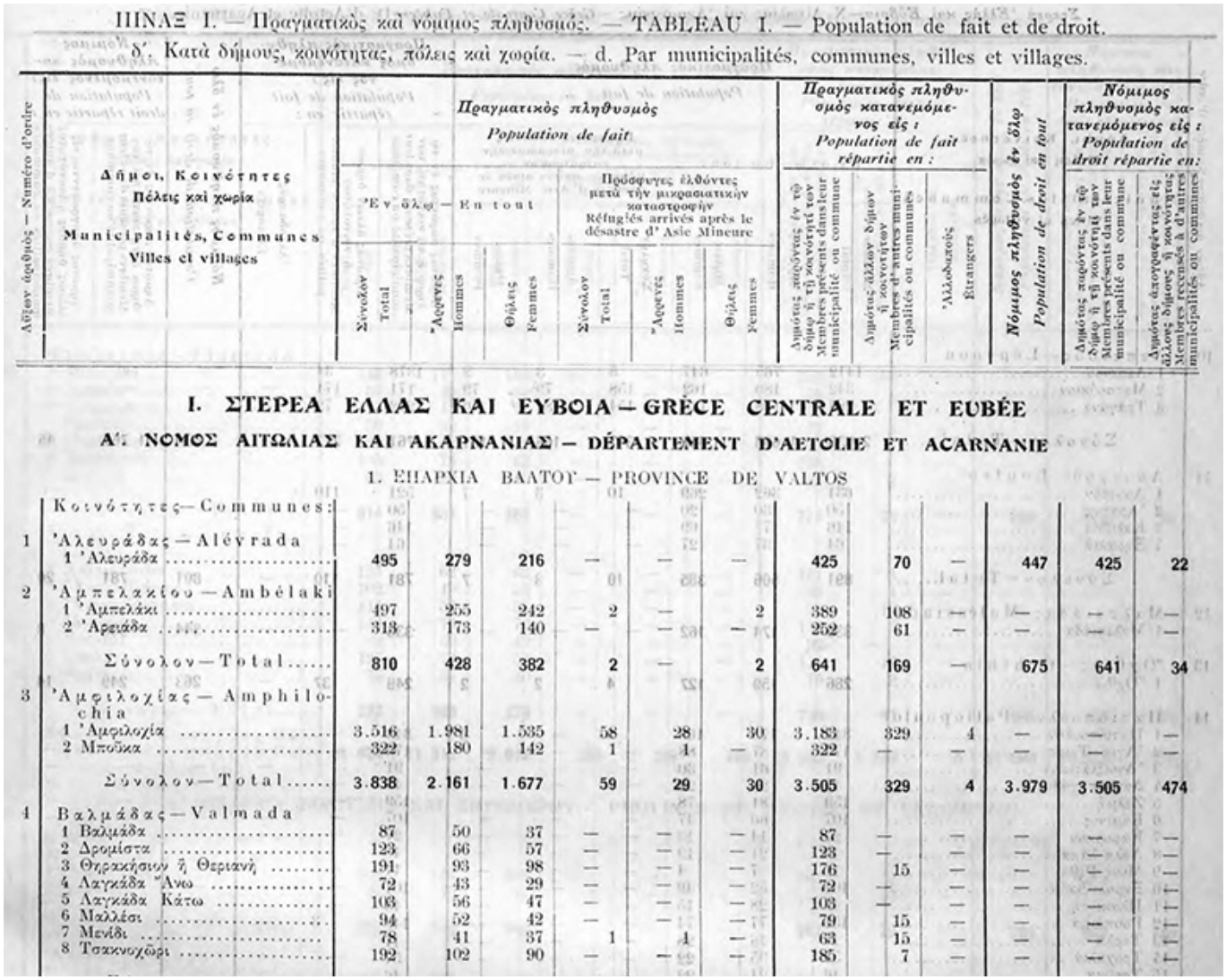
Il sera établi un bulletin individuel pour chaque personne (y compris les nouveaux-nés), qui s'est trouvée dans chaque demeure à minuit dans la nuit du 15 au 16 mai 1928; il en est de même pour chaque personne qui s'est trouvée hors de sa demeure mais qui y retournera dans la matinée suivante.

Sera passible de sanctions pénales, d'après la loi sur la statistique, toute personne qui s'y refuserait ou qui donnerait de renseignements faux. Le chef de famille est responsable du recensement des membres de chaque famille.

Avant de répondre aux questions 2, 3, 4, 9 et 10 lire attentivement les instructions du verso.

1. Nom et prénoms . . . . . homme ou femme . . . . . .
2. Quel âge avez-vous ? . Pour les enfants au dessous d'une année: Marquer les mois . . ou les jours . . (Lire instruction 1 au verso).
3. Dans quelle municipalité ou commune de l'Etat êtes-vous né ? . . . . . indiquer la province. . . . . (Lire instruction 2 au verso).
   Si vous êtes né hors de la Grèce indiquer le pays. . . . . .
4. Si vous-êtes sujet Hellène, à quelle municipalité ou commune de l'Etat appartenez-vous ? . . . . . de quelle province ? . . . . . (Lire instruction 3 au verso).
   Si vous êtes sujet étranger de quel Etat ? . . . . . .
5. Etes-vous célibataire, marié, veuf ou divorcé ? . . . . . .
6. Savez-vous lire et écrire ? (oui ou non). . . . . .
7. Quelle est votre religion ? . . . . . Si vous êtes chrétien indiquer le rite: orthodoxe, catholique ou protestant.
8. Quelle est votre langue maternelle ? (celle que vous avez appris étant petit). . . . . Quelle est celle que vous parlez actuellement chez vous ? . . . . . .
9. Etes-vous venu en Grèce comme réfugié ou échangeable ? (oui ou non). . . . . (Lire instruction 4 au verso)
   Si oui, quand êtes-vous venu définitivement ? avant ou après le désastre de l'Asie Mineure ? . . . . . d'où êtes-vous venu ? (Asie Mineure, Pont, Caucase, Bulgarie, Thrace etc.).
10. Quelle est votre profession principale ? . . . . . (Lire instruction 5 au verso).
    Exercez-vous en outre d'autre profession ? laquelle ? . . . . . .

(b) Municipal-level data



ΠΙΝΑΞ Ι. — Πραγματικὸς καὶ νόμιμος πληθυσμός. — TABLEAU I. — Population de fait et de droit.

δ'. Κατὰ δήμους, κοινότητας, πόλεις καὶ χωρία. — d. Par municipalités, communes, villes et villages.

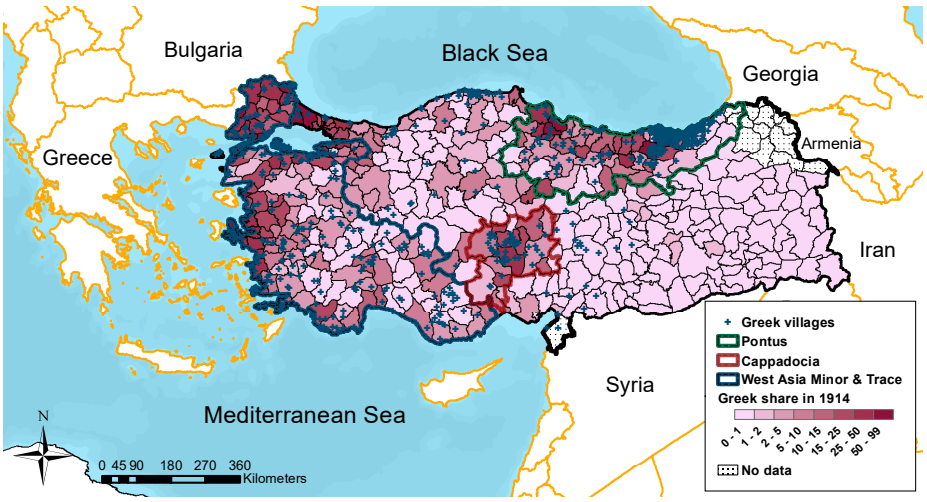
Ἀριθμὸς ἀριθμός — Numéro d'ordre	Δῆμοι, Κοινότητες, Πόλεις καὶ χωρία — Municipalités, Communes, Villes et villages	Πραγματικὸς πληθυσμός — Population de fait: Ἐν ὅλῳ — En tout: Σύνολον Total	Ἄρρενες Hommes	Θήλεις Femmes	Πρόσφυγες ἐλθόντες μετὰ τὴν μικρασιατικὴν καταστροφήν — Réfugiés arrivés après le désastre d'Asie Mineure: Σύνολον Total	Ἄρρενες Hommes	Θήλεις Femmes	Πραγματικὸς πληθυσμὸς κατανεμόμενος εἰς: Population de fait répartie en: Δημόται παρόντες ἐν τῷ δήμῳ ἢ τῇ κοινότητι των — Membres présents dans leur municipalité ou commune	Δημόται ἄλλων δήμων ἢ κοινοτήτων — Membres d'autres municipalités ou communes	Ἀλλοδαποί — Etrangers	Νόμιμος πληθυσμὸς ἐν ὅλῳ — Population de droit en tout	Νόμιμος πληθυσμὸς κατανεμόμενος εἰς: Population de droit répartie en: Δημόται παρόντες ἐν τῷ δήμῳ ἢ τῇ κοινότητι των — Membres présents dans leur municipalité ou commune	Δημόται ἀπογεγραμμένοι εἰς ἄλλους δήμους ἢ κοινότητας — Membres inscrits à d'autres municipalités ou communes
	Ι. ΣΤΕΡΕΑ ΕΛΛΑΣ ΚΑΙ ΕΥΒΟΙΑ — GRÈCE CENTRALE ET EUBÉE													
	Α'. ΝΟΜΟΣ ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝΑΝΙΑΣ — DÉPARTEMENT D'AETOLIE ET ACARNANIE													
	1. ΕΠΑΡΧΙΑ ΒΑΛΤΟΥ — PROVINCE DE VALTOS													
	Κοινότητες — Communes:													
1	Ἀλευράδας — Alévrada													
	1 Ἀλευράδα	495	279	216	—	—	—	425	70	—	447	425	22	
2	Ἀμπελακίου — Ambélaki													
	1 Ἀμπελάκι	497	255	242	2	—	2	389	108	—	—	—	—	
	2 Ἀρειάδα	313	173	140	—	—	—	252	61	—	—	—	—	
	Σύνολον — Total	810	428	382	2	—	2	641	169	—	675	641	34	
3	Ἀμφιλοχίας — Amphilochia													
	1 Ἀμφιλοχία	3.516	1.981	1.535	58	28	30	3.183	329	4	—	—	—	
	2 Μπούκα	322	180	142	1	1	—	322	—	—	—	—	—	
	Σύνολον — Total	3.838	2.161	1.677	59	29	30	3.505	329	4	3.979	3.505	474	
4	Βαλμάδας — Valmada													
	1 Βαλμάδα	87	50	37	—	—	—	87	—	—	—	—	—	
	2 Δρομίστα	123	66	57	—	—	—	123	—	—	—	—	—	
	3 Θρακήσιον ἢ Θεριανή	191	93	98	—	—	—	176	15	—	—	—	—	
	4 Λαγκάδα Ἄνω	72	43	29	—	—	—	72	—	—	—	—	—	
	5 Λαγκάδα Κάτω	103	56	47	—	—	—	103	—	—	—	—	—	
	6 Μαλλέσι	94	52	42	—	—	—	79	15	—	—	—	—	
	7 Μενίδι	78	41	37	1	1	—	63	15	—	—	—	—	
	8 Τσακνοχώρι	192	102	90	—	—	—	185	7	—	—	—	—	

Notes. Data source: The 1928 population census of the Kingdom of Greece.

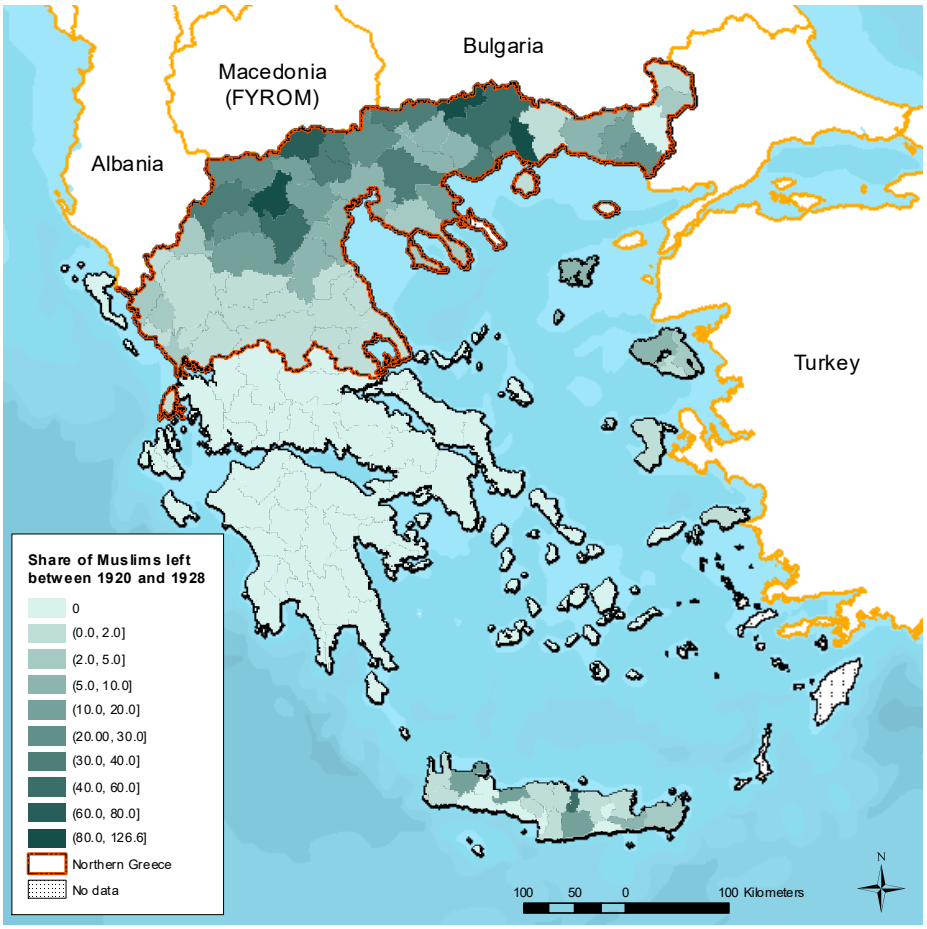
C Historical maps: Muslim population in Greece

Figure A.2: The population exchange of Greek-Orthodoxes and Muslims

(a) Greek-Orthodoxes in Ottoman Turkey in 1914

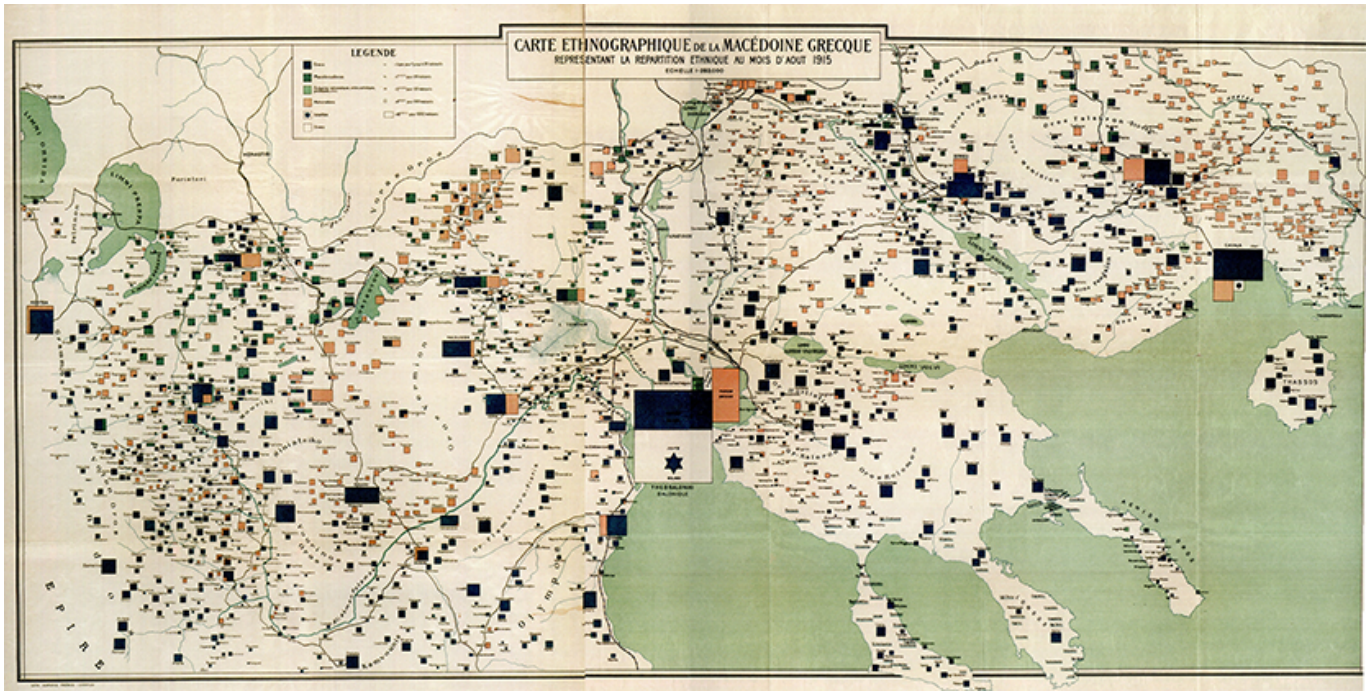


(b) Muslims in Greece in 1920



Notes. This figure presents the spatial distribution of Greek Orthodox living within the borders of the Republic of Turkey as of 1914 (panel (a)) and that of Muslims who left Greece between 1920 and 1928 (panel (b)) at the province level. The darker the purple color, the higher the number of Greek Orthodox relative to the 1914 population. The darker the green color, the higher the number of Muslims left between 1920 and 1928 relative to the 1920 population. The dark orange line in panel (b) represents the borders of the Northern Greece. Data sources: The 1914 population census of the Ottoman Empire for the number of Greek Orthodox; the 1929 population census of the Kingdom of Greece for the number of Muslim population left.

Figure A.3: Ethnographic map of Greek Macedonia: presenting the ethnic distribution in August 1915

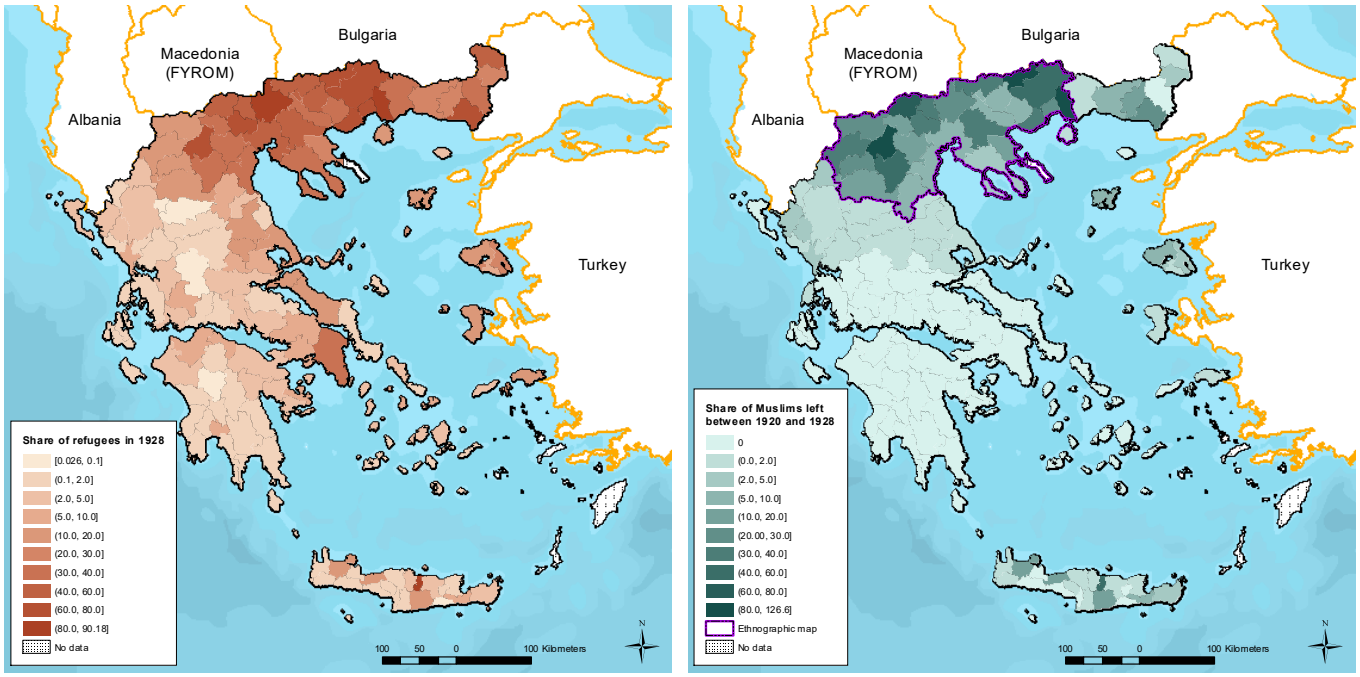


Notes. This figure presents the distribution of population by ethnoreligious groups at the settlement level in Macedonia in 1915. The title of the map reads: "Carte ethnographique de la Macedoine Grecque: presentant la repartition ethnique au mois d'aout 1915" [Ethnographic map of Greek Macedonia: presenting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

Figure A.4: Share of refugees and Muslims at the province level & Muslim localities in Macedonia in 1915

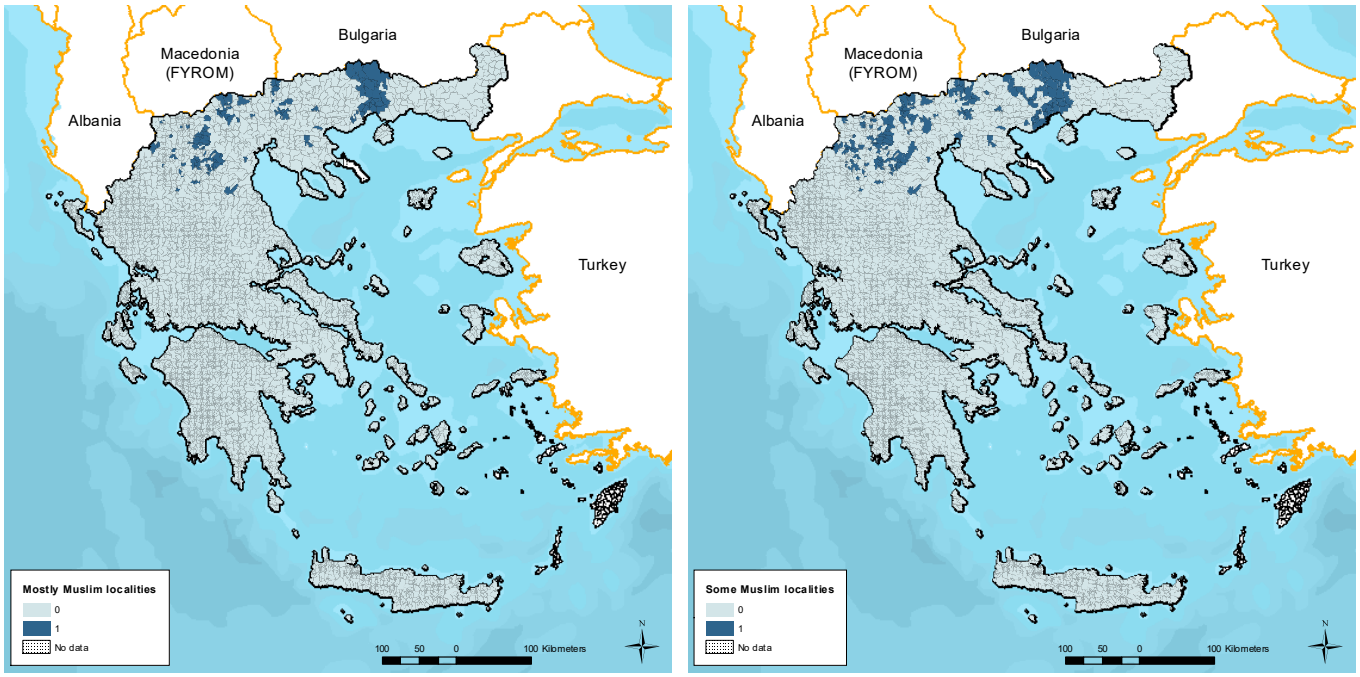
(a) Share of refugees in 1928 at the province level

(b) Share of Muslims left between 1920 and 1928 at the province level



(c) Mostly-Muslim localities in 1915

(d) Some-Muslim localities in 1915



Notes. This figures presents the spatial distribution of refugees in 1928 (panel (a)) and Muslims who left between 1920 and 1928 (panel (b)) at the province level and that of mostly-Muslim and some-Muslim localities in Macedonia in 1915 (panels (c) and (d)). The purple line in panel (b) represents the Macedonia region for which data on the distribution of population by ethnoreligious groups in 1914 are available. Data sources: The 1928 population census of the Kingdom of Greece for number of refugees and number of Muslims left; Güvenç (2010) for the map showing the distribution of ethnoreligious groups in Macedonia in 1915.

D Descriptive Statistics

Table A.1: Timeline of events

May 1919	Greek Army lands in Smyrna (Izmir)
August 27, 1922	Turkish Offence
September 9, 1922	Turkish Army enters Smyrna
September 13, 1922	Persecutions trigger mass exodus of Greek Orthodoxes
April 1923	Refugee census in Greece
July 1923	Lausanne Peace Treaty and population exchange agreement
November 1923	The Refugee Settlement Commission (RSC) starts its operations
May 1928	Greek population census
January 1930	Dissolution of the RSC

Table A.2: Summary statistics at the municipality level:
Outcomes in 1971 and characteristics of dwellings built after 1945

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Outcomes in 1971					
Population density (log) in 1971	3.4	1	-1.8	8.3	4376
Dwelling quality in 1971 (Z-score)	0	1	-2.1	4.2	4376
- Electricity in 1971 (%)	59	31.7	0	100	4376
- Sewage inside in 1971 (%)	69.1	23.1	0	100	4376
- Bathing facilities in 1971 (%)	5.1	8.9	0	79.3	4376
- Tap water inside in 1971 (%)	22.8	24.3	0	100	4376
Share of employment in non-agriculture in 1971 (%)	23.9	20.8	0	100	4376
Share of employment in manufacturing in 1971 (%)	4.5	6.5	0	54.5	4376
Panel B: Characteristics of buildings built after 1945					
- Electricity in 1991, built after 1945 (%)	94.3	8.8	0	100	4388
- Sewage inside in 1991, built after 1945 (%)	95.7	8.7	0	100	4388
- Bathing facilities in 1991, built after 1945 (%)	63.9	21.1	0	100	4388
- Tap water inside in 1991, built after 1945 (%)	85.5	16.9	0	100	4388

Notes. This table presents the summary statistics for the the main outcomes of interest in 1971 and the characteristics of dwellings built after 1945 as observed in 1991. Data source: Greek Statistical Office.

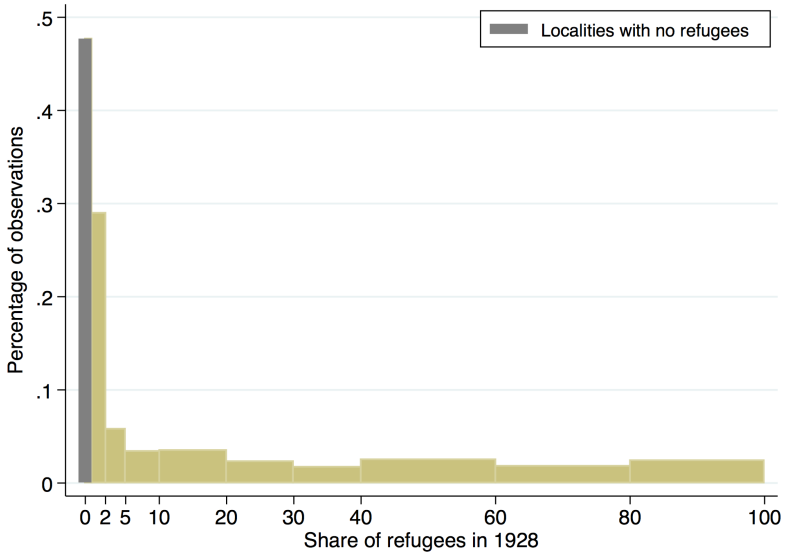
Table A.3: Urbanisation and industrialisation in Greece in the 20th century

Census year	Share of population in urban areas	Share of employment in agriculture
1920	22.9	53.5
1928	31.1	49.3
1940	32.8	
1951	37.7	51.3
1961	43.3	53.8
1971	53.2	41.2
1981	58.1	28.1
1991	58.8	18.6

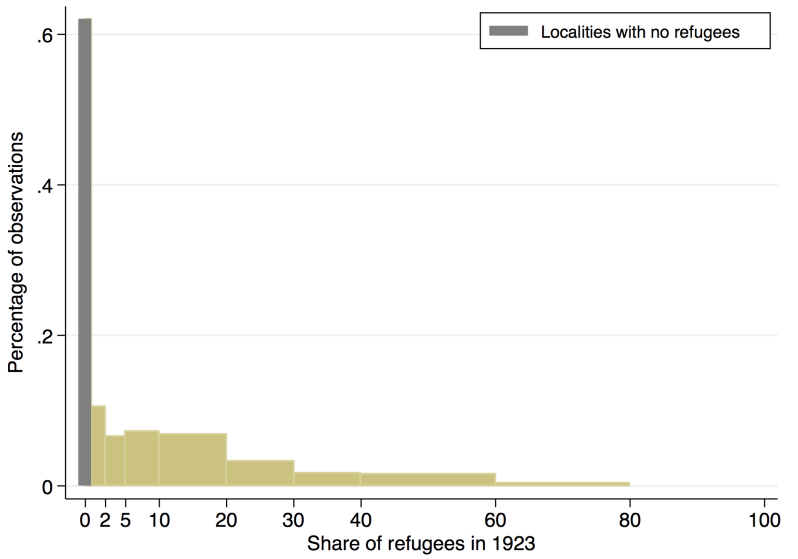
Notes. This table presents the urbanization rate and employment share in agriculture across census years between 1920 and 1991. Urban areas are defined as municipalities with more than 10,000 inhabitants. Data sources: Greek population censuses of 1928, 1940, 1951, 1961, 1971, 1981, and 1991.

Figure A.5: Histogram of the share of refugees at the municipality level

(a) Share of refugees in 1928



(b) Share of refugees in 1923



Notes. This figure presents the histogram of the share of refugees in 1928 (panel (a)) and in 1923 (panel (b)) at the municipality level. Data sources: The 1923 refugee census and the 1928 population census of the Kingdom of Greece.

E Determinants of resettlement location

Table A.4: Determinants of refugee settlement at the municipality level

Outcome:	Share of refugees in 1928	Share of refugees in 1923	Change btwn 1923–1928
	(1)	(2)	(3)
Log population density in 1920	0.212	0.730	-0.518
	(0.370)	(0.717)	(0.673)
Dummy: city in 1920	3.718*	10.476***	-6.757***
	(2.193)	(1.426)	(1.290)
Dummy: island	4.644***	24.736***	-20.092***
	(1.098)	(1.212)	(1.057)
Distance to railway network in 1920 (km)	-0.047	-0.012	-0.035
	(0.033)	(0.022)	(0.024)
Distance to shoreline (km)	-0.005	-0.082***	0.077**
	(0.049)	(0.020)	(0.039)
Distance to border (km)	-0.080	-0.009	-0.071
	(0.062)	(0.017)	(0.059)
Altitude (m)	-0.003	-0.003	-0.000
	(0.003)	(0.002)	(0.002)
Mean annual precipitation	-0.013	0.003	-0.016
	(0.012)	(0.007)	(0.012)
Mean annual temperature	0.211	-0.316*	0.528*
	(0.369)	(0.168)	(0.312)
Crop suitability score	0.996**	1.251**	-0.255
	(0.443)	(0.583)	(0.373)
N	4,388	4,388	4,388
R2	0.60	0.54	0.47
Province FE	Y	Y	Y
Average outcome	7.05	4.19	2.86

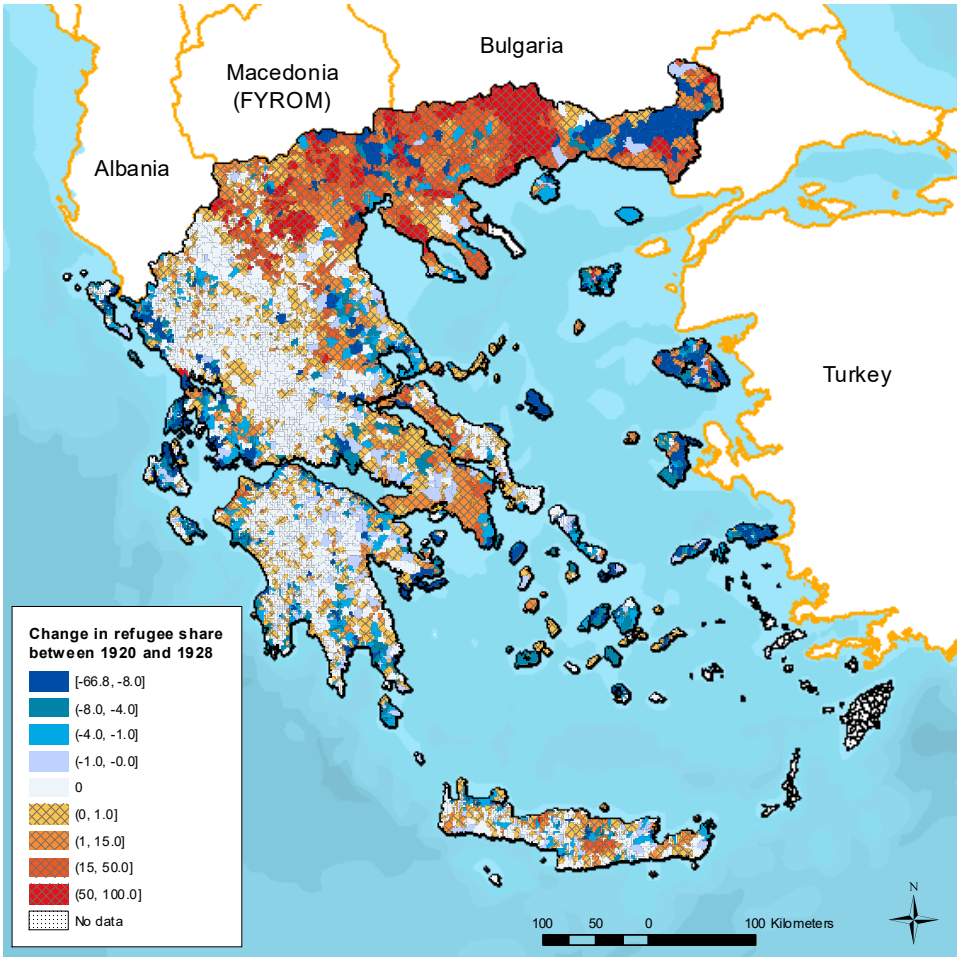
Notes. This table examines the determinants of the share of refugees in 1923 and 1928 and the change in the share of refugees between the two years across municipalities. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. All specifications controls for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, shoreline, and national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null betweeen municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Determinants of refugee settlement at the province level

Outcome:	Share of refugees in 1928					
	(1)	(2)	(3)	(4)	(5)	(6)
Share of employment in non-agriculture in 1920	-0.198**			-0.056		
	(0.091)			(0.048)		
Share of employment in manufacture in 1920		-0.468***			-0.152	
		(0.163)			(0.099)	
Literacy rate in 1920 (15-65 year olds)			-0.368**			0.063
			(0.153)			(0.076)
Share of Muslims left between 1920 and 1928				0.551***	0.542***	0.575***
				(0.067)	(0.068)	(0.066)
N	176	176	176	176	176	176
R2	0.73	0.74	0.74	0.88	0.88	0.88
Region FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Average of variable of interest	38.31	16.76	45.33	38.31	16.76	45.33
Average outcome	12.99	12.99	12.99	12.99	12.99	12.99

Notes. This table examines the determinants of the distribution of refugees in 1928 across provinces. The unit of analysis is the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). All specifications control for region fixed effects (for the following 10 regions of Greece: Crete, Peloponnese, Central Greece and Euboea, Cyclades Islands, Aegean Islands, Ionian Islands, Epirus, Thessaly, Macedonia and Western Thrace) and a set of province-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the observation unit was a city in 1920 and whether it is an island; the distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the province. Heteroskedasticity-robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

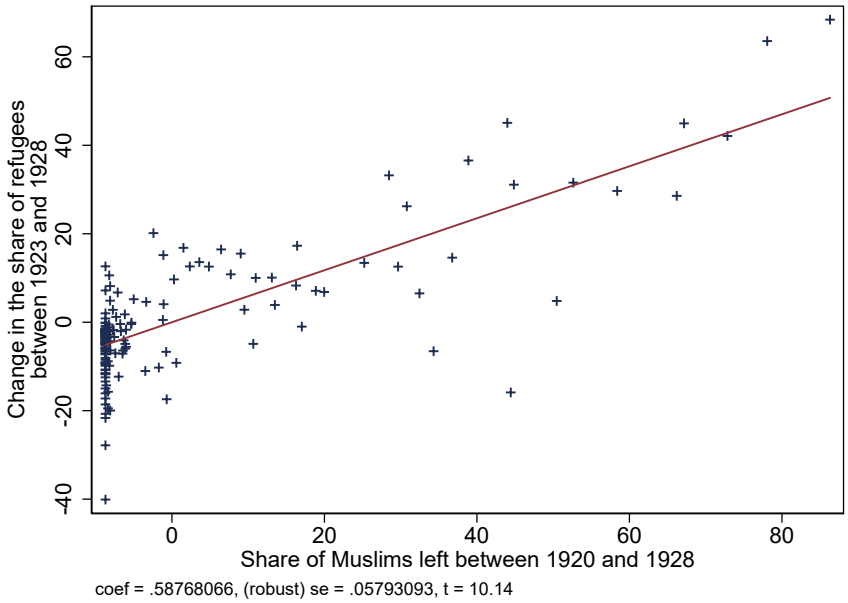
Figure A.6: Change in the share of refugees between 1923 and 1928



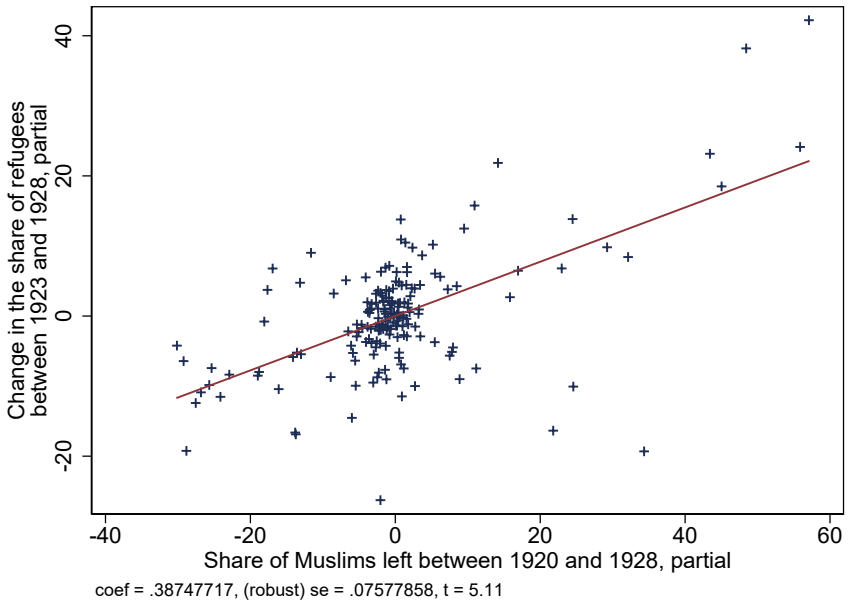
Notes. This figure visualizes the change in the share of refugees between 1923 and 1928 at the municipality level. The darker the color is, the greater the change in the share of refugees in a municipality is. Blue color represents a fall in the share of refugees between 1923 and 1928, whereas the red color represents an increase. Data sources: The 1923 refugee census and the 1928 populaton census of the Kingdom of Greece.

Figure A.7: Change in the share of refugees between 1923 and 1928 and the share of Muslims left between 1920 and 1928 at the province level

(a) Without controls



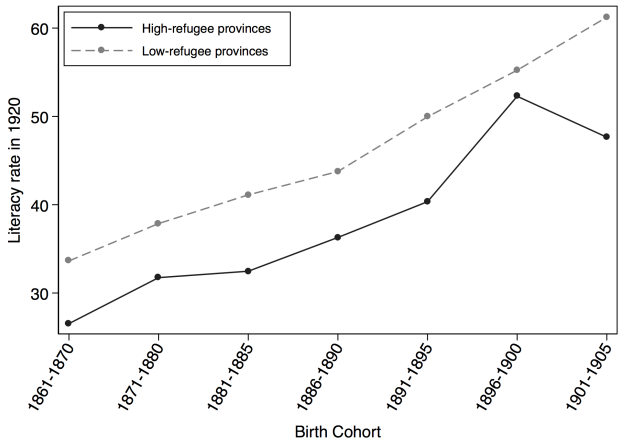
(b) With controls



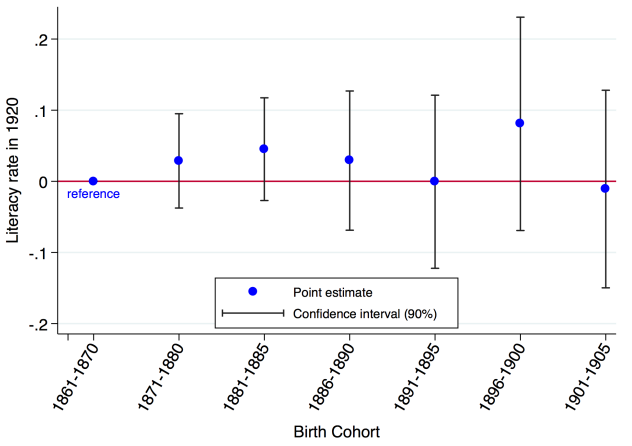
Notes. This figure presents the correlation between the change in the share of refugees between 1923 and 1928 and the share of Muslims left between 1920 and 1928 at the province level. The unit of analysis is the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). Panel (a) presents the raw correlation between the two shares, whereas panel (b) presents the correlation conditional on region fixed effects (for the following 10 regions of Greece: Crete, Peloponnese, Central Greece and Euboea, Cyclades Islands, Aegean Islands, Ionian Islands, Epirus, Thessaly, Macedonia and Western Thrace) and a set of province-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the observation unit was a city in 1920 and whether it is an island; the distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the province. Heteroskedasticity-robust standard errors are reported. Data sources: The 1923 refugee census and the 1928 populaton census of the Kingdom of Greece; see notes of Table 1 for the sources of province-level covariates.

Figure A.8: Pre-trends in literacy rates in 1920

(a) Trends in average literacy by province



(b) Test of Parallel Pre-trend



Notes. This figure presents the average differences in trends in literacy rates across birth cohorts observed in 1920, i.e., before the arrival of refugees, and across provinces that received varying levels of refugees. The figure in panel (a) plots the average literacy rate in 1920 by birth cohorts within high-refugee and low-refugee provinces (defined as above and below of the median of the share of refugees in 1928, respectively). The figure in (b) plots, for each birth cohort c, the point estimates β_c (along with 90% confidence intervals) obtained from the following regression: $Lit_{pc} = \sum_{c \neq c_0} \mathbb{1}_{\{cohort=c\}} \times (\beta_c \; REF_p + \alpha X'_p) + \delta_p + \delta_c + \varepsilon_{pc}$, where Lit_{pc} is the 1920 literacy rate of birth cohort c in province p, REF_p is the share of refugees in 1928, $\mathbb{1}_{\{cohort=c\}}$ is a dummy variable taking one if the birth cohort is equal to c and 0 otherwise, and X_p is a vector of geographic and population controls – the same as those included in Table A.5. δ_p and δ_c stand for province and birth cohort fixed-effects, respectively. The unit of analysis is the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). Standard errors clustered at the level of the unit of analysis. The reference cohort c_0 are individuals born between 1861 and 1870. The coefficient β_c measures whether the difference in literacy rate between cohort c and c_0 is significantly higher (or lower) in provinces that received more refugees in 1928 (relative to the average province). We cannot reject that the trends in literacy rates are the same across provinces with varying shares of refuges in 1928.

F Prosperity outcomes in 1971 and housing amenities

Table A.6: Development and labor-market outcomes in 1971

Outcome:	Population density (log) in 1971	Dwelling Quality in 1971	Share of employment in non-agriculture in 1971	Share of employment in manufacturing in 1971
	(1)	(2)	(3)	(4)
Share of refugees in 1928	1.956***	2.998***	0.571***	0.161***
	(0.393)	(0.538)	(0.199)	(0.035)
Share of refugees in 1928, squared	-0.024***	-0.032***	-0.007***	-0.002***
	(0.004)	(0.006)	(0.002)	(0.000)
N	4,376	4,376	4,376	4,376
R2	.759	.463	.357	.287
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Share of refugees maximizing outcome	40.1	47.4	41.1	41.8
Linear effect when marginal effect is positive	1.194	1.836	0.393	0.1
Oster's delta	1.6	3	6.2	4.9
Average outcome	335.8	0	23.8	4.5

Notes. This table presents the effects of refugees in 1928 on development and labor-market outcomes in 1971. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees in 1928 varies between 0 and 100. The outcome variable in column 1 is expressed in logs ($\times 100$). The outcome variable in columns 2 is standardized (z-scores). The outcome variables in columns 3 and 4 are expressed in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. To compute Oster's delta, we focus on the linear effects of refugees on a sample of municipalities where the marginal effect is positive in the quadratic specification. We report the estimated linear effects on this sample. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Refugee resettlement and household amenities in 1991

Outcome:	Electricity in 1991	Sewage inside in 1991	Bathing facilities in 1991	Tap water inside in 1991
	(1)	(2)	(3)	(4)
Share of refugees in 1928	0.090***	0.049***	0.409***	0.138**
	(0.025)	(0.016)	(0.108)	(0.056)
Share of refugees in 1928, squared	-0.001***	-0.001***	-0.004***	-0.002***
	(0.000)	(0.000)	(0.001)	(0.001)
N	4,388	4,388	4,388	4,388
R2	.248	.154	.414	.329
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Share of refugees maximizing outcome	36.1	43.8	47.6	43.1
Linear effect when marginal effect is positive	0.032	0.034	0.225	0.089
Oster's delta	2.6	2.1	2	1.2
Average outcome	93.1	94.4	56.3	82.3

Notes. This table presents the effects of refugees in 1928 on household amenities in 1991. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees in 1928 varies between 0 and 100. The outcome variables are expressed in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. To compute Oster's delta, we focus on the linear effects of refugees on a sample of municipalities where the marginal effect is positive in the quadratic specification. We report the estimated linear effects on this sample. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Refugee resettlement and household amenities in 1971

Outcome:	Electricity in 1971	Sewage inside in 1971	Bathing facilities in 1971	Tap water inside in 1971
	(1)	(2)	(3)	(4)
Share of refugees in 1928	0.242***	0.403***	0.299***	0.755***
	(0.083)	(0.098)	(0.072)	(0.086)
Share of refugees in 1928, squared	-0.003***	-0.004***	-0.003***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
N	4,376	4,376	4,376	4,376
R2	.391	.364	.385	.376
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Share of refugees maximizing outcome	38.8	45.1	47.2	51.7
Linear effect when marginal effect is positive	0.114	0.268	0.193	0.448
Oster's delta	.7	5.1	5.1	5
Average outcome	59.0	69.1	5.1	22.7

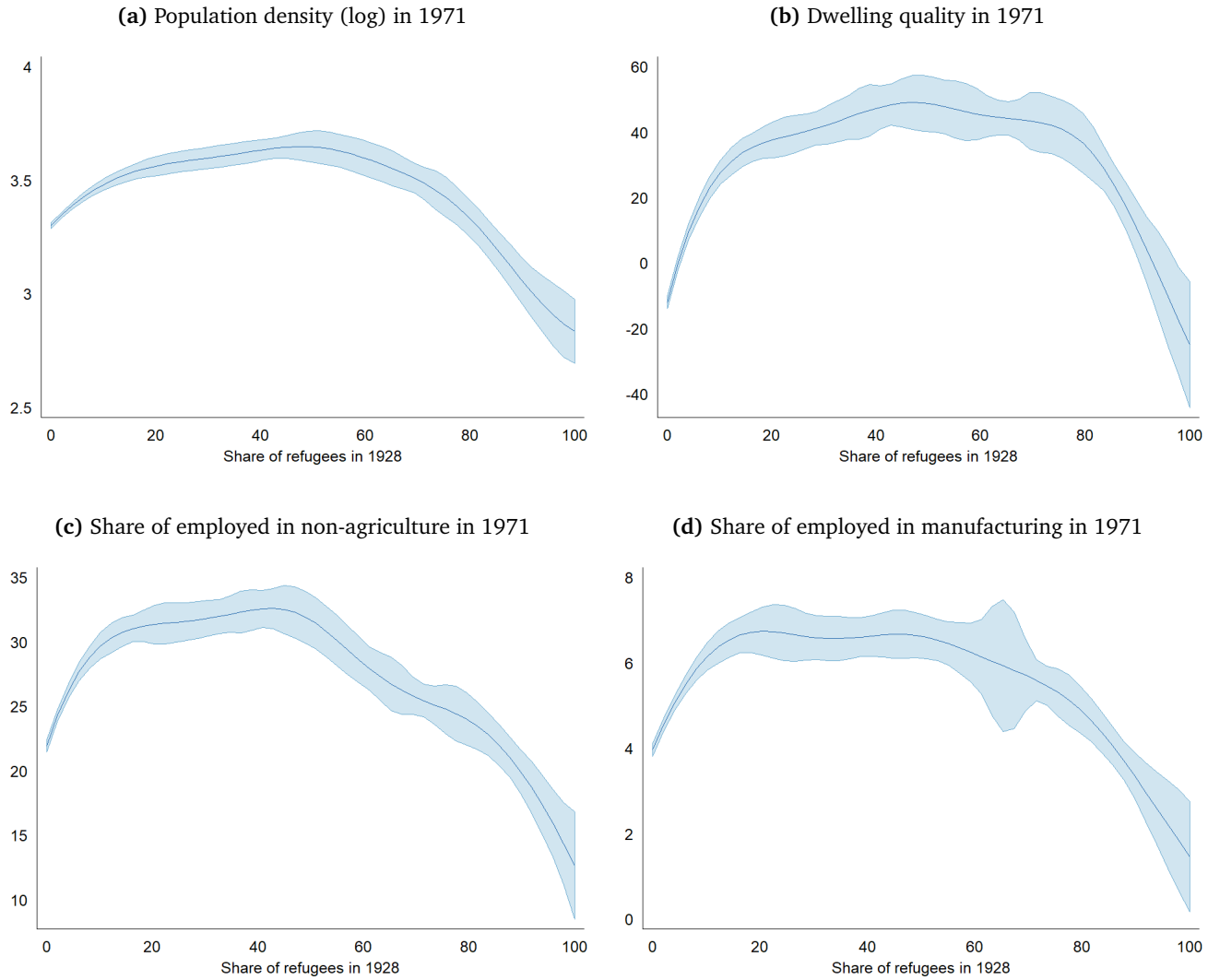
Notes. This table presents the effects of refugees in 1928 on household amenities in 1971. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees in 1928 varies between 0 and 100. The outcome variables are expressed in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. To compute Oster's delta, we focus on the linear effects of refugees on a sample of municipalities where the marginal effect is positive in the quadratic specification. We report the estimated linear effects on this sample. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Refugee resettlement and household amenities in 1991, among dwellings built after 1945

Outcome:	Electricity in 1991	Sewage inside in 1991	Bathing facilities in 1991	Tap water inside in 1991
	(1)	(2)	(3)	(4)
Share of refugees in 1928	0.084***	0.044***	0.408***	0.123**
	(0.027)	(0.016)	(0.098)	(0.056)
Share of refugees in 1928, squared	-0.001***	-0.000**	-0.004***	-0.001**
	(0.000)	(0.000)	(0.001)	(0.001)
N	4,388	4,388	4,388	4,388
R2	.193	.128	.364	.289
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Share of refugees maximizing outcome	35.6	46.4	47.3	4
Linear effect when marginal effect is positive	0.034	0.026	0.229	1.951
Oster's delta	2.8	1.8	2.6	5.2
Average outcome	94.2	95.7	63.9	85.5

Notes. This table presents the effects of refugees in 1928 on household amenities in 1991 for dwellings built after 1945. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees in 1928 varies between 0 and 100. The outcome variables are expressed in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. To compute Oster's delta, we focus on the linear effects of refugees on a sample of municipalities where the marginal effect is positive in the quadratic specification. We report the estimated linear effects on this sample. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.9: Semiparametric effects of refugees in 1928 on development levels in 1971 at the municipality level



Notes. This figure presents the kernel-weighted local polynomial fits showing the relationship between the share of refugees in 1928 and development levels in 1971 at the municipality level, controlling for province fixed effects; log population in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude; mean annual precipitation and temperature; and crop suitability. We perform these estimations using the `semipar` command provided by the statistical software Stata 15, which is based on Robinson (1988)'s double residual estimator. Standard errors are clustered at the province level and 90% confidence interval are reported.

G Robustness of main results

G.1 Province panel data estimates

Table A.10: Short-term effects of refugees on natives occupational choices of natives

A Outcome:	Employment share in non-agriculture among natives		
	(1)	(2)	(3)
Share of refugees in 1928 $\times$ $\mathbb{1}$\{Year = 1928\}	0.314***	0.361***	0.350***
	(0.082)	(0.106)	(0.087)
N	352	352	352
R2	0.99	0.99	0.99
Year FE	Y	Y	Y
Province FE	Y	Y	Y
Controls $\times$ Year FE	Y	Y	Y
Region $\times$ Year FE		Y	Y
Department $\times$ Year FE			Y
Average outcome	37.2	37.2	37.2

Notes. This table presents the effects of refugees in 1928 on the change in the employment share in non-agriculture among non-refugees between 1920 and 1928. The unit of analysis are the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). There are 176 units of observation that are stable across census years; 1920 serves as the baseline year. The share of refugees varies between 0 and 100. The outcome variable is expressed in shares, varying between 0 and 100. All specifications control for province and census-year fixed effects and interactions of census-year dummies with a set of province-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the observation unit was a city in 1920 and whether it is an island; the distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the province. In addition, in columns 2 and 3, we include region $\times$ year fixed effects and department $\times$ year fixed effects. Standard errors clustered at the level of the unit of analysis are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Robustness: Time-varying effects of refugees on structural transformation (Province-level panel-data specification)

Outcome:	Share of employment in non-agriculture		
	(1)	(2)	(3)
Share of refugees in 1928 × 𝟙{Year = 1928}	0.236***	0.206**	0.290***
	(0.082)	(0.086)	(0.094)
Share of refugees in 1928 × 𝟙{Year = 1951}	0.165**	0.150*	0.243**
	(0.083)	(0.083)	(0.121)
Share of refugees in 1928 × 𝟙{Year = 1971}	0.332***	0.391***	0.571***
	(0.119)	(0.127)	(0.163)
Share of refugees in 1928 × 𝟙{Year = 1991}	0.418***	0.458***	0.603***
	(0.129)	(0.147)	(0.155)
Share of Muslims left btwn 1920 and 1928 × 𝟙{Year = 1928}	-0.096	-0.052	-0.151**
	(0.061)	(0.057)	(0.064)
Share of Muslims left btwn 1920 and 1928 × 𝟙{Year = 1951}	-0.039	-0.030	-0.115
	(0.057)	(0.059)	(0.078)
Share of Muslims left btwn 1920 and 1928 × 𝟙{Year = 1971}	-0.058	-0.027	-0.165
	(0.090)	(0.085)	(0.106)
Share of Muslims left btwn 1920 and 1928 × 𝟙{Year = 1991}	-0.025	-0.019	-0.112
	(0.105)	(0.117)	(0.113)
N	871	871	871
R2	0.96	0.96	0.97
Year FE	Y	Y	Y
Province FE	Y	Y	Y
Controls × Year FE	Y	Y	Y
Region × Year FE		Y	Y
Department × Year FE			Y
Average outcome	44.7	44.7	44.7

Notes. This table presents the time-varying effects of refugees in 1928 on the employment share in non-agriculture between 1920 and 1991, controlling for the time-varying effects of the share of Muslims left between 1920 and 1928. The unit of analysis is the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). There are 176 units of observation that are stable across census years; 1920 serves as the baseline year. The share of refugees varies between 0 and 100. The outcome variable is expressed in shares, varying between 0 and 100. Moreover, all specifications control for province and census-year fixed effects and interactions of census-year dummies with a set of province-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the observation unit was a city in 1920 and whether it is an island; the distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the province. In addition, in columns 2 and 3, we include region × year fixed effects and department × year fixed effects. Standard errors clustered at the level of the unit of analysis are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.2 Spatial spillovers estimates

Table A.12: Spillover effects: quadratic specification

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)*
Share of refugees in 1928	2.166***	0.930***	1.141***	0.352**	0.077***	0.163***
	(0.575)	(0.277)	(0.338)	(0.148)	(0.018)	(0.041)
Share of refugees in 1928, squared	-0.027***	-0.011***	-0.014***	-0.004***	-0.001***	-0.002***
	(0.006)	(0.004)	(0.004)	(0.001)	(0.000)	(0.000)
Share of refugees among neighbors	0.482	2.920***	1.522**	-0.076	0.198***	0.047
	(0.522)	(1.090)	(0.680)	(0.086)	(0.057)	(0.074)
Share of refugees among neighbors, squared	-0.008	-0.031**	-0.019**	0.004***	-0.002***	0.000
	(0.011)	(0.015)	(0.009)	(0.001)	(0.000)	(0.001)
N	4,348	4,348	4,348	4,348	4,348	4,348
R2	.758	.661	.331	.429	.419	.346
Share of refugees maximizing outcome	40.9	42.6	41.8	41.0	39.2	41.9
Share of refugees among neighbors max. outcome	32.2	46.9	40.9	9.7	40.8	-84.1
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y

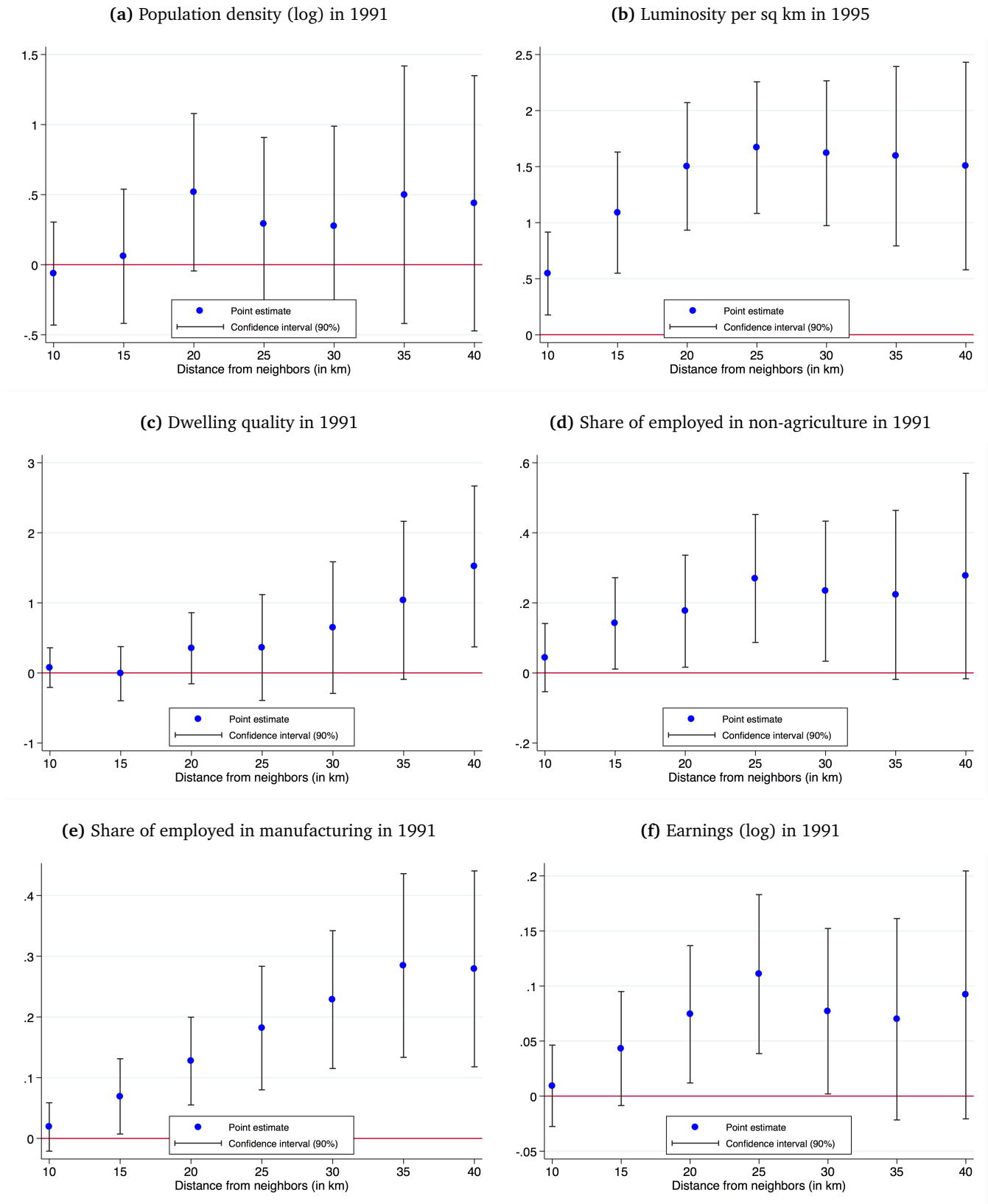
Note: This table presents the quadratic spillover effects of refugees among neighboring municipalities in 1928 on development levels and labour market outcomes in 1991/1995. Neighbouring municipalities are defined as spatially contiguous municipalities of first and second degree (i.e. neighbors of neighbors). The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The sample excludes 50 municipalities that correspond to an entire island and thus that have no spatially contiguous neighbours. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 6, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Spillover effects: 1971 outcomes

Outcome:	Population density (log) in 1971	Dwelling Quality in 1971	Share of employment in non-agriculture in 1971	Share of employment in manufacturing in 1971
	(1)	(2)	(3)	(4)
Share of refugees in 1928	1.926***	2.879***	0.546***	0.152***
	(0.393)	(0.546)	(0.191)	(0.033)
Share of refugees in 1928, squared	-0.024***	-0.031***	-0.007***	-0.002***
	(0.004)	(0.006)	(0.002)	(0.000)
Share of refugees among neighbors	0.099	0.835***	0.030*	0.054**
	(0.326)	(0.220)	(0.018)	(0.025)
N	4,336	4,336	4,336	4,336
R2	.759	.467	.358	.301
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y

Notes. This table presents the spillover effects of refugees among neighbouring municipalities in 1928 on development levels and labour market outcomes in 1971. Neighbouring municipalities are defined as spatially contiguous municipalities of first and second degree (i.e. neighbors of neighbors). The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The sample excludes 50 municipalities that correspond to an entire island and thus that have no spatially contiguous neighbours. The share of refugees varies between 0 and 100. The outcome variable in columns is expressed in logs ($\times$ 100). The outcome variable in columns 2 is standardized (z-scores). The outcome variables in columns 3 and 4 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.10: Sensitivity of estimated spillover effects to definitions of spatial neighbors



Notes. This figure plots the estimated spillover effects of refugees in neighboring municipalities, where whether two municipalities are considered to be spatial neighbours depends on the air distance between their centroids. Specifically, it displays the effect of an increase of a 1 percentage points in the share of refugees in 1928 among neighboring municipalities within a given radius: 10, 15, 20, 25, 30 or 40 kms. Each reported point estimate is obtained from a different regression using the same empirical specification as in Table 5. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart.

G.3 Spatial matching estimates and 1923 refugee distribution

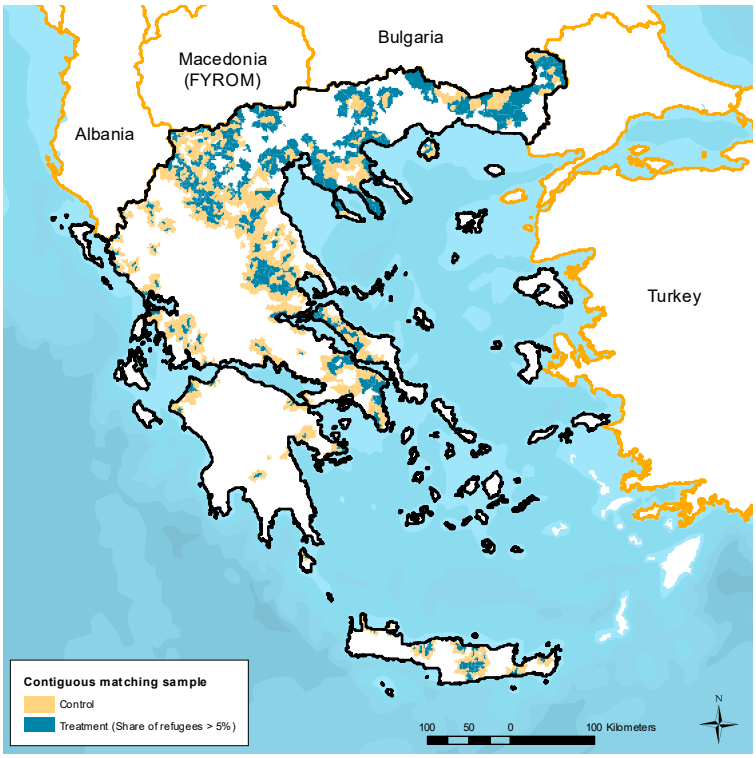
Table A.14: Matching estimator balancing test: Municipality characteristics

Standarized Dependent variable:	Pop. density (log) in 1920	Island	Distance to Railway	Distance to Shore	Distance to Borders	Altitude	Average annual precipitation	Average annual temperature	Suitability of crops
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Spatial matching estimator: contiguous municipalities									
Treated, $\mathbb{1}\{\text{Share of refugees} > 5\%\}$	0.149**	-0.056	-0.115*	0.022	-0.137**	-0.305***	-0.186***	0.055	0.326***
	(0.071)	(0.064)	(0.064)	(0.072)	(0.066)	(0.068)	(0.053)	(0.067)	(0.067)
N	1,059	1,059	1,059	1,059	1,059	1,059	1,059	1,059	1,059
Panel B: Spatial matching estimator: contiguous municipalities and propensity score									
Treated, $\mathbb{1}\{\text{Share of refugees} > 5\%\}$	-0.077	-0.051	-0.077	0.022	-0.106	-0.061	-0.088	-0.056	0.076
	(0.078)	(0.076)	(0.077)	(0.081)	(0.077)	(0.074)	(0.066)	(0.074)	(0.076)
N	774	774	774	774	774	774	774	774	774

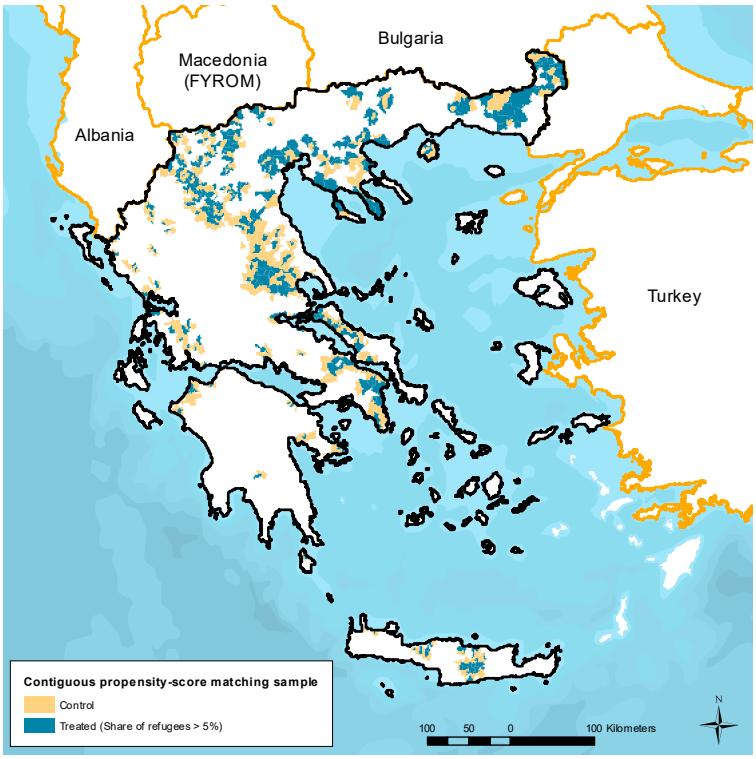
Notes. This table presents the mean differences in observables between treated and control municipalities based on a spatial matching estimator that compares spatially contiguous municipalities. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Treated municipalities are defined as municipalities with at least 5% refugees in 1928. Panel A presents the results for all contiguous treated and control municipalities, whereas panel B presents the results for a sample of contiguous treated and control municipalities that are matched in terms of observables using a propensity score, excluding the least similar 20% of control-treatment municipality pairs. The outcome variables are standardized (z-scores). Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.11: Spatial matching samples

(a) Spatial matching sample



(b) Spatial matching with propensity score sample



Notes. This figure presents the samples used in the spatial matching estimator. Treated municipalities are defined as municipalities with at least 5% refugees in 1928. The blue municipalities are the treatment municipalities, and the yellow municipalities are the matched control municipalities. Panel (a) presents the distribution of all treatment and control municipalities based on our spatial matching estimator whereas panel (b) excludes from this sample the least similar 20% of control-treatment municipality pairs according to a propensity score based on observables. Data sources: The 1928 population census of the Kingdom of Greece.

Table A.15: Robustness to different estimation strategy and different sample

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Matching Estimator: spatially contiguous neighbors and propensity score						
Treated, $\mathbb{1}\{\text{Share of refugees} > 5\%\}$	22.496**	17.541***	12.336**	4.691**	1.220***	2.276***
	(10.346)	(4.513)	(4.996)	(2.113)	(0.274)	(0.740)
N	774	774	774	774	774	774
R2	.012	.01	.006	.008	.003	.014
Average outcome in control group	341.0	24.8	28.8	50.3	8.5	640.7
Panel B: Intensive margin, sample: Share of refugees $> 0.1\%$						
Share of refugees in 1928	2.094***	1.115***	1.308***	0.328***	0.070***	0.144***
	(0.502)	(0.265)	(0.309)	(0.111)	(0.018)	(0.041)
Share of refugees in 1928, squared	-0.025***	-0.013***	-0.015***	-0.004***	-0.001***	-0.002***
	(0.005)	(0.004)	(0.003)	(0.001)	(0.000)	(0.000)
N	2,228	2,228	2,228	2,228	2,228	2,228
R2	.788	.654	.384	.498	.456	.396
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	41.3	43.9	43.6	41.3	37.2	4
Average outcome	360.7	3	22.9	55.4	8.5	642.4

Notes. This table presents the robustness of the effects of refugees in 1928 to different identifying variations. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Panel A reports the results obtained by comparing spatially contiguous treatment and control municipalities that are matched in terms of observables using a propensity score, excluding the least similar 20% of control-treatment municipality pairs. Treated municipalities are defined as municipalities with at least 5% refugees in 1928. Panel B reports the results at the intensive margin, comparing only municipalities with at least 0.1% of refugees in 1928. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16: Robustness to using the share of refugees in 1923

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Municipal-level estimates, OLS						
Share of refugees in 1923	2.092***	1.975***	1.800**	0.762***	0.125**	0.382***
	(0.584)	(0.615)	(0.752)	(0.227)	(0.053)	(0.117)
Share of refugees in 1923, squared	-0.011	-0.021**	-0.019	-0.009***	-0.001	-0.005***
	(0.010)	(0.010)	(0.012)	(0.003)	(0.001)	(0.002)
N	4,388	4,388	4,388	4,388	4,388	4,388
R2	.756	.64	.329	.427	.401	.345
Province FE	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	98.6	46.8	48.4	44.4	48.4	40.1
Average outcome	332.7	0	0	5	7	640.6
Panel B: Matching Estimator: spatially contiguous neighbors and propensity score						
Treated (Sh. refugees in 1923 $>$ 5%)	32.095***	21.955***	16.847***	6.547***	1.603***	3.271***
	(7.738)	(3.408)	(3.508)	(1.815)	(0.593)	(0.779)
N	988	988	988	988	988	988
R2	.023	.019	.012	.02	.008	.03
Average among control municipalities	346.6	29.7	22.8	48.5	7.2	639.8
Panel C: Intensive margin, sample: Share of refugees $>$ 0.1%						
Share of refugees in 1923	2.111***	1.729***	1.744***	0.504***	0.055	0.253***
	(0.604)	(0.508)	(0.603)	(0.183)	(0.038)	(0.086)
Share of refugees in 1923, squared	-0.009	-0.014*	-0.015	-0.005**	-0.000	-0.003**
	(0.010)	(0.008)	(0.010)	(0.002)	(0.001)	(0.001)
N	1,666	1,666	1,666	1,666	1,666	1,666
R2	.785	.675	.457	.547	.496	.492
Province FE	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	123.4	62.4	57.6	52.6	110.9	48.6
Average outcome	375.9	45.7	26.2	57.3	8.6	643.3

Notes. This table shows the effects of refugees in 1923 exploiting different identifying variations. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Panel A reports the robustness of our baseline estimates using the distribution of refugees in 1923. Panel B reports the results obtained by comparing spatially contiguous treatment and control municipalities that are matched in terms of observables using a propensity score, excluding the 20% least similar control-treatment municipality pairs. Treated municipalities are defined as municipalities with at least 5% refugees in 1923. Panel C reports the results at the intensive margin, comparing only municipalities with at least 0.1% of refugees in 1923. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.4 Standard-Errors

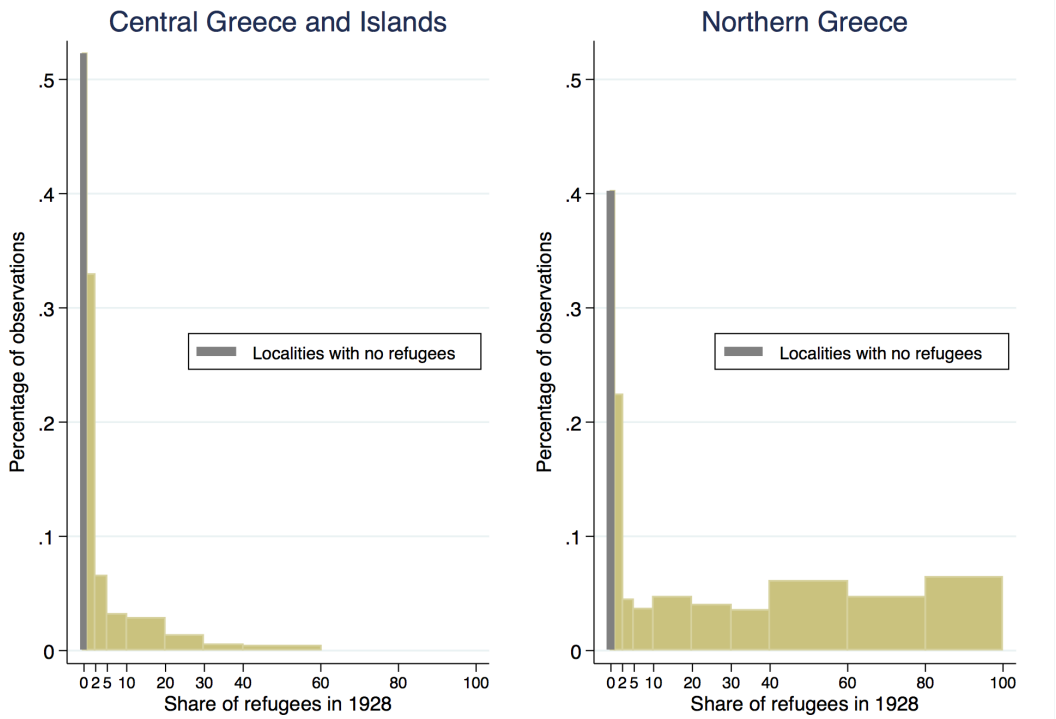
Table A.17: Robustness of inference to different ways of treating the standard errors

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Share of refugees in 1928	2.217	1.246	1.284	0.377	0.093	0.179
(1) Conley correction: 20 km radius	(0.308)***	(0.264)***	(0.286)***	(0.096)***	(0.029)***	(0.037)***
(2) Conley correction: 50 km radius	(0.459)***	(0.293)***	(0.297)***	(0.129)***	(0.023)***	(0.045)***
(3) Conley correction: 100 km radius	(0.602)***	(0.319)***	(0.343)***	(0.154)**	(0.020)***	(0.057)***
(4) Conley correction: 200 km radius	(0.505)***	(0.299)***	(0.470)***	(0.150)**	(0.011)***	(0.062)***
(5) Conley correction: 250 km radius	(0.432)***	(0.219)***	(0.423)***	(0.163)***	(0.014)***	(0.067)***
(6) Clustering at province level	(0.375)***	(0.301)***	(0.294)***	(0.103)***	(0.040)**	(0.038)***
Share of refugees in 1928, squared	-0.027	-0.014	-0.015	-0.005	-0.001	-0.002
(1) Conley correction: 20 km radius	(0.003)***	(0.003)***	(0.003)***	(0.001)***	(0.000)***	(0.000)***
(2) Conley correction: 50 km radius	(0.005)***	(0.004)***	(0.003)***	(0.001)***	(0.000)***	(0.000)***
(3) Conley correction: 100 km radius	(0.007)***	(0.005)***	(0.004)***	(0.002)***	(0.000)***	(0.001)***
(4) Conley correction: 200 km radius	(0.006)***	(0.004)***	(0.005)***	(0.002)***	(0.000)***	(0.001)***
(5) Conley correction: 250 km radius	(0.005)***	(0.003)***	(0.005)***	(0.002)***	(0.000)***	(0.001)***
(6) Clustering at province level	(0.004)***	(0.004)***	(0.003)***	(0.001)***	(0.000)**	(0.000)***
N	4,388	4,388	4,388	4,388	4,388	4,388
R2	.756	.64	.329	.427	.401	.345
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y

Notes. This table reports the robustness of the estimated effects of refugees in 1928 to different assumptions about the Variance-Covariance (VCV) matrix. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Each model (1) to (6) reports the standard errors (in parentheses) obtained from a different regression with varying assumptions regarding the correlation in residuals across municipalities. In models (1) to (5), standard errors are corrected for spatial correlation following Conley (1999) and Colella et al. (2019), assuming that the spatial correlation is null between municipalities that are more than 20kms, 50 kms, 100 kms, 200 kms, and 250 kms apart. In model (6), we correct the standard errors for correlations across municipalities within the same province. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H Regional heterogeneity

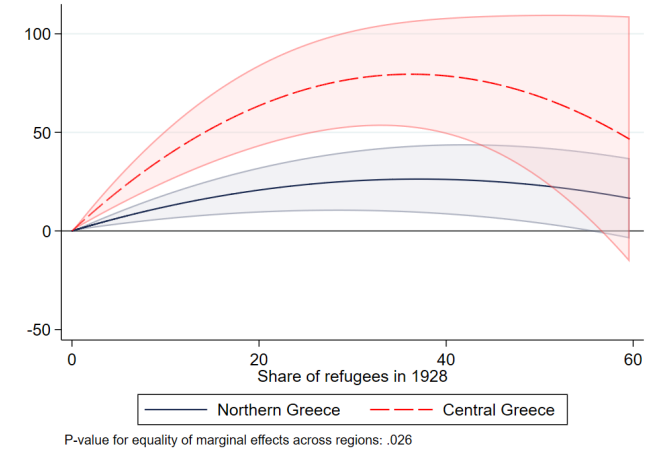
Figure A.12: Histogram of the share of refugees in 1928 by region at the municipality level



Notes. This figure presents the histogram of the share of refugees in 1928 separately for Central Greece and Islands and for Northern Greece. Data source: The 1928 population census of the Kingdom of Greece.

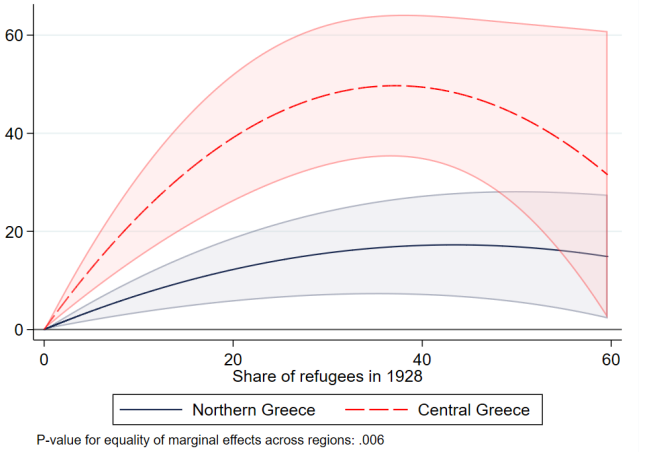
Figure A.13: Regional heterogeneity in estimated effects

(a) Population density (log) in 1991



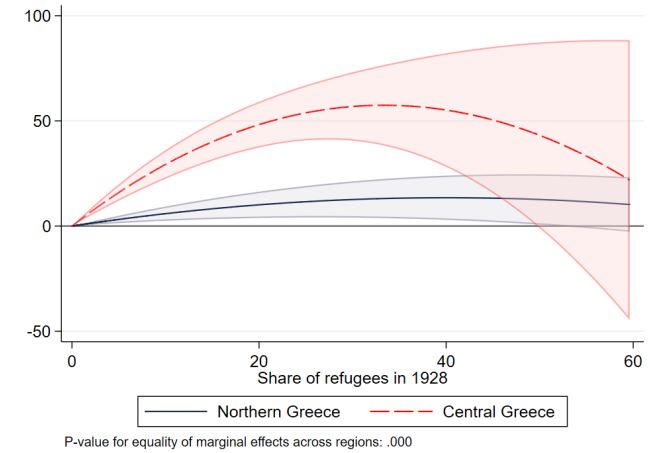
P-value for equality of marginal effects across regions: .026

(b) Luminosity per sq km in 1995



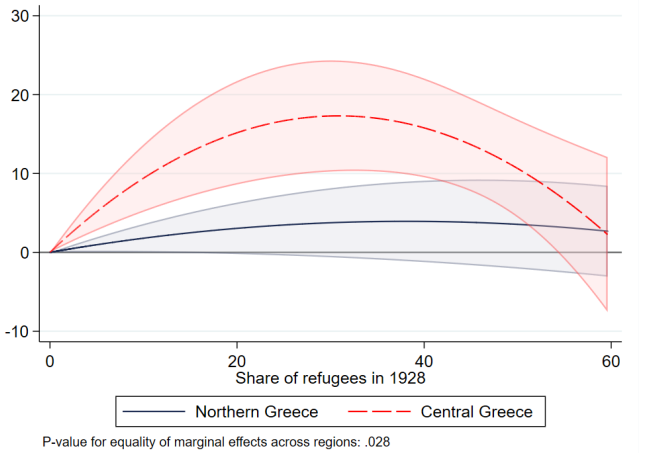
P-value for equality of marginal effects across regions: .006

(c) Dwelling quality index in 1991



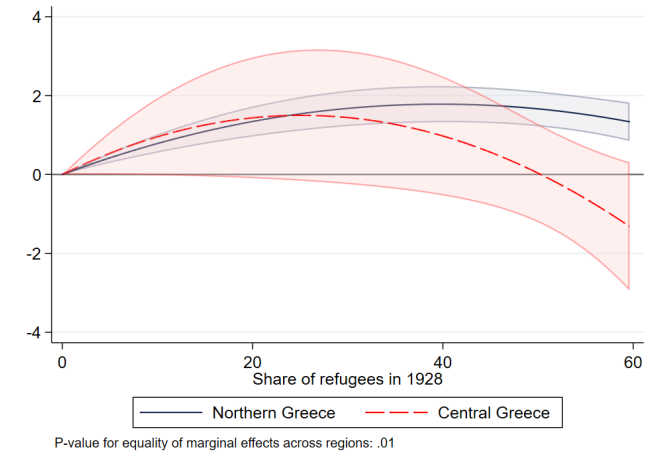
P-value for equality of marginal effects across regions: .000

(d) Share of employment in non-agriculture in 1991



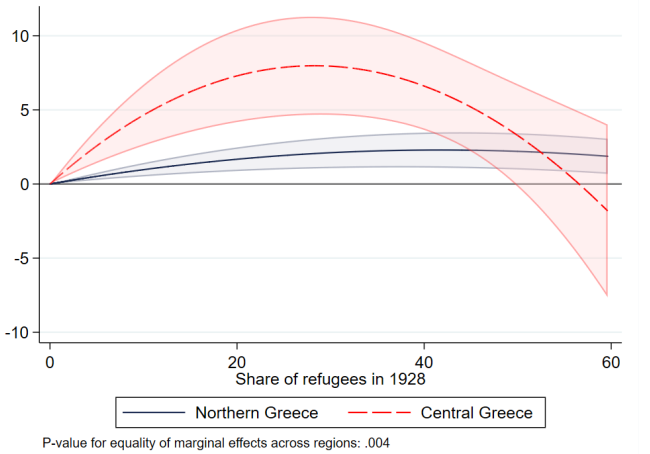
P-value for equality of marginal effects across regions: .028

(e) Share of employment in manufacturing in 1991



P-value for equality of marginal effects across regions: .01

(f) Earnings (log) in 1991

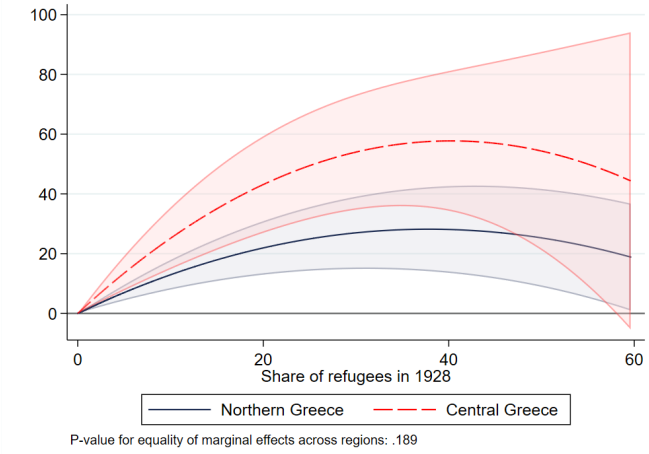


P-value for equality of marginal effects across regions: .004

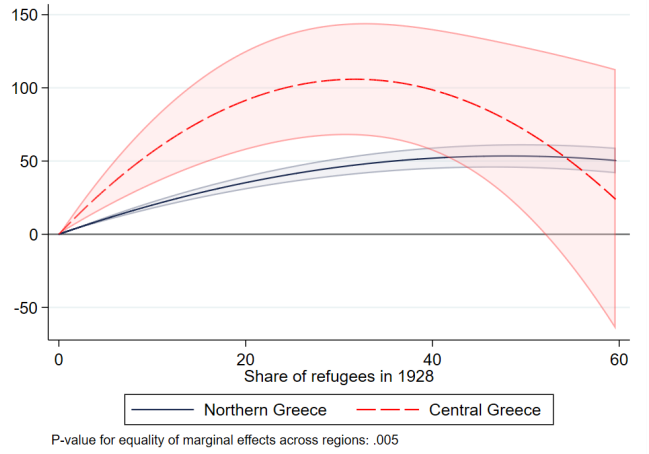
Notes. This figure plots the estimated quadratic effects of refugees in 1928 on development and labor market outcomes in 1991/1995 using the regression coefficients reported in Table 7, along with corresponding 90% confidence intervals. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Red dashed lines represent the estimated quadratic effects across municipalities within Central Greece, whereas the blue solid lines represent those across municipalities within Northern Greece. Standard errors are corrected for spatial correlation following Conley (1999) and Colella et al. (2019). We assume that the spatial correlation is null between municipalities that are more than 100 kms apart.

Figure A.14: Regional heterogeneity: 1971 outcomes

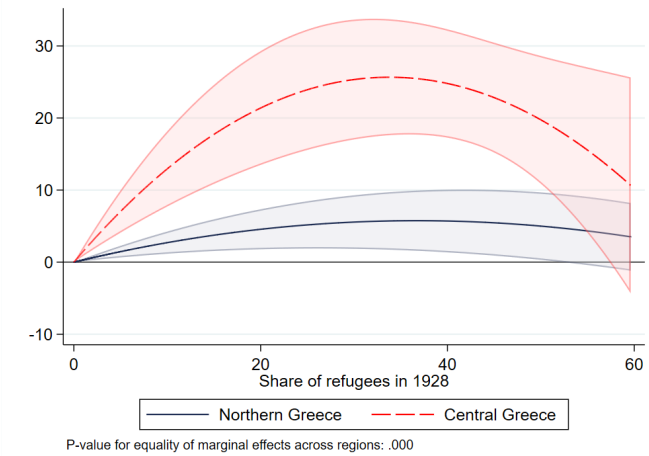
(a) Population density (log) in 1971



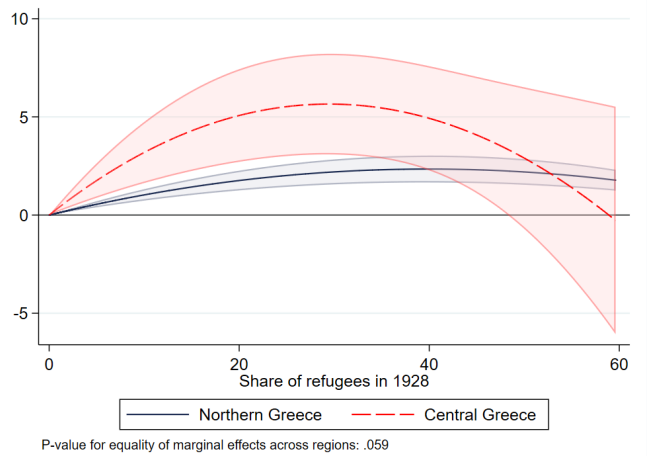
(b) Dwelling quality index in 1971



(c) Share of employed in non-agriculture in 1971



(d) Share of employed in manufacturing in 1971



Notes. This figure plots the estimated quadratic effects of refugees in 1928 on development and labor market outcomes in 1971 using the same specifications as those reported in Table 7, along with corresponding 90% confidence intervals. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Red dashed lines represent the estimated quadratic effects across municipalities within Central Greece, whereas the blue solid lines represent those across municipalities within Northern Greece. Standard errors are corrected for spatial correlation following Conley (1999) and Colella et al. (2019). We assume that the spatial correlation is null between municipalities that are more than 100 kms apart.

I Mechanisms: Skill complementarity in Northern Greece

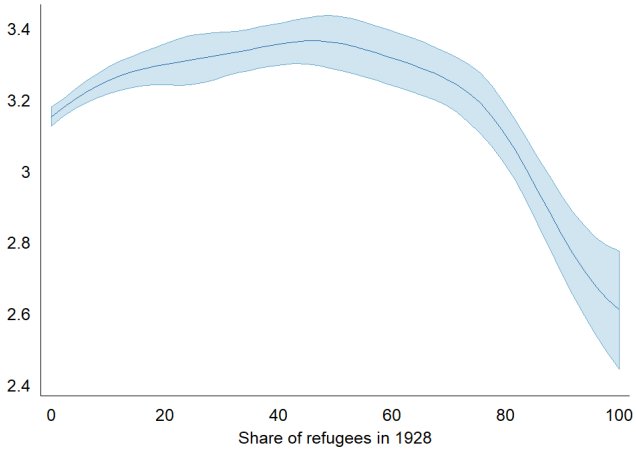
Table A.18: Transfer of textile-specific knowledge: Robustness

Outcome:	Share of textile sector in 1991		Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991
	(1)	(2)	(3)*	(4)
Share of refugees in 1928	0.088***	0.082***	0.104	0.069***
	(0.012)	(0.019)	(0.126)	(0.013)
Share of refugees in 1928, squared	-0.001***	-0.001***	-0.002**	-0.001***
	(0.000)	(0.000)	(0.001)	(0.000)
High suitability for cotton	0.989	0.565	-3.542	0.328
	(0.706)	(0.740)	(2.631)	(0.762)
Share of refugees in 1928 * High suitability for cotton		0.023**	0.096**	0.032**
		(0.009)	(0.047)	(0.013)
High suitablity for cereals		-0.687**	-2.890	-1.793***
		(0.313)	(1.883)	(0.479)
Share of refugees in 1928 * High suitablity for cereals		-0.005	0.086*	0.013
		(0.014)	(0.044)	(0.020)
N	1,650	1,650	1,650	1,650
R2	.505	.507	.431	.391
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Share of refugees maximizing outcome	44.2	41.9	22.4	32.4
Average outcome	3.8	3.8	53.6	9.2

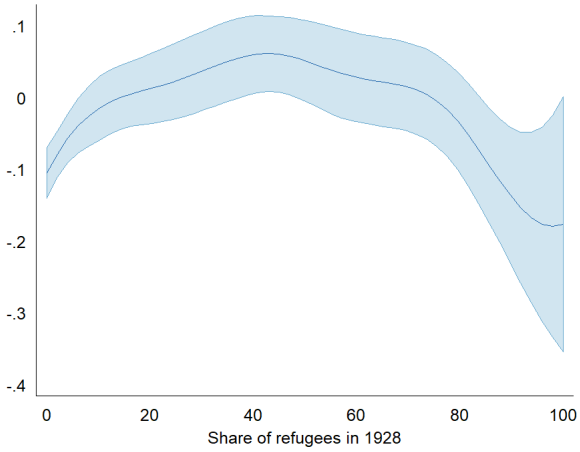
Notes. This table presents the robustness of the estimated heterogeneous effects of refugees in 1928 on industrialization levels in 1991 by suitability for cotton cultivation to controlling for the heterogenous effects of refugees in 1928 by suitability for cereal cultivation. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The analysis is restricted to municipalities in Northern Greece. "High suitability for cotton" is a dummy variable indicating whether a municipality is in the top tercile of the sample in terms of suitability for cotton cultivation. "High suitability for cereals" is a dummy variable indicating whether a municipality is in the top tercile of the sample in terms of suitability for cereal cultivation, namely wheat, oat, and barley. The share of refugees varies between 0 and 100. All outcome variables are expressed in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 3, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.15: Semiparametric effects of refugees in 1928, Northern Greece

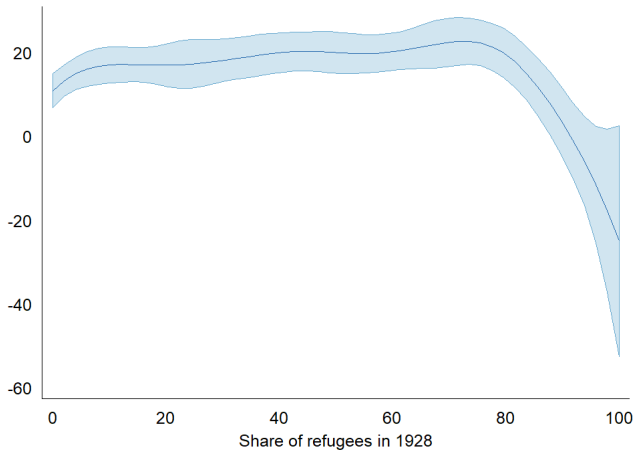
(a) Population density (log) in 1991



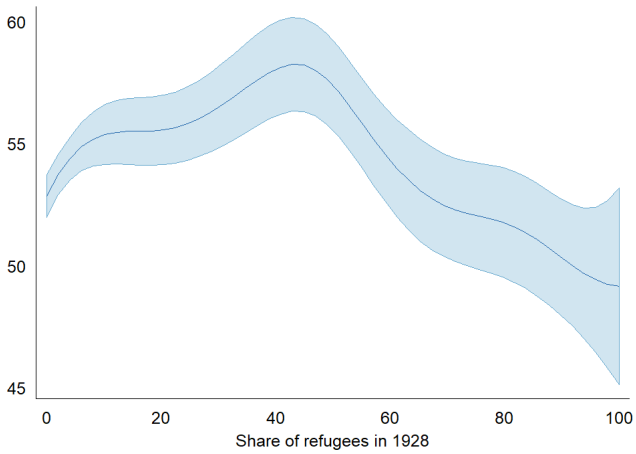
(b) Luminosity per sq km in 1995



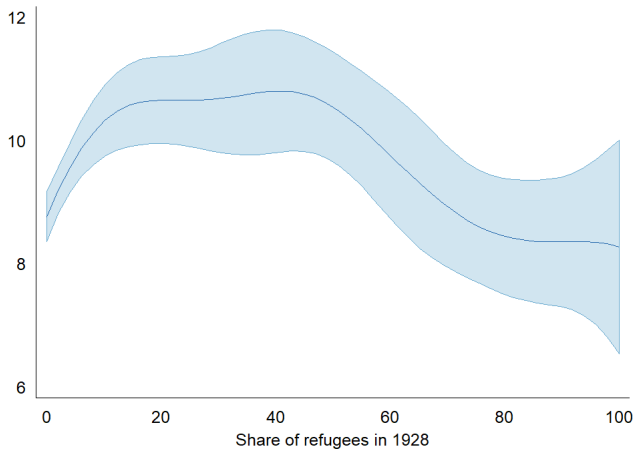
(c) Dwelling quality in 1991



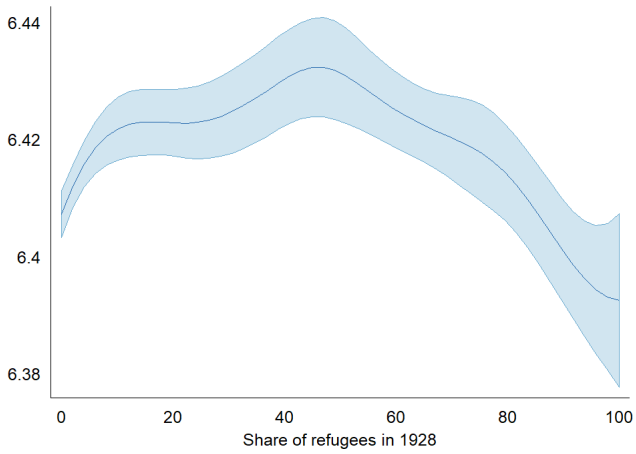
(d) Share of employed in non-agriculture in 1991



(e) Share of employed in manufacturing in 1991



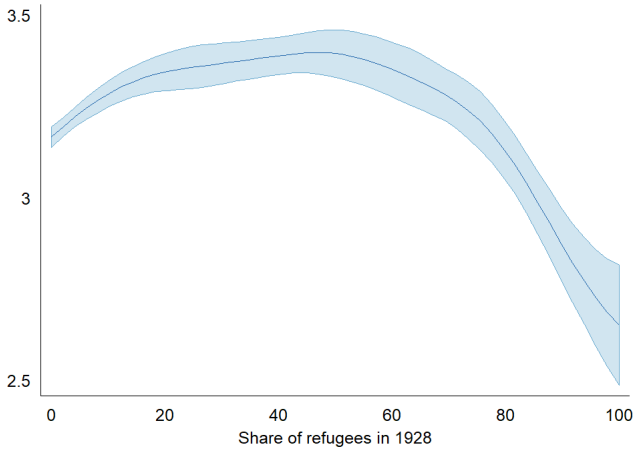
(f) Earnings (log) in 1991



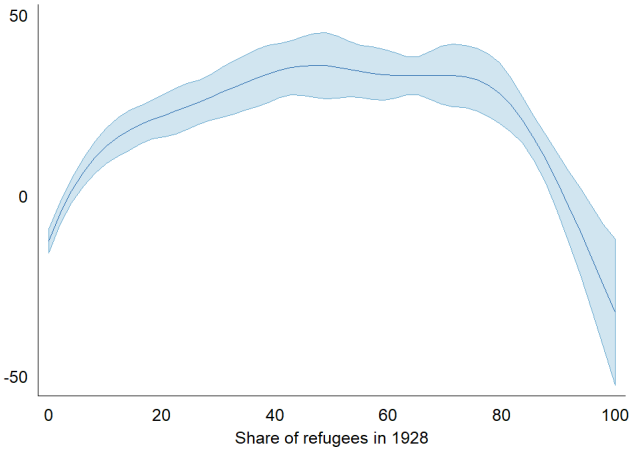
Notes. This figure presents the kernel-weighted local polynomial fits showing the relationship between the share of refugees in 1928 and development levels in 1991/1995 across municipalities within Northern Greece, controlling for province fixed effects; log population in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude; mean annual precipitation and temperature; and crop suitability. We perform these estimations using the `semipar` command provided by the statistical software Stata 15, which is based on Robinson (1988)'s double residual estimator. Standard errors are clustered at the province level and 90% confidence interval are reported.

Figure A.16: Semiparametric effects of refugees in 1928, Macedonia: 1971 outcomes

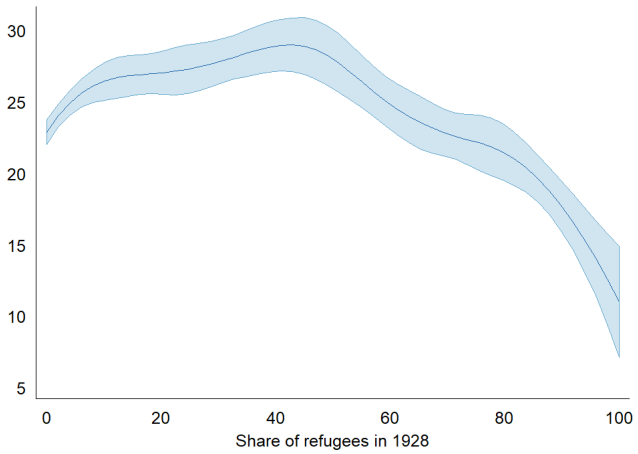
(a) Population density (log) in 1971



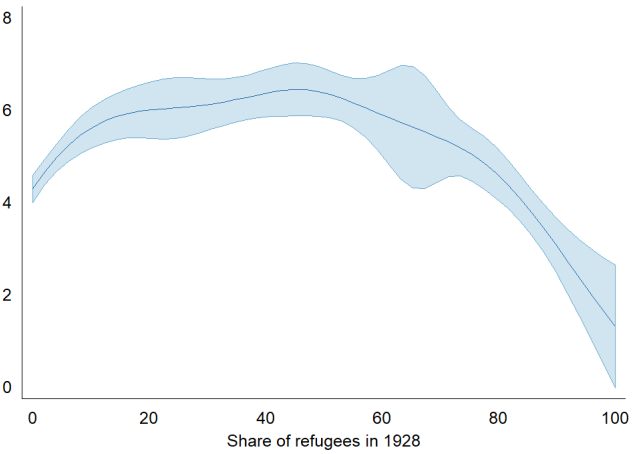
(b) Dwelling quality in 1971



(c) Share of employed in non-agriculture in 1971



(d) Share of employed in manufacturing in 1971



Notes. This figure presents the kernel-weighted local polynomial fits showing the relationship between the share of refugees in 1928 and development levels in 1971 across municipalities within Northern Greece, controlling for province fixed effects; log population in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude; mean annual precipitation and temperature; and crop suitability. We perform these estimations using the `semipar` command provided by the statistical software Stata 15, which is based on Robinson (1988)'s double residual estimator. Standard errors are clustered at the province level and 90% confidence interval are reported.

Table A.19: Non-linear effects of refugees in Northern Greece

Outcome:	Population density (log)		Luminosity per sq km	Dwelling quality		Share of employment in non-agriculture		manufacturing		Earnings (log)
	1971	1991	1995	1971	1991	1971	1991	1971	1991	1991
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Relative effects of dummy variables:										
β_1: Share of refugees btw. 1-20% relative to less 1%	10.355**	10.228**	17.203***	36.953***	20.390***	5.710***	4.774***	1.543***	1.888**	2.807***
	(4.107)	(4.464)	(5.559)	(4.093)	(3.333)	(1.440)	(1.079)	(0.581)	(0.807)	(0.538)
β_2: Share of refugees btw. 20-60% relative to 1-20%	9.511	5.908	-2.095	9.105**	-11.772**	-0.590	-0.371	0.188	-0.605	-0.654
	(7.232)	(6.853)	(3.211)	(3.944)	(4.947)	(1.646)	(2.400)	(0.176)	(0.632)	(0.653)
β_3: Share of refugees btw. 60-100% relative to 20-60%	-34.523***	-33.036***	-10.117	-15.899***	-4.322	-8.616***	-6.025***	-2.197***	-1.869**	-1.600***
	(9.326)	(10.497)	(8.056)	(5.689)	(5.746)	(1.045)	(1.649)	(0.595)	(0.730)	(0.253)
N	1,642	1,650	1,650	1,642	1,650	1,642	1,650	1,642	1,650	1,650
R2	.699	.705	.6	.497	.33	.308	.43	.263	.398	.292
Province FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Average outcome	322.37	318.45	-6	2.31	13.63	23.96	53.66	4.78	9.27	641.28

Notes. This table presents the relative effects of refugees depending on their share in the municipal population. The reported β coefficients are estimated using the following regression: $Y_{mp} = \beta_1 * \mathbb{1}\{\text{ Refugee share } \geq 1\%\} + \beta_2 * \mathbb{1}\{\text{Refugee share } \geq 20\%\} + \beta_3 * \mathbb{1}\{\text{Refugee share } \geq 60\%\} + \alpha X'_{mp} + \delta_p + \varepsilon_{mp}$. All specifications control for province fixed effects and a set of municipal-level covariates, X_{mp} which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; mean annual precipitation and temperature, and crop suitability of the municipality. In addition, the set of covariates, we include a quadratic function of altitude. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.20: Test for non-linear selection of refugees within Northern Greece

Dependent variable:		Population density	Suitability	Distance to			Average annual	
	Altitude (m)	(log) in 1920	of crops	Railway in 1920	Shore	Borders	temperature	precipitation
	(1)*	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relative effects of dummy variables:								
β_1: Share of refugees btw. 1-20% relative to less 1%	-107.011***	0.221***	0.126**	-1.490***	0.439	-0.916	-0.014	-6.165***
	(15.898)	(0.075)	(0.060)	(0.378)	(0.692)	(1.112)	(0.031)	(1.336)
β_2: Share of refugees btw. 20-60% relative to 1-20%	-170.207***	-0.001	0.004	-1.261	0.446	-2.761**	-0.060	6.632*
	(31.770)	(0.107)	(0.067)	(0.927)	(1.663)	(1.311)	(0.066)	(3.493)
β_3: Share of refugees btw. 60-100% relative to 20-60%	69.117***	-0.068	-0.030	-0.985	0.151	-1.897**	-0.001	-3.235
	(23.646)	(0.095)	(0.034)	(0.872)	(1.198)	(0.900)	(0.048)	(3.124)
N	1,659	1,659	1,659	1,659	1,659	1,659	1,659	1,659
R2	.473	.354	.754	.922	.9	.952	.944	.957
Province FE	Y	Y	Y	Y	Y	Y	Y	Y
Elevation control	NO	Y	Y	Y	Y	Y	Y	Y
Average outcome	532.15	3.18	.2	38.12	47.25	65.22	12.29	734.73

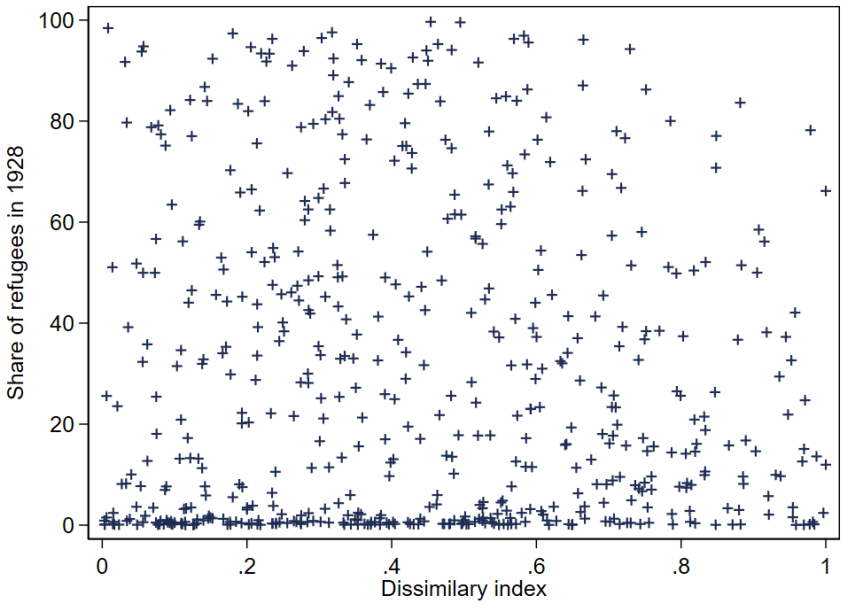
Note: This table presents the differences in population density and geographical characteristics between municipalities with different shares of refugees. The βs coefficients are estimated using the following regression: $Y_{mp} = \beta_1 * \mathbb{1}\{$ Refugee share $\geq 1\%\} + \beta_2 * \mathbb{1}\{$Refugee share $\geq 20\%\} + \beta_3 * \mathbb{1}\{$Refugee share $\geq 60\%\} + \delta_p + \varepsilon_{mp}$. All specifications control for province fixed effects. Except for column 1, all regressions also include as controls a quadratic function of the municipality's altitude. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 1, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.21: Spatial segregation between refugees and natives within municipalities in Northern Greece: 1971 outcomes

Outcome:	Population density (log) in 1971	Dwelling Quality in 1971	Share of employment in non-agriculture in 1971	Share of employment in manufacturing in 1971
	(1)	(2)*	(3)	(4)
Share of refugees in 1928	1.227***	2.164***	0.396***	0.163***
	(0.442)	(0.385)	(0.099)	(0.046)
Share of refugees in 1928, squared	-0.018***	-0.023***	-0.005***	-0.002***
	(0.004)	(0.004)	(0.001)	(0.001)
Dissimilary index (Z-score)	-2.232	-8.421**	-1.902***	-0.745***
	(2.104)	(4.053)	(0.666)	(0.251)
N	558	558	558	558
R2	.747	.56	.533	.485
Province FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Share of refugees maximizing outcome	33.04	47.78	36.83	42.36
Average outcome	341.7	41.7	25.3	5.6

Notes. This table presents the effects of spatial segregation between refugees and natives in 1928 across settlements within a municipality on development levels and labour market outcomes in 1971. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The analysis is restricted to municipalities in Northern Greece with at least two settlements (or villages) and hosting some refugees in 1928. The spatial "Dissimilarity index" is standardized (z-score). The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects, the number of settlements within the municipality, and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. However, in column 2, we relax this assumption to 50 kms only, as otherwise we are not able to estimate the standard errors due to lack of independent variation across clusters. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.17: Spatial segregation and share of refugees in 1928 in Northern Greece



Notes. This figure plots the share of refugees in 1928 against the dissimilarity index showing the spatial segregation between refugees and natives in 1928 across settlements within a municipality. Data source: The 1928 population census of the Kingdom of Greece.

J Alternative Mechanisms

J.1 Muslim economic legacy

Table A.22: Mechanism: Historical Muslim presence, Municipal-level estimates

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All municipalities						
Share of refugees in 1928	2.187***	0.964***	1.025***	0.253	0.071***	0.149**
	(0.700)	(0.333)	(0.289)	(0.163)	(0.014)	(0.063)
Share of refugees in 1928, squared	-0.022***	-0.009**	-0.012***	-0.003	-0.001***	-0.002**
	(0.006)	(0.004)	(0.003)	(0.002)	(0.000)	(0.001)
Mostly-Muslim locality in 1915	-50.991***	-20.968***	-2.429	-8.437***	-2.323**	-2.416**
	(16.939)	(5.549)	(17.402)	(2.157)	(1.034)	(0.950)
N	3,348	3,348	3,348	3,348	3,348	3,348
R2	.761	.638	.337	.417	.449	.343
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	49.3	55.65	41.8	45.2	45.55	46.4
Average outcome	332.6	0	0	53.9	7.0	640.5
Coefficients without Muslims localities in controls						
Share of refugees in 1928	2.285	1.005	1.029	.269	.075	.154
Share of refugees in 1928, squared	-.027	-.011	-.012	-.004	-.001	-.002
Panel B: Municipalities in Northern Greece						
Share of refugees in 1928	1.420***	0.551**	0.897**	0.071	0.048***	0.091***
	(0.356)	(0.235)	(0.377)	(0.093)	(0.009)	(0.020)
Share of refugees in 1928, squared	-0.014***	-0.004***	-0.010**	-0.001	-0.001***	-0.001***
	(0.003)	(0.001)	(0.004)	(0.001)	(0.000)	(0.000)
Mostly-Muslim locality in 1915	-48.136***	-19.710**	-2.692	-9.020***	-2.662**	-2.628***
	(17.248)	(8.035)	(19.427)	(1.441)	(1.035)	(0.824)
N	994	994	994	994	994	994
R2	.709	.554	.291	.399	.41	.307
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	49.25	65.75	43.85	43.85	41.65	50.15
Average outcome	318.4	-6	13.6	53.6	9.2	641.2
Coefficients without Muslims localities in controls						
Share of refugees in 1928	1.489	.58	.901	.084	.052	.095
Share of refugees in 1928, squared	-.019	-.006	-.01	-.002	-.001	-.001

Notes. This table presents the effects of refugees in 1928 and those of the historical presence of Muslims on development levels and labour market outcomes in 1991/1995. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. Panel A reports the estimated effects for all Greece, whereas panel B reports them for municipalities in Northern Greece. The share of refugees varies between 0 and 100. Data on whether a municipality was a Mostly-Muslim locality, i.e., where Muslims made more than 75% of the local population, in Macedonia in 1915 come from Güvenç (2010). For municipalities outside of Macedonia, we assume that they were not a Mostly-Muslim locality in 1915 if the share of Muslims left between 1920 and 1928 is less than 0.1% of the province population. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.23: Mechanism: Historical Muslim presence, Province-level estimates

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Share of refugees in 1928	3.067**	0.063	3.351***	0.428**	0.218**	0.134
	(1.275)	(0.629)	(1.070)	(0.215)	(0.108)	(0.089)
Share of refugees in 1928, squared	-0.026**	0.007	-0.023**	-0.002	-0.001	-0.000
	(0.011)	(0.007)	(0.010)	(0.002)	(0.002)	(0.001)
Share of Muslims left between 1923 and 1928	-0.726	-0.341	-0.615	-0.080	-0.026	-0.021
	(0.594)	(0.253)	(0.424)	(0.115)	(0.063)	(0.050)
N	176	176	176	176	176	176
R2	.928	.916	.671	.686	.461	.727
Region FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Share of refugees maximizing outcome	59.2	-4.3	73.6	136.1	84.3	141.3
Average outcome	422.9	0	0	67.7	9.9	648.4
Coefficients without Share of Muslims in controls						
Share of refugees in 1928	2.713	-.104	3.051	.389	.205	.123
Share of refugees in 1928, squared	-.031	.005	-.027	-.002	-.001	-.001

Notes. This table presents the effects of refugees in 1928 and those of the share of Muslims left between 1920 and 1928 on development levels and labour market outcomes in 1991/1995. The unit of analysis is the urban and rural areas of provinces as defined by the 1920 population census of the Kingdom of Greece. (See section B.3 in the Supplementary Appendix for details). The share of refugees in 1928 and the share of Muslims left between 1923 and 1928 vary between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for region fixed effects and a set of province-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the observation unit was a city in 1920 and whether it is an island; the distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the province. Heteroskedasticity-robust standard errors are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J.2 Infrastructure Investement

Table A.24: Heterogeneous effects by region within Northern Greece

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Northern Greece, Macedonia						
Share of refugees in 1928	1.392***	0.624**	1.219***	0.090	0.075**	0.102***
	(0.160)	(0.282)	(0.338)	(0.131)	(0.035)	(0.030)
Share of refugees in 1928, squared	-0.018***	-0.007**	-0.013***	-0.002	-0.001***	-0.001***
	(0.002)	(0.003)	(0.003)	(0.001)	(0.000)	(0.000)
N	674	674	674	674	674	674
R2	.697	.589	.229	.429	.433	.386
Share of refugees maximizing outcome	38.2	46.3	46.5	27.4	37.3	43.5
Average outcome	317.72	13.54	40.03	55.57	12.26	642.1
Panel B: Northern Greece, Macedonia excluded						
Share of refugees in 1928	1.496	2.145***	1.062	0.637***	0.172*	0.263***
	(0.924)	(0.380)	(0.652)	(0.220)	(0.093)	(0.098)
Share of refugees in 1928, squared	-0.024**	-0.029***	-0.012	-0.005***	-0.002	-0.003***
	(0.010)	(0.005)	(0.008)	(0.002)	(0.001)	(0.001)
N	976	976	976	976	976	976
R2	.726	.585	.328	.451	.218	.23
Share of refugees maximizing outcome	31.8	36.8	45.7	58.7	50.2	45.7
Average outcome	318.95	-19.5	-4.60	52.35	7.22	640.72
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y

Notes. This table presents the heterogeneous effects of refugees in 1928 by regions within Northern Greece. Panels A and B present the effects of refugees in Macedonia and in other regions than Macedonia, respectively. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs ($\times$ 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.25: Public-work programs in Macedonia: heterogeneous impacts

Outcome:	Population density (log) in 1991	Luminosity per sq km in 1995	Dwelling quality in 1991	Share of employment in non-agriculture in 1991	Share of employment in manufacturing in 1991	Earnings (log) in 1991
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: RSC-built dispensaries in 1929						
Share of refugees in 1928	1.324***	0.612**	1.164***	0.074	0.073**	0.100***
	(0.140)	(0.298)	(0.343)	(0.129)	(0.034)	(0.032)
Share of refugees in 1928, squared	-0.019***	-0.007**	-0.013***	-0.002*	-0.001***	-0.001***
	(0.003)	(0.003)	(0.003)	(0.001)	(0.000)	(0.000)
Share of refugees × Dispensary < 10km	0.456***	0.058	0.341**	0.109	0.016	0.013
	(0.108)	(0.158)	(0.164)	(0.082)	(0.029)	(0.024)
Dispensary < 10km	-29.243***	-11.137	-30.862*	-4.833	-0.721	-1.484
	(4.486)	(7.265)	(18.374)	(3.938)	(1.608)	(1.071)
N	674	674	674	674	674	674
R2	.701	.59	.24	.433	.433	.387
Panel B: RSC Reclamation works						
Share of refugees in 1928	1.259***	0.608**	1.238***	0.064	0.064*	0.094***
	(0.113)	(0.267)	(0.346)	(0.132)	(0.038)	(0.030)
Share of refugees in 1928, squared	-0.017***	-0.007**	-0.013***	-0.001	-0.001***	-0.001***
	(0.002)	(0.003)	(0.003)	(0.001)	(0.000)	(0.000)
Share of refugees × Reclamation < 10km	0.640***	0.079	-0.079	0.124	0.056	0.039
	(0.204)	(0.303)	(0.176)	(0.093)	(0.035)	(0.032)
Reclamation < 10km	-20.006***	-2.613	13.762	-4.770	-1.590	-1.599
	(7.586)	(21.231)	(11.569)	(5.270)	(1.971)	(1.926)
N	674	674	674	674	674	674
R2	.699	.589	.23	.431	.435	.387
Province FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y

Notes. This table presents the heterogeneous effects of refugees in 1928 with respect to presence of public investments by the Refugee Settlement Commission (RSC) within 10 km distance to the municipality. Panels A and B present the heterogeneous effects of refugees with respect to presence of RSC-built dispensaries in 1929 and that of RSC reclamation works, respectively. The unit of analysis is the municipality as defined by the 1920 population census of the Kingdom of Greece. The share of refugees varies between 0 and 100. The outcome variables in columns 1 and 6 are expressed in logs (× 100). The outcome variables in columns 2 and 3 are standardized (z-scores). The outcome variables in columns 4 and 5 are in shares, varying between 0 and 100. All specifications control for province fixed effects and a set of municipal-level covariates, which comprises: log population density in 1920; dummy variables indicating whether the municipality was a city in 1920 and whether it is an island; distances to railway network in 1920, to shoreline, and to national borders; altitude, mean annual precipitation and temperature, and crop suitability of the municipality. Standard errors corrected for spatial correlation following Conley (1999) and Colella et al. (2019) are reported in parentheses. We assume that the spatial correlation is null between municipalities that are more than 100 kms apart. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J.3 Second-generation refugees

Table A.26: Share of second-generation refugees in Greece (European Social Survey: rounds 2002, 2004, and 2010)

	Percentage of respondents with		
Decade of birth	Parents born in Turkey	Foreign-born parents (outside Turkey)	Number of respondents
1910–1920	2.82	2.82	71
1920–1930	10.91	4.65	559
1930–1940	14.70	3.09	1102
1940–1950	9.18	3.13	991
1950–1960	3.59	1.64	976
1960–1970	1.32	1.23	1137
1970–1980	0.36	1.19	1097
1980–1990	0.26	2.08	768
1990–2000	0.00	7.58	211

Notes. This table shows the share of respondents with at least one parent born in Turkey and with at least parent born in another foreign country other than Turkey across decades of birth. Data Sources: European Social Survey, Greece sample for rounds 2002, 2004, and 2010.

Table A.27: Mechanism: Second-generation refugees vs. natives (ESS in the 2000s)

Dependent var.	Educational attainment			Last or current occupation						Household income		
	Junior High School			Agriculture			High-skilled occupation			in deciles		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Parents born in Turkey	-0.012	0.044	0.031	-0.029	-0.060	-0.056	-0.076***	-0.042*	-0.041*	-0.118	0.228	0.156
	(0.031)	(0.032)	(0.029)	(0.041)	(0.040)	(0.040)	(0.022)	(0.025)	(0.025)	(0.170)	(0.172)	(0.167)
N	1,715	1,715	1,715	1,305	1,305	1,305	1,305	1,305	1,305	1,281	1,281	1,281
R2	0.08	0.21	0.31	0.08	0.29	0.33	0.07	0.15	0.21	0.16	0.29	0.32
Region FE		Y	Y		Y	Y		Y	Y		Y	Y
Parental characteristics			Y			Y			Y			Y
Average outcome	.207	.207	.207	.431	.431	.431	.126	.126	.126	3.93	3.93	3.95

Notes. This table presents the mean differences in outcomes between second-generation refugees and others among a sample of individuals born in Greece between 1923 and 1950. Data sources: The 2002, 2004 and 2010 rounds of the European Social Survey in Greece. Parental characteristics include mother's and father's educational attainment and their employment statuses when respondents was aged 14. The outcome variable in columns 1 to 3 is a dummy variable indicating whether the respondent graduated from junior high school. The outcome variable in columns 4 to 6 is a dummy variable indicating whether the last or current occupation of respondents is in agriculture sector. The outcome variable in columns 7 to 9 is a dummy variable indicating whether the last or current occupation of respondents is a high-skilled occupation. High-skilled occupations comprise the first three ISCO-88 major groups, that is, "Legislators/Senior officials/Managers", "Professionals", and "Technicians and Associate Professionals". The outcome variable in columns 10 to 12 is an ordered categorical variable showing to which decile of the distribution household income corresponds. All regressions include as controls respondent's gender and year of birth, as well as survey rounds fixed-effects. Parental characteristics include father's and mother's education, employment type and whether they were alive when respondent was aged 14. Heteroskedasticity-robust standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

References

Alesina, Alberto, and Ekaterina Zhuravskaya. 2011. "Segregation and the Quality of Government in a Cross Section of Countries." *American Economic Review*, 101(5): 1872–1911.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2019. "Inference with Arbitrary Clustering." *IZA Discussion Paper No 12584*.

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

Fischer, G., F. Nachtergaele S. Prieler H.T. van Velthuizen L. Verelst D. Wiberg. 2008. *Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008)*. IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Güvenç, Sefer. 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.

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

Masters, Bruce. 2009. "Millet." In *Encyclopedia of the Ottoman Empire*. , ed. Gábor Ágoston and Bruce Masters. New York:Facts On File.
˙Ministry of National Economy

Ministry of National Economy. 1950. *Atlas des municipalites et commmunes de la Grece: suivant division administrative du 31/12/1948*. General Statistical Service of Greece: Athens.

Reardon, Sean F, and Glenn Firebaugh. 2002. "Measures of multigroup segregation." *Sociological methodology*, 32(1): 33–67.

Robinson, P. M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica*, 56(4): 931–954.