Geography, Land Ownership and Literacy: Historical Evidence from Greek Regions

Benos, Nikos and Karagiannis, Stelios and Tsitou, Sofia

University of Ioannina

2022

Geography, Land Ownership and Literacy: Historical Evidence from Greek Regions

Nikos Benos[*]

Stelios Karagiannis[**]

Sofia Tsitou[***]

Abstract

This study examines the impact of land ownership on literacy rates in a cross-section of Greek provinces around 1900. Consistent with our theoretical framework (Galor et al., 2009), we find that the dominance of large properties has a substantial adverse effect on human capital accumulation. Thus, our evidence explains a substantial part of provincial differences in terms of human capital in early 20th century Greece for the first time. This differs from much of the literature, because Greece was at the early stages of the transition to the industrial era during the period examined.

Keywords Landownership - Human capital - Geography - Regional analysis

JEL Classification I20, N34, Q15, C21

***Acknowledgment**:* The financial support of this research by the Hellenic Foundation for Research and Innovation under the 'First call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant' is gratefully acknowledged. This work was part of the project OSCICEEP (2084) .

[*] Department of Economics, Ioannina, Greece.

[**] European Training Foundation, Turin, Italy.

[***] Department of Economics, University of Macedonia, Thessaloniki, Greece.

Corresponding author: Nikos Benos, Department of Economics, Ioannina Campus, Ioannina 45100, Greece. *Phone*: (+30) 26510 05955, Fax: (+30) 26510 05092 *Email:* nbenos@uoi.gr

1. Introduction

The role of geographical factors, institutional factors and human capital formation has been the center of a discussion about the origin of the differential timing of the transition from stagnation to growth and the important changes in income distribution which have taken place since the onset of the industrial revolution. There is a huge literature analyzing geography as a very important factor for comparative economic development (Baten and Hippe, 2018; Easterly, 2007; Adamopoulos and Restuccia, 2018). Also, a number of papers support that inequality in the distribution of land has affected human capital formation and the transition from an agricultural to an industrial economy. Galor et al. (2009, GMV) provides the theoretical foundation behind the fact that inequality in land ownership has an adverse effect on human capital accumulation. During the transition from agricultural to industrial economy, major conflict arose between agricultural landholders and capitalists. Landholders benefit much less from an increase in human capital of their workers than capitalists: human capital raises productivity of workers more in industry than in agriculture because land and human capital are less complementary. Return on land declines as wages of workers rise due to higher education that individuals obtain and educated workers have stronger incentives to migrate to industrial areas than less educated workers. Departure of workers from land is contrary to the interests of landholders. Landowners inhibit educational policies aimed to augment general education. Consequently, land ownership is obstacle to human capital formation, thus factor slowing down industrialization and economic growth. Land ownership is a driving force behind divergence in income per capita with long-run implications even today. So, variations in the distribution of land ownership across countries generated variations in the industrial composition of the economies. This is because the transition to an industrial economy raises the importance of human capital in the production process, reflecting its complementarity with physical capital and technology. GMV confirm the hypothesis that inequality in land distribution has a negative correlation with human capital formation via US state-level evidence for the first decades of the twentieth century. Countries with high land inequality are overtaken in the procedure of industrialization. Education improves the productivity

of industrial labour more than that of agricultural labour. In contrast, economies with low land inequality promote public education, benefit from industrialization and leach into a cursory process of development. In order to address causality they instrument landownership inequality through the interaction between global changes in the relative price of agricultural crops that are associated with economies of scale and variation in climatic characteristics across states. More recent studies focus on literacy rates and the channels through which land ownership, geographic factors and other determinants influence them. For instance, Baten and Hippe (2018) find that the distribution of land ownership is a mechanism behind the correlation of human capital and geographic factors using a European multi-country regional level.

This research aims to shed light on land ownership as driving force of the divergence in educational attainment and economic development across regions in modern Greece. In line with GMV theoretical predictions, we show that inequality in land ownership has a negative correlation with human capital formation. Specifically, this paper examines the validity of GMV theory for the first time in a regional Greek sample of provinces around 1900. This differs from most literature, which focuses on industrial economies because Greece was at the early stages of the transition to the industrial era during early 20th century. In order to examine empirically the above theoretical hypothesis we build an entirely novel data set with Greek historical data, regarding literacy rates (male, female, total, refugees) at the provincial level based on the 1928 Census, which are the first ones to include an extended educational attainment coverage at the provincial levels. Generally, we use more spatially disaggregated data than the related literature. Our dependent variables are literacy rates, and the rest of the explanatory variables are similar to e.g. Baten and Hippe (2018). According to historical sources large landholdings in Greece were concentrated in areas where chiftliks (large manors farmed for the market) existed, while these sources focus on the different forms of land tenure in various regions of Greece. We estimate regressions with literacy rates as dependent variables and land ownership, geographic factors and other controls as independent. To further strengthen our results (i.e. with respect to endogeneity), we also provide treatment effect estimations to address endogeneity. Moreover, we construct some graphs in

order to sketch the idea that these historical data affect long-run development. Our research suggests that inequality in the distribution of land ownership adversely affected human capital formation, and thus the process of the transition from an agricultural to an industrial economy.

The rest of the paper is structured as follows: in Section 2 we revisit the theoretical framework on literacy, land ownership and geography, in Section 3 we describe the data and the methodology used in this study, in Section 4 we present our empirical results, in Section 5 we present the effect of our data on the long-run development and Section 6 concludes.

2. Literature Review

There are quite a few studies which focus on human capital formation and its determinants. Recent work focuses on literacy rates and the channels through which land ownership and geographic factors influence them. As mentioned above, Galor et al. (2009) argues that land inequality has negatively affected human capital promoting educational institutions (e.g. public schooling, child labour regulations) in the United States, using variation in the distribution of land ownership and educational expenditure, across states and over time, for the period 1900-1940. Deininger and Squire (1998) use cross-country data and panel data for 103 countries in the period 1960-1990 and find a negative relationship between initial inequality in land distribution and long-run growth. Land ownership is negatively correlated with education, while high levels of education lead to more investment.

Erickson (2004) uses two new measures of land ownership that capture the level of landlessness within the countries. The first one is land inequality Gini and the second is the agricultural population divided by the total number of holdings. They do not find a significant relationship between land ownership, institutions and development but they support that high land ownership across agricultural populations leads to low levels of education. However, their results contend that land ownership does not impinge on the relationship between geographic endowments and income inequality, which is a very surprising inference because previous studies support the opposite. Easterly (2007) confirms the idea that inequality affects

development, schooling and institutions using cross-country data for developing countries with different agricultural endowments- specifically the land suitability for wheat to that suitability for sugarcane. High inequality in land ownership constitutes the most important drawback to welfare, high quality institutions and high schooling.

A panel data analysis of Wegenast (2009) suggests that the agricultural production system in Asia and Latin America constitutes a major factor of different educational outcomes. Countries with higher agricultural plantations provide less broadly based educational policies in contrast to countries organized around family farming. The panel data analysis and OLS regressions conclude that exports of plantation crops, as proxy for the political strength of the agrarian elite, reduce secondary education attainment levels and government's investments in secondary schooling but these same exports are associated with higher tertiary education levels. Other agricultural products such as cereals, animals or forestry do not affect the different types of educational categories.

Baten and Hippe (2018) examine the relationship between geographic factors and literacy in more than 300 European regions in the 19th century. The indicator of numeracy is the ABCC index which is a proxy for basic numerical skills. The main geographic variables are soil suitability (cereal, pasture, potato, sugar) temperature, precipitation, land size, altitude and ruggedness. They argue that human capital is negatively correlated with land ownership, which is also related to geographic factors. Specifically, capitalists benefit more from an increase in human capital of their workers than landholders. For this reason landowners do not promote policies that promote education. Consequently, land ownership plays an important role in human capital formation, industrialization and economic growth.

Moreover, Adamopoulos and Restuccia(2018) argue that low agricultural productivity in poor countries is not due to poor land endowments. They examine the role of geography and land quality for differences in agricultural productivity using micro-geography data for 162 countries from Global Agro-Ecological Zones (GAEZ) and spatial analysis and try to answer if these differences exist because of geography or economic choices. If all farmers worldwide use the same productivity procedures,

then land quality is not a restriction for poor countries but the issue is what crops are produced, where they are produced within the country and how efficiently they are produced in each cell of land.

Some papers analyze this negative relationship for specific countries. Specifically, Beltrán Tapia and Martínez-Galarraga (2018) confirm the negative relationship between the fraction of farm laborers and literacy rates (female and male) by using a large historical Spanish dataset in mid-19th century, i.e. before industrialization. They examine if demand or supply factors are responsible for this negative relationship and find that in places where there are landowning elites that obstruct education development, demand factors are also crucial. For societies with high land ownership, the economic situation of rural poor is uncertain and there are only few investments on education because of the budget constraints of large segments of the population. In rural areas large landowners obstruct public support of education but in urban areas this does not happen because of the existence of other elites. The example of Prussia is examined by Cinnirella and Hornung (2016), using a unique country-level census database. They examine the negative relationship between landownership concentration and primary education considering agricultural features, demand and supply factors in 19th century. In order to deal with endogeneity, they use an instrumental variable (exogenous variation in farm size because of differences in soil texture), country and time fixed effects, and confirm the negative effect of landownership concentration on education and suggest that the effect is indeed causal and weakens over time. Regions with poorer soil quality have a lower population density, a weaker demand for land and are characterized by higher average farm sizes.

Other historical studies yield insignificant and sometimes contradictory results. Goni (2018) supports the same idea that high landownership concentration has a significant negative effect on education by using a dataset on 1387 School Boards and 32 counties in the period 1871-1899 in England. This effect is significant only for changes that began after the Industrial Revolution in England. High land ownership reduces the ratio of state to private schools, the number and salaries of teachers and the facilities per pupil and as a consequence the attribution of children declines. He

argues that the relationship between land ownership and state education is causal and uses two instrumental variables: soil texture and the massive redistribution of land after the Norman conquest of England in 1066. The estimations with these instrumental variables provide two different results: first-stage estimations provide a strong persistence in land ownership over eight centuries, from 1066 to the late-nineteenth century and second-stage estimations support that the effect of landownership concentration on state education is causal. The consequences of the negative effects of land ownership are confined in areas that landowners are political figures that promote land elites.

Overall, some studies support the significant impact of land ownership, geography and other factors on human capital formation whereas some others do not. However, there is not a study using Greek data, which investigates the historical determinants of human capital formation considering the effect of land ownership in this process. We are based on the literatures of human capital formation, geography and land ownership, by studying a cross-section of all provinces in Greece in 1928, i.e. we use the earliest available data on educational attainment before the establishment of compulsory primary education in Greece. Finally, we construct data which allow us to examine the impact of land ownership on educational outcomes before the implementation of a comprehensive land reform which started in the mid-1920s and redistributed large land holdings to landless peasants.

3. Data and Methodology

3.1 Data

The empirical analysis aims to test the hypothesis that land ownership and geographic factors have an impact on human capital formation in a cross-section of Greek provinces around 1900. We examine the validity of GMV theory (Galor et al., 2009) in a regional Greek sample for the first time. This differs from much of the literature because Greece was at the early stages of the transition to the industrial era during the period studied. We build an entirely novel data set with Greek historical data, regarding literacy rates (the dependent variable) at the provincial level based on the Censi 1928 of the Hellenic Statistical Authority. Our sample consists of 142

provinces in 1928. We use this period because it is the first one including an extended coverage at the provincial level around 1900 (Galor et al., 2009; Cinnirella and Hornung, 2016; use also Census data for literacy rates). We use male, female, total and refugee literacy rates. We use refugees to control for possible human capital externalities from the refugees to the natives and make sure that the estimated effect of land inequality and geography on literacy rates does not include the impact of refugee literacy. Moreover, we construct maps that illustrate the geographical distribution of literacy rates in Greek provinces in 1928. We combine the data of literacy rates from the Hellenic Statistical Authority with the coordinates of Greek provinces which are available in Geodata, which is a Greek geospatial database.

The first three maps present male, female and total literacy rates in Greek provinces in 1928 for native people (refugees not included). As expected, province "Attica" has one of the highest literacy rates (male, female, total). The lowest literacy rates are detected in Western Thrace and Epirus, and the rest highest literacy rates are detected in Peloponnese, Cyclades and Ionian Islands regions.

(Insert here maps 1, 2 and 3)

Furthermore, we create maps for Greek provinces in 1928 including the total population (native and refugees) and maps for the distribution of literacy rates of refugees (male, female, total) in 1928.

For total literacy rates (refugees included) in 1928, "Attica" in Central Greece has again one of the highest literacy rates. Territories with high literacy rates are Peloponnese, Ionian Islands and Cyclades and regions with low literacy rates are Western Thrace and Epirus.

(Insert here maps 4, 5 and 6)

The distribution of refugee literacy rates is very different from the distribution of native literacy rates. High literacy rates are detected in Crete and Cyclades.

(Insert here maps 7, 8 and 9)

The main control variable of interest is land ownership which captures the part of land owned by large landholders for the whole Greece based on relevant historical sources. The assumption is that land ownership is one of the main driving forces of

the divergence in terms of human capital accumulation between Greek regions. This in turn is expected to have serious long-run implications for the spatial distribution of income per capita across regions in modern Greece. Galor et al. (2009) and Baten and Hippe (2018) use quantitative Census data for land ownership. Baten and Hippe (2018) calculate the share of large holdings by dividing the total area of holdings larger than 50 ha by the total area of all holdings. For Greece this is impossible because only few quantitative data exist but at higher regional levels (i.e. territory level) and it is not sure if these data are valid, because there is no a credible definition of large land ownership in 1928. As a result we have studied economic history sources in order to construct a qualitative variable that measures land ownership. According to these sources (Vergopoulos, 1975; MCGrew, 1985; Petmezas, 2003; Petmezas and Kostis, 2006; Kontogiorgi, 2006; Petmezas, 2012) large landholdings in Greece were concentrated in areas where *chiftliks (*large manors farmed for the market) existed, while these sources focus on the different forms of land tenure in different regions of Greece. *Chiftliks* first appeared in 16th century and became widespread in 18th century because of the weakening of the central political authority of the Ottoman Empire and the large expansion in international trade. According to Ottoman law there is no private ownership and theoretically all regions of the Empire belong to the Sultan. The *chiftlik* system is associated with the development of local administrators, beys and aghas. In 1858 by Land Code aghas and beys had the right to full ownership of estates which they had been cultivating for ten years. By the end of the nineteenth century the land tenure system divided the land into two categories: the *chiftliks*, the large freehold estates, and the head villages, Christian villages which were under the supervision of the central government and not of a local bey or agha. *Chiftliks* consisting of twenty to thirty families of tenant farmers—rarely up to fifty—were located in the fertile lowlands close to main roads and were conferred as the estates of a landlord. Head villages were mainly mountain villages or at best situated in the foothills, where there were better hygienic conditions and cultivation of land was easier. They are characterized by local autonomy since the major intervention of the state in village life was in the field of tax-farming. Another important social group was that of the transhumant shepherds. They had lots of sheep and goats, spent the summer on the highland pastures on the mountain ranges and migrated to the valleys,

where they spent the winter on rented pastures. They offered more to the existing money economy than the farmers. The *chiftlik* system developed as a result of attempts to commercialize Ottoman agriculture and deal with the demands from Europe for products such as wheat, cotton and wool. *Chiftlik* was a combination of agriculture and pasture and not only a large agricultural venture. The landowner ensured his annual income by the combination of winter grains (wheat, barley, etc.), spring crops and pastureland rented by shepherds. During and after the Balkan Wars there is a defective transition from the Ottoman system of land rights to one based on private ownership principles. The *chiftliks* started getting under the control of the Greek state, a state that depended on guarantees. On the grounds that the majority of large estates had remained in the hands of Muslim landlords, Greek government was afraid of expropriating this land because of the diplomatic complications with Turkey. During the first decade of Greek administration, the position of the chiftlik planters impaired roughly. In addition after liberation, the state also rented the abandoned *chiftliks* of Muslim landlords who had migrated. Regional differences in land ownership can generally be viewed as very stable over time according to these sources until approximately 1922, i.e. the arrival of mass refugee inflows. This is why mass expropriations of large landholdings started soon afterwards in order to distribute this land to the refugees. In this study land ownership is a dummy variable taking values 0 and 1, where 0 corresponds to regions with low share of large land holdings and 1 corresponds to regions with high share of large land holdings. According to the theoretical framework (see literature review) we expect the sign of this variable to be negative. In this respect, we construct a map for the geographical distribution of land ownership in Greek provinces referring to the period before the expropriation of large properties by the state.

(Insert here map 10)

Provinces with high share of large land holdings are found in Central Greece/Euboea (e.g. provinces of “Attikis, “Thivon”, “Istiaias”), Thessaly, Epirus, Macedonia and in province “Didymoteichou” in Western Thrace. Provinces characterized by low share of large land holdings are mainly located in Peloponnese, Crete, Ionian Islands, Cyclades, Aegean Islands and Western Thrace.

We enrich the estimated specifications with control variables consistent with the literature (e.g. León, 2004; Clark and Gray, 2014; Baten and Hippe, 2018; Adamopoulos and Restuccia, 2018; Goñi, 2022).

First, we use geographic factors as the channel behind the correlation of land ownership with literacy rates. The geographic factors are soil yield (wheat, potato, sugar, olive, rice, and pasture), temperature, altitude, ruggedness and precipitation. Geographic factors are very stable over time. We use soil yield factors instead of soil suitability factors, because the latter are highly correlated with each other, and we cannot use them all together in our estimates. Altitude (the elevation above sea level in meters) is collected from the Census of 1951 from Hellenic Statistical Authority and ruggedness is the standard deviation of altitude. The remaining eight geographic variables are obtained using GIS techniques. Specifically, we have downloaded Greek maps that contain the data on each variable from the database IPUMS Terra (Integrated Population and Environmental Data). Then we have taken the coordinates of Greek municipalities from Geodata, a Greek geospatial database and combined maps and coordinates using the QGIS software (open-source geographic information system) in order to extract the geographic data at municipal level. Finally, in order to collect these geographic data at the provincial level we have found the municipalities belonging to every province and taken the median of the data of these municipalities. Following the relevant literature (Wegenast, 2009; Beltrán Tapia and Martínez-Galarraga, 2018; Baten and Hippe, 2018) we expect the sign of the variables to be negative for some factors and positive for others.

Finally, we use four additional control variables. These are taken from the Census of the Hellenic Statistical Authority. The first one is the land area for Greek provinces. The second is the population density which is calculated by the division of population with land area. The third is the urbanization rate, a proxy for regions with large population to control for urbanization effects on literacy (Baten and Hippe, 2018 also use this variable). Specifically, it is a ratio calculated by the division of ten thousand with the population in 1928[1]. The last control variable is the ratio of Bulgarians/Turks in 1920 which is the division of the number of Bulgarians/Turks with the population in 1920.

Table 1 presents the description of the data, the period and the data sources.

(Insert here Table 1)

In addition, summary statistics of the variables of Greek provinces in 1928 are reported in Table 2. The literacy rates take values between zero and one because they are ratios.

(Insert here Table 2)

Table 3 provides the correlations between geographic factors in Greek provinces. Variables with high correlation, such as wheat and sugar yield, are not included together in the estimations to avoid multicollinearity.

(Insert here Table 3)

3.2 Methodology

We pursue empirical analysis considering regional (department) fixed effects because the regressions are likely to suffer from endogeneity in the form of omitted variable bias. If there are omitted variables, and these variables are correlated with the variables in the model, then fixed effects models provide a means for controlling for this bias. The idea is that whatever effects the omitted variables have on the subject at one time, they will also have the same effect at a later time; hence their effects will be constant or fixed over time. In addition, we allow for clustered standard errors at the regional level (department and prefecture levels) to account for spatial error dependence. For example, clustered standard errors at the department levels assume that errors of the spatial units are correlated within departments, but independent between departments.

We also use treatment effects (regression adjustment and nearest-neighbor matching) to address endogeneity. We use the average treatment effects (ATEs) to estimate treatment effects from observed data and the potential-outcome means (POMs) to estimate the distribution of individual-level treatment effects. A potential-

[1]This is the threshold sets by the Hellenic Statistical Authority in the Census to define that a settlement constitutes an urban center.

outcome model specifies the potential outcomes that each individual would obtain under each treatment level, the treatment assignment process, and the dependence of the potential outcomes on the treatment assignment process. The term treatment effect is defined as the average causal effect of a variable (land ownership) on an outcome variable (literacy rates). A treatment effect is the change in an outcome caused by a subject, often an individual, getting one treatment instead of another. The defining characteristic of observational data is that treatment status is not randomized, and it is not possible to observe a specific subject having received the treatment and having not received the treatment. Moreover, that implies that the outcome and treatment are not necessarily independent. A classic solution to this problem is to randomize the treatment. High costs or ethical issues rule out this solution in many observational datasets. The treatment effect estimations allow us to estimate the efficacy of treatments using observational data. All the estimators require some form of the following three assumptions. The independent and identically distributed (i.i.d.) sampling assumption, ensures that the outcome and treatment status of each individual are unrelated to the outcome and treatment status of all the other individuals in the population. The conditional-independence (CI) assumption means once we control for all observable variables, the potential outcomes are independent of treatment assignment. The third assumption is the overlap assumption. This assumption ensures that each individual could receive any treatment level and more formally this assumption states that each individual have a positive probability of receiving treatment.

The regression adjustment (RA) method extends the idea of using sample means to estimate treatment effects by using a regression model to predict potential outcomes adjusted for covariates. Regression adjustment fits separate regressions for each treatment level and uses averages of the predicted outcomes over all the data to estimate the POMs. The estimated ATEs are differences in the estimated POMs.

Matching estimators use an average of the outcomes of the nearest individuals to impute the missing potential outcome for each sampled individual. Matching estimators are based on the idea of comparing the outcomes of subjects that are as similar as possible with the sole exception of their treatment status. Nearest-neighbor matching (NNM) is accomplished by calculating the “distance” between pairs of

observations with regard to a set of covariates and then "matching" each subject to comparable observations that are closest to it. NNM is nonparametric in that no explicit functional form for either the outcome model or the treatment model is specified. This flexibility comes at a price; the estimator needs more data to get to the true value than an estimator that imposes a functional form. More formally, the NNM estimator converges to the true value at a rate slower than the parametric rate, which is the square root of the sample size, when matching on more than one continuous covariate.

4. Empirical Results

To better understand the relationship between land ownership, geographic factors and literacy rates, we estimate the following regression model:

$$LT_{ig} = \beta_0 + \beta_1 landown + \beta_2 popdens + \beta_3 land + \beta_4 urban + \beta_5 geography + \beta_6 foreign + \mu_i + \varepsilon_i, \quad (1)$$

where *i* denotes the specific region, *g* denotes the gender, *LT* denotes the different literacy rates, *landown* is a dummy variable for land ownership, *popdens* is the population density, *land* is the land area, *urban* is the urbanization rate, *geography* denotes the different geographic factors, *foreign* denotes the ratio of Bulgarians/Turks in 1920, μ_i represents regional fixed effects and ε comprises the non-observed influences on the *LT*.

We start with fixed effect estimation techniques. Table 11 provides fixed effect estimations of literacy rates on land ownership for Greek provinces in 1928. We use fixed effects at department level and cluster standard errors at department level (Columns 1, 2, 3) and prefecture level (Columns 4, 5, 6). Land ownership is negatively correlated with literacy rates. Pop density and urbanization are strongly and positively correlated with literacy rates. Pasture yield and precipitation have strong negative relationships with literacy rates.

(Insert here Table 4)

The majority of the related studies start empirical analysis with OLS estimations in order to interpret the negative relationship between land ownership and literacy rates (Easterly, 2007; Wegenast, 2009; Cinnirella and Hornung, 2016; Baten and Hippe, 2018). The next Table (Table 5) provides OLS estimations of male,

female and total literacy rates of Greek provinces in 1928 with fixed effects at department level. Cluster standard errors are at department level (Columns 1, 2, 3) and prefecture level (Columns 4, 5, 6). Regions with higher land ownership tend to have lower male, female and total literacy rates in 1928. Population density and urbanization continue to have the same signs (positive) and statistical significance. Precipitation and pasture yield have significant negative impacts on literacy rates. The result that geographic factors have an impact on literacy rates is supported by the related literature, but the signs of the coefficients are different. For example, Baten and Hippe (2018) support that pasture yield has a positive impact on literacy and population density has a negative impact on literacy. The signs of the coefficients of these variables are in contrast with our results. The inference that we can draw from these estimations is that the signs of the main variables are the same regardless of the estimation technique.

(Insert here Table 5)

For Greek provinces in 1928 except from the native literacy rates we also have data for total literacy rates (native and refugee) and only for refugees. To this end we include in our analysis some Tables with estimations for total literacy rates (native and refugee). For these estimations we have total literacy rate as dependent variable and also use refugee literacy rate as independent variable. For land ownership, population density and urbanization the signs and the statistical significance are the same as the estimations only with native literacy rates but not so strong. Table 6 illustrates fixed effect regressions with male, female and total literacy rates as dependent variables in Greek provinces in 1928. Fixed effects are at department level and cluster standard errors are at department level (Columns 1, 2, 3) and prefecture level (Columns 4, 5, 6). Table 7 (fixed effects at department level) provides OLS estimations. Cluster standard errors are at department level (Columns 1, 2, 3) and prefecture level (Columns 4, 5, 6). Land ownership, pasture yield and precipitation have statistically significant negative relationship with literacy rates. Pop density and urbanization have positive and statistically significant signs.

(Insert here Tables 6 and 7)

Next, we use treatment effects (regression adjustment and nearest-neighbor

matching) to address endogeneity. In these estimations the effect of geographic factors on literacy comes through land ownership. Table 8 presents the regression adjustment model of native male, female and total literacy rates in Greek provinces in 1928. We use the average treatment effect (ATE), and the potential-outcome means with robust standard errors. For all these estimations the dependent outcome variables are male, female and total literacy rates and the independent outcome variables are population density, urbanization rate and the rate of Bulgarians/Turks. The treatment variable is land ownership. Land ownership has negative statistically significant relationships with female and total literacy rates in 1928. For example the strong presence of large land holdings reduces female literacy by 5.6%.

(Insert here Table 8)

Table 9 provides the nearest-neighbor matching model with the average treatment effect for native male, female and total literacy rates for Greek provinces in 1928. We use robust standard errors and bias adjustment. For these estimations the dependent outcome variables are male, female and total literacy rates and the independent outcome variables are population density, urbanization rate and the ratio of Bulgarians/Turks. The treatment variable is land ownership. Landownership is negatively correlated with female and total literacy rates. These negative correlations are also statistically significant at the 1% level of significance for female population and at the 5% level of significance for total population. So, the presence of large land holdings in a province reduces total literacy on average by 4.2%.

(Insert here Table 9)

5. Long-Run Development

The last few decades there is a motivation of researching the evolutionary roots of comparative economic development across regions and countries. Specifically, many researchers focus on the influences of human evolution and the composition of human traits on comparative economic development across societies. Moreover, there is a literature (Nunn, 2009; Caicedo, 2018) that examines the long-term consequences of a historical human capital intervention. Following this literature, we provide an informal inquiry on whether historical literacy rates in 1928 still have an impact on educational attainment (approximately sixty years later (1981)

and economic development about eighty years later (2000)). We use 1981 Census data for male, female and total secondary and tertiary education attainment for provinces which correspond to provinces in 1928. Regarding, economic development we use 2000 Census data for GDP per capita. We use 1981 and 2000 Census data because they are the latest available data at provincial level. Using scatter plots we examine if historical human capital formation influences long-run outcomes. The first three figures illustrate male, female and total secondary education attainment versus male, female and total literacy rates in 1928. The lines are linear trends and depict the positive influence of literacy rates in secondary education attainment. Correlations between literacy in 1928 and secondary education attainment in 1981 for provinces are 0.42, 0.53 and 0.54 for males, females and total population respectively and are statistically significant at the 1% level.

(Insert here Figures 1-3)

Next, we present male, female and total tertiary education attainment versus male, female and total literacy rates in 1928 (Figures 4-6). The lines are again a linear trend, and the figures provide positive influence of literacy rates on tertiary education attainment. Correlations between literacy in 1928 and tertiary education attainment in 1981 for provinces are 0.38, 0.45 and 0.49 for males, females and total population respectively and are statistically significant at the 1% level.

(Insert here Figures 4-6)

The last three figures provide GDP per capita versus male, female and total literacy rates in 1928 (Figures 7-9). The lines are linear trends and depict positive influence between literacy rates and GDP per capita except for figure with male literacy rate (Figure 7). Correlations between literacy in 1928 and GDP per capita in 2000 for provinces are 0.35 and 0.25 for females and total population and are statistically significant at the 1% level of significance for females and at the 5% level of significance for total population. There is no statistically significant correlation between male literacy rates and GDP per capita.

(Insert here Figures 7-9)

We show that human capital transmission across generations is a main driver of economic outcomes. More generally, the findings that we present underscore the

importance of historical institutions and interventions for driving economic growth in the long run.

6. Concluding Remarks

Theory suggests that the concentration of landownership has been a major factor of human capital formation and economic growth (Galor et al., 2009; Baten and Hippe, 2018; Easterly, 2007; Adamopoulos and Restuccia, 2018). The increase in the demand for human capital in the process of industrialization and on the onset of the demographic transition is the main force in the transition from stagnation to growth. As the demand for human capital emerged, differences in the concentration of landownership generated variations in human capital formation, technological changes and the timing of industrialization.

This study examines and analyses the relationship between land ownership and literacy, as proposed by Unified Growth Theory. Empirical results suggest that unequal distribution of land has a substantial negative correlation with literacy. These results were found to be robust across alternative specifications. This negative relationship is based on theoretical grounds. We find a relationship between geographic factors and literacy in Greek provinces around 1900. Pasture yield and precipitation are always negatively correlated with literacy rates in estimations with literacy as dependent variable. We also control for several other explanatory factors (urbanization, population density, land area). Population density has positive, statistically significant and strong relationship with literacy rates for all estimation techniques. Land ownership has a negative, statistically significant relationship with all literacy rates. Urbanization is strongly positively correlated with literacy rates in Greek provinces in 1928.

Although we include in our analysis a large set of geographic variables (soil yield, temperature, altitude, precipitation etc.) the effect of land ownership on literacy rates does not change. We suggest that, in earlier phases of industrialization, the distribution of land has an important negative correlation with the development of literacy. A more equal distribution of land may help to foster educational attainment, economic growth and income distribution. Despite the fact that the agrarian reform dismantling large land holdings took place in the period between 1920 and 1930, the

effects of early land ownership on literacy and development seem to persist in the long run.

References

Adamopoulos, T., and Restuccia, D. (2022). Geography and agricultural productivity: Cross-country evidence from micro plot-level data. *The Review of Economic Studies,* 89(4):1629-1653.

Baten, J., and Hippe, R. (2018). Geography, land inequality and regional numeracy in Europe in historical perspective. *Journal of Economic Growth*, 23:79-109.

Cinnirella, F., and Hornung, E. (2016). Landownership concentration and the expansion of education. *Journal of Development Economics*, 121:135–152.

Deininger, K., and Squire, L. (1998). New ways of looking at old issues: inequality and growth. *Journal of Development Economics*, 57(2):259–287.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2):755–776.

Erickson, L. A. (2004). Dimensions of land inequality and economic development. *IMF working paper No. 04/158*.

Galor, O., Moav, O., and Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *The Review of Economic Studies*, 76(1):143–179.

Goñi, M. (2022). Landed elites and education provision in England: evidence from school boards, 1871-99. *Journal of Economic Growth,* 1-47.

Gray, R., and Clark, G. (2014). Geography is not destiny: geography, institutions and literacy in England, 1837–1863. Oxford Economic Papers, 66(4):1042–1069.

Hellenic Statistical Authority, (1862). *Census of 1861.* Athens, Greece.

Hellenic Statistical Authority*,* (1871). *Census of 1870.* Athens, Greece.

Hellenic Statistical Authority, (1884). *Tables of provinces of Epirus and Thessalia in the census 1881*. Athens, Greece.

Hellenic Statistical Authority, (1929). *Census of 1928.* Athens, Greece.

Hellenic Statistical Authority, (1961). *Census of 1951.* Athens, Greece.

Hellenic Statistical Authority, (1915). *Enumeration of the new provinces of Greece inhabitants of the year 1913.* Athens, Greece.

Kontogiorgi, E. (2006). Population exchange in Greek Macedonia: The rural settlement of refugees 1922–1930. *Oxford University Press on Demand*.

Kostis, K., and Petmezas, S. (2006). The development of Greek Economy in 19th century. *Alexandria Editions,* Athens, (in Greek).

León, A. (2004). The effect of education on fertility: Evidence from compulsory schooling laws. *Working Paper, University of Pittsburgh.*

McGrew, W. W. (1985). Land and revolution in modern Greece, 1800–1881: the transition in the tenure and exploitation of land from Ottoman rule to independence. *Kent State University Press*.

Michalopoulos, S., and Papaioannou, E. (2017). The Long Economic and Political Shadow of History: Europe and the Americas. *CEPR Working Paper.*

Petmezas, S. (2003). The Greek agricultural economy in the 19th century. *University of Crete Editions*, Crete, (in Greek).

Petmezas, S. (2012). Introduction in the history of Greek agricultural economy in the interwar period. *Alexandreia Editions*, Athens, (in Greek).

Tapia, F. J. B., and Martinez-Galarraga, J. (2018). Inequality and education in pre-industrial economies: evidence from Spain. *Explorations in Economic History*, 69:81-101.

Vergopoulos, K. (1975). The agricultural issue in Greece-the problem of social integration of agriculture*. Exantas editions*, Athens, (in Greek).

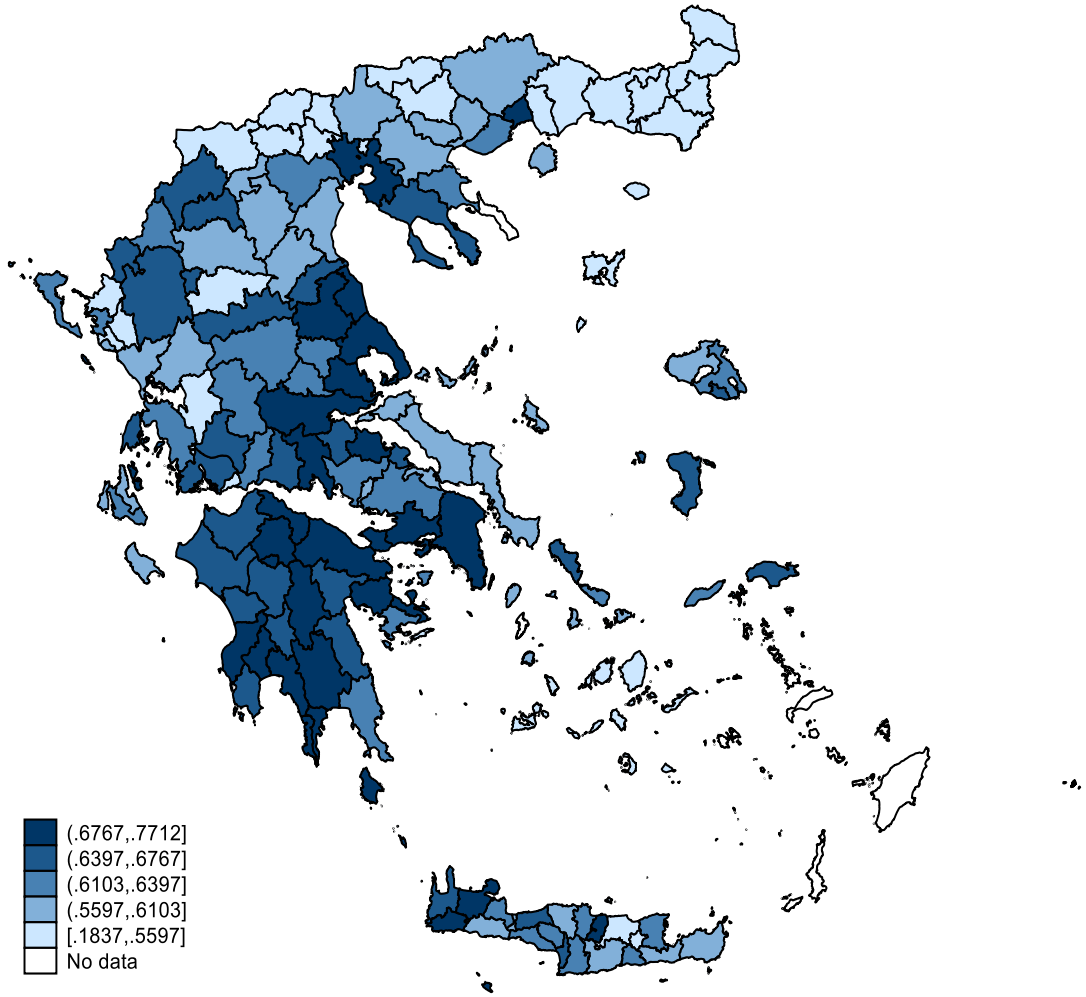
Wegenast, T. (2009). The legacy of landlords: educational distribution and development in a comparative perspective. *Zeitschrift für vergleichende Politikwissenschaft*, 3(1):81–107.

Appendix

Maps

Map 1

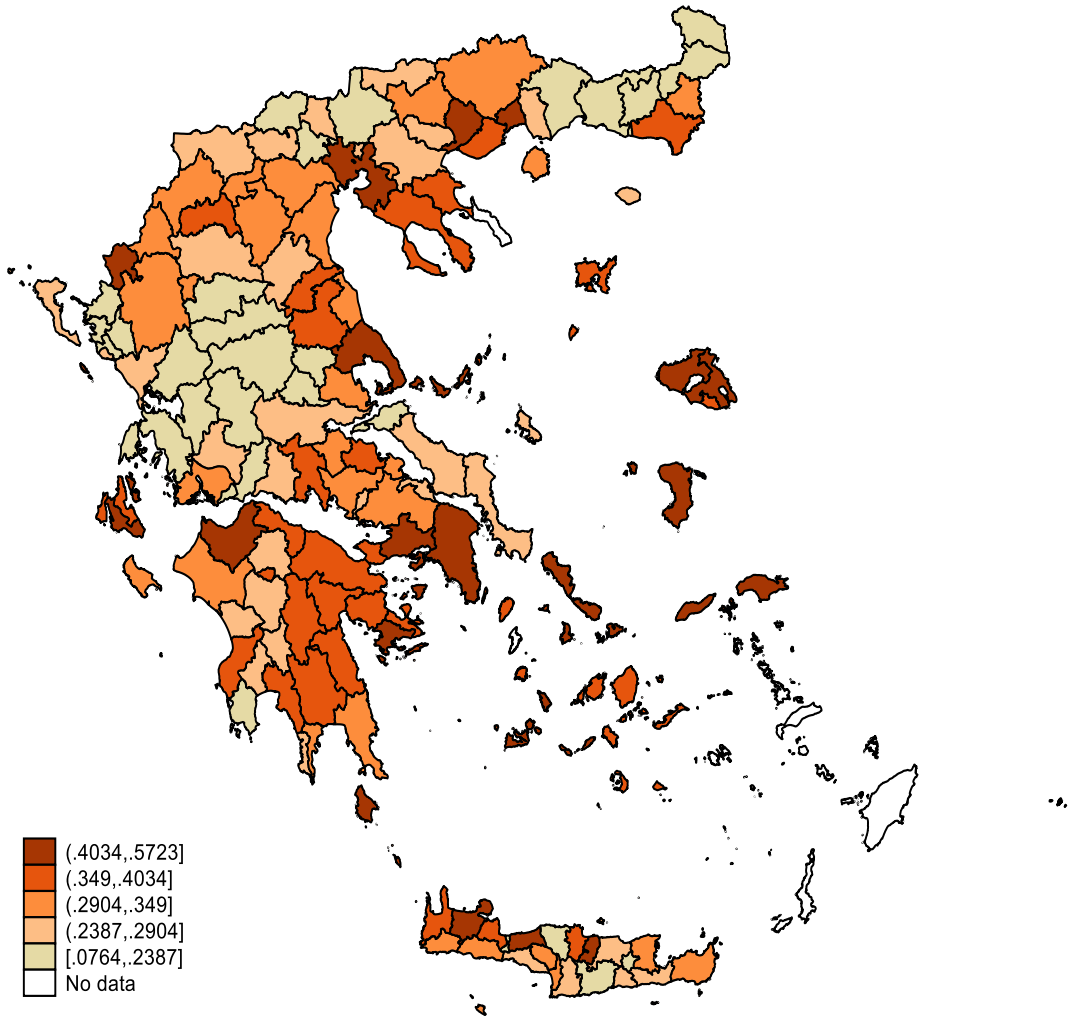
Native Male Literacy Rate 1928



Lowest literacy rates: Sapon(Western Thrace), Komotinis(Western Thrace), Orestiados(Western Thrace), Xanthis(Western Thrace). **Highest literacy rates:** Aigialeias(Peloponnese), Kythiron(Peloponnese), Attikis(Central Greece/Euboea), Paxon(Ionian Islands).

Map 2

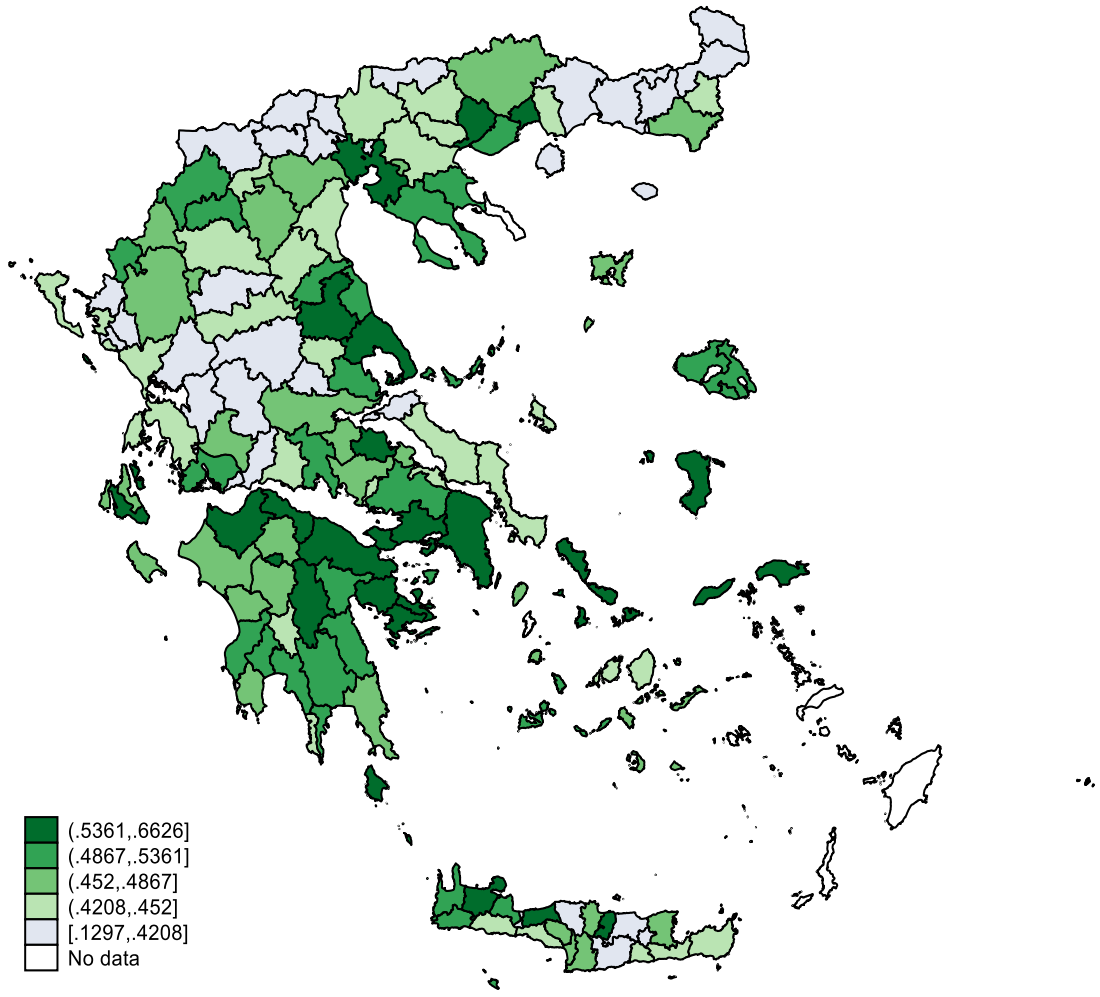
Native Female Literacy Rate 1928



Lowest literacy rates: Sapon(Western Thrace), Orestiados(Western Thrace), Margaritiou(Epirus), Paramythias(Epirus). **Highest literacy rates:** Ydras(Peloponnese), Attikis(Central Greece/Euboea), Tinou(Cyclades), Zichnis/Fylidos(Macedonia).

Map 3

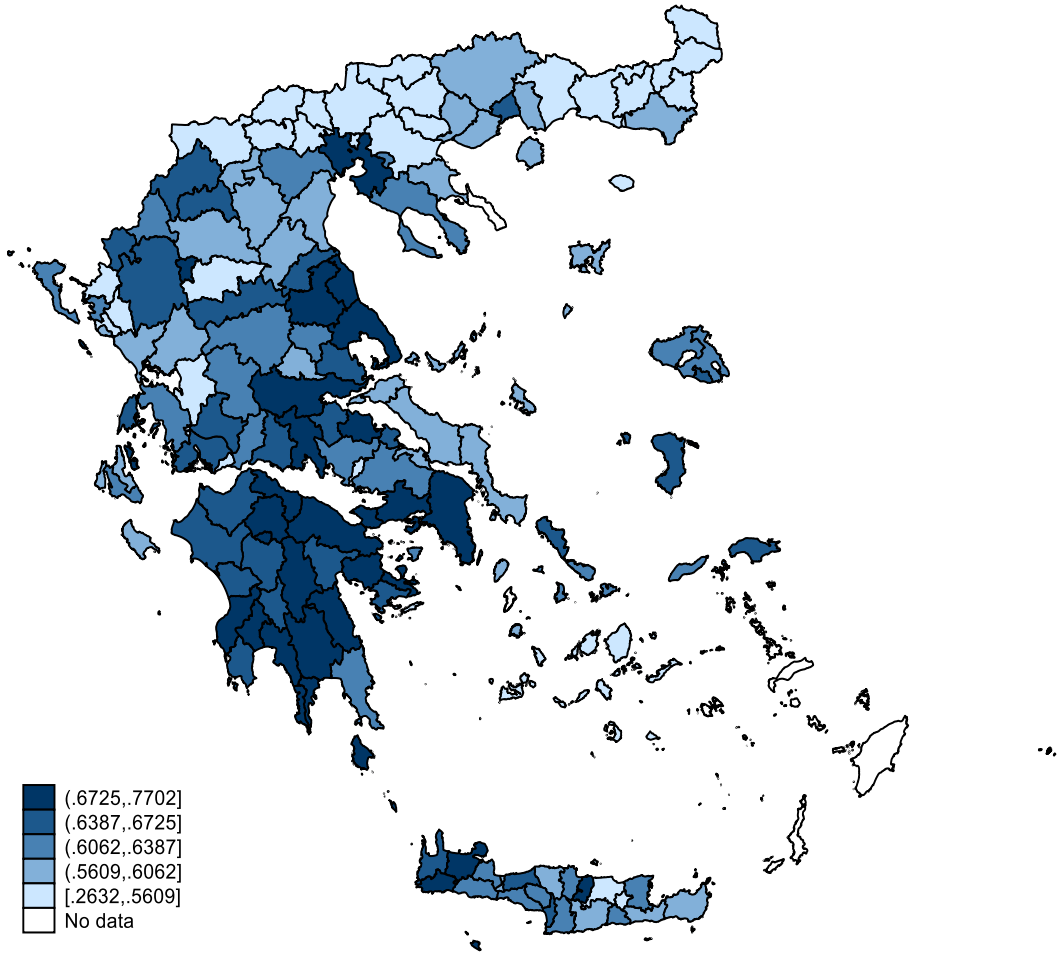
Native Total Literacy Rate 1928



Lowest literary rates: Sapon(Western Thrace), Orestiados(Western Thrace), Komotinis(Western Thrace), Margaritiou(Epirus). **Highest literacy rates:** Tinou(Cyclades), Ithakis(Ionian Islands), Paxon(Ionian Islands), Attikis(Central Greece/Euboea).

Map 4

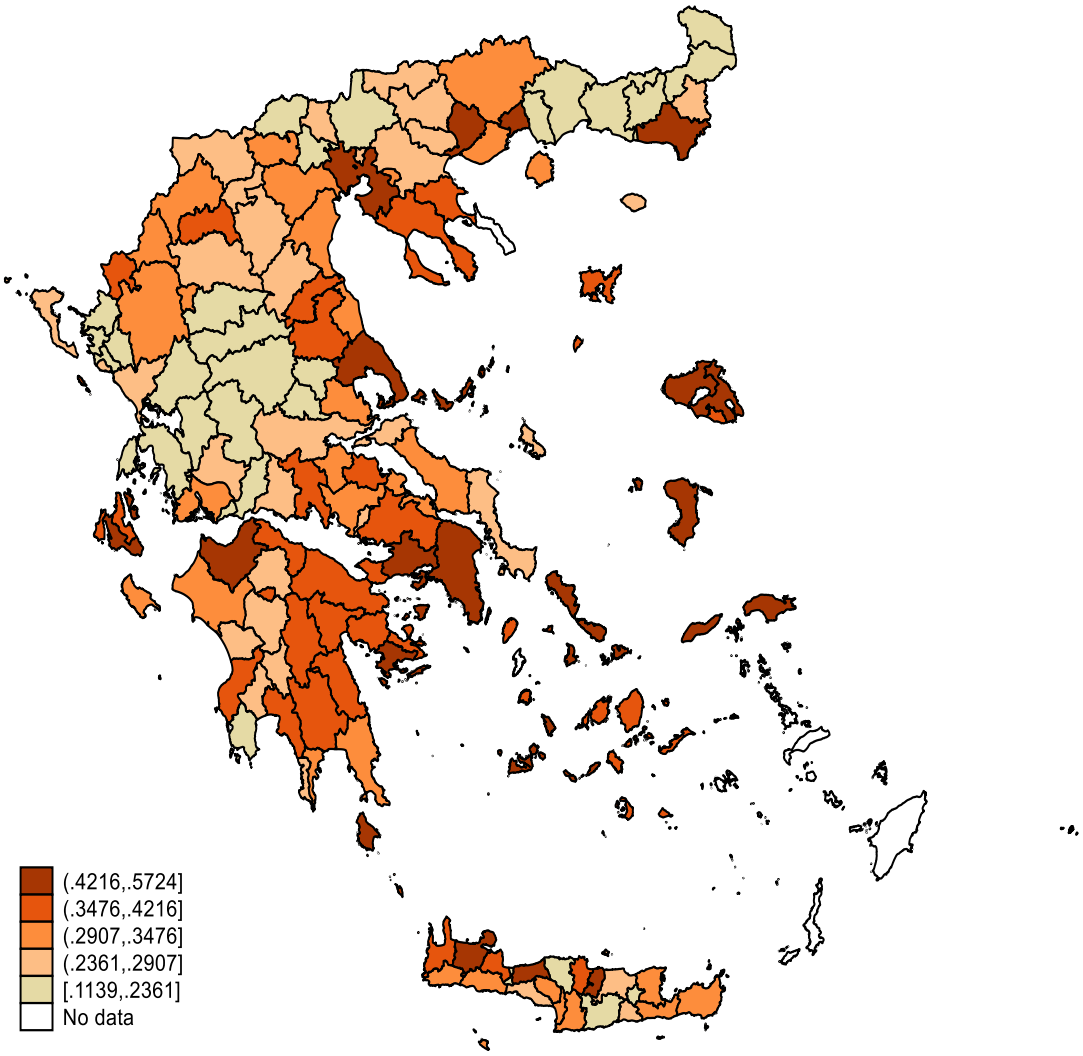
Male Literacy Rate 1928



Lowest literacy rates: Sapon(Western Thrace), Margaritiou(Epirus), Orestiados(Western Thrace), Didymoteichou(Western Thrace). **Highest literacy rates:** Mantineias(Peloponnese), Attikis(Central Greece/Euboea), Kythiron(Peloponnese), Paxon(Ionian Islands).

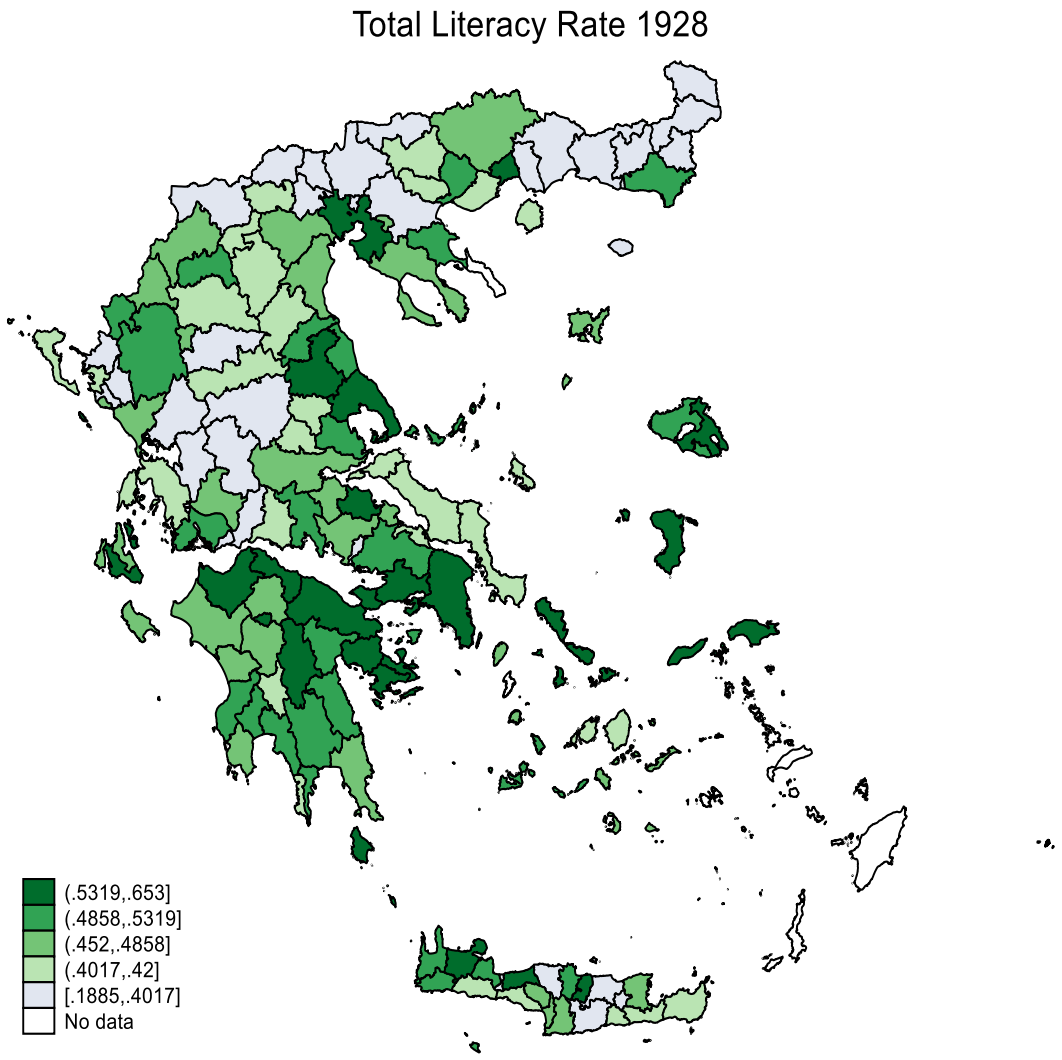
Map 5

Female Literacy Rate 1928



Lowest literacy rates: Sapon(Western Thrace), Paramythias(Epirus), Margaritiou(Epirus), Valtou(Central Greece/Euboea). **Highest literacy rates:** Ithakis(Ionian Islands), Ydras(Peloponnese), Attikis(Central Greece/Euboea), Tinou(Cyclades).

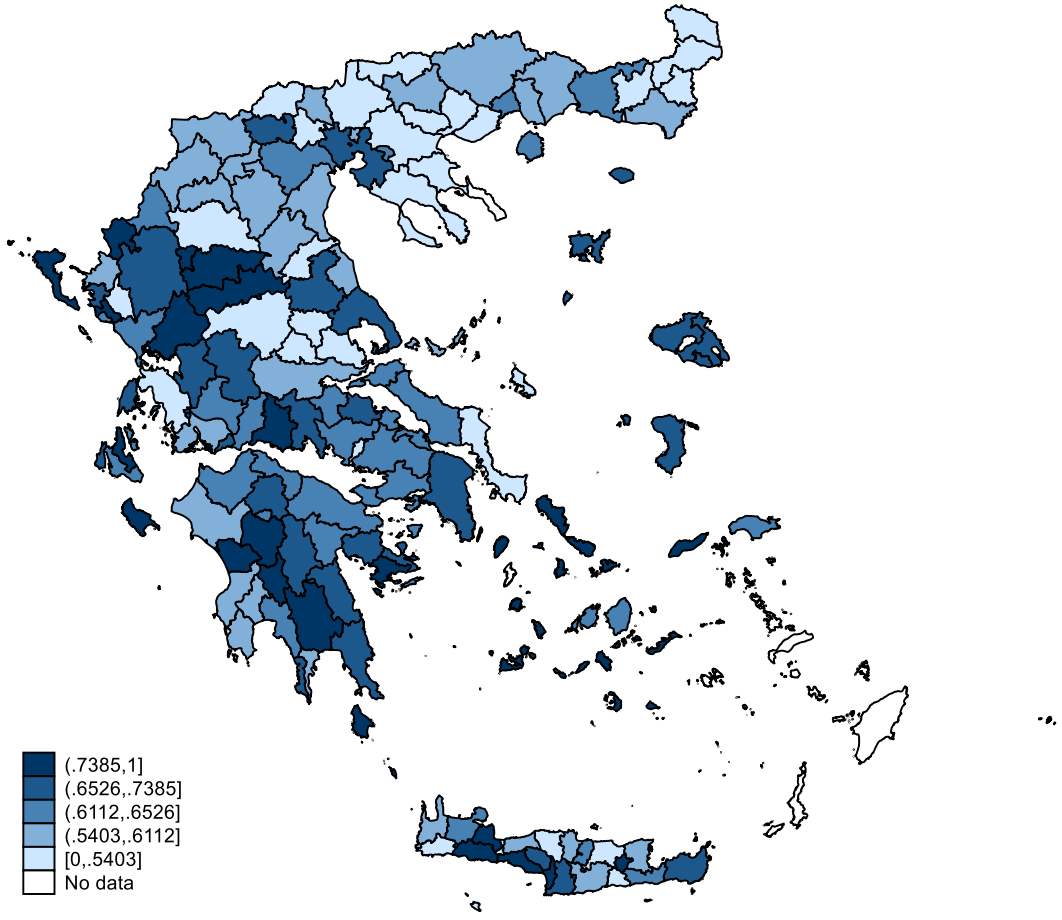
Map 6



Lowest literacy rates: Sapon(Western Thrace), Margaritiou(Epirus), Filiaton(Epirus), Orestiados(Western Thrace). **Highest literacy rates:** Tinou(Cyclades), Ithakis(Ionian Islands), Paxon(Ionian Islands), Attikis(Central Greece/Euboea).

Map 7

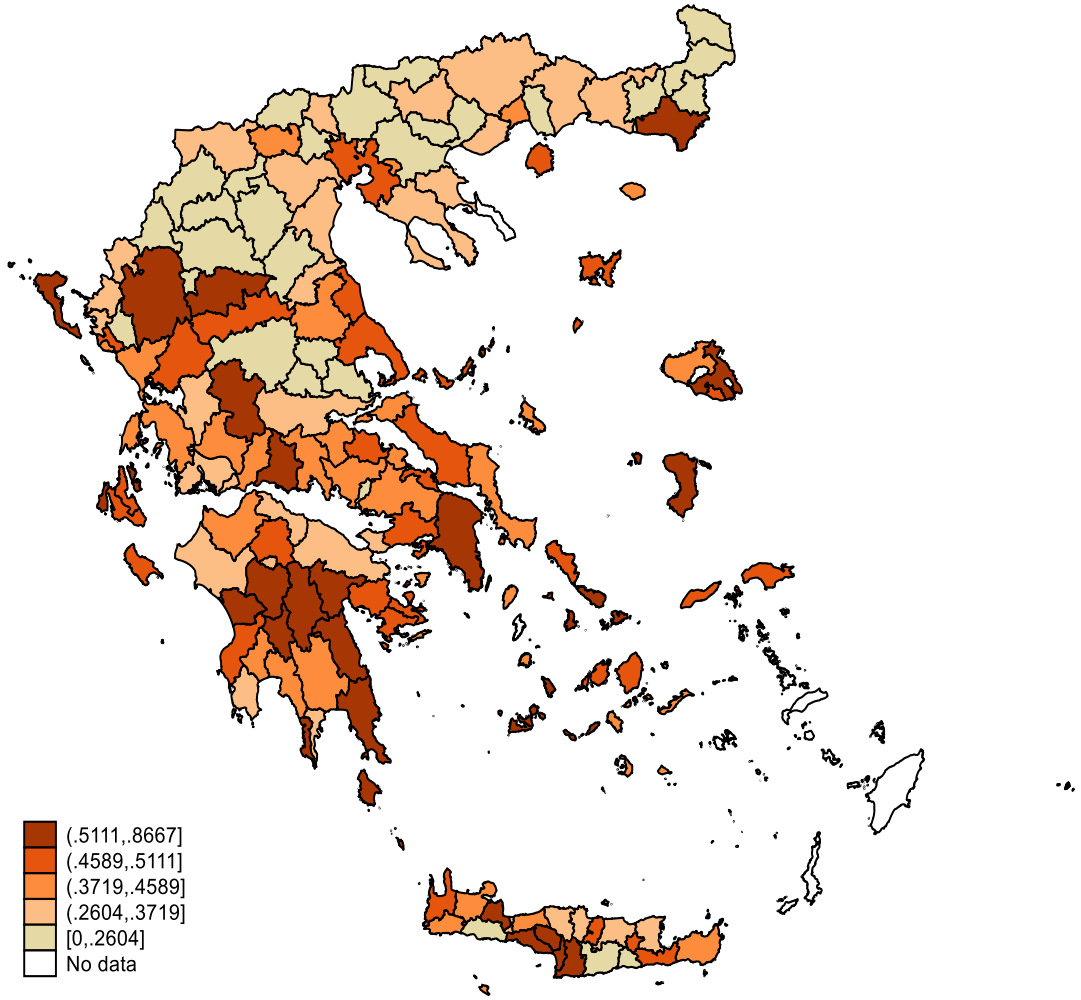
Male Refugee Literacy Rate 1928



Lowest literacy rates: Paxon (Ionian Islands), Didymoteichou(Western Thrace), Domokou(Central Greece/Euboea), Farsalon(Thessaly). **Highest literacy rates:** Syrou(Cyclades), Milou(Cyclades), Metsovou(Epirus), Sfakion(Crete).

Map 8

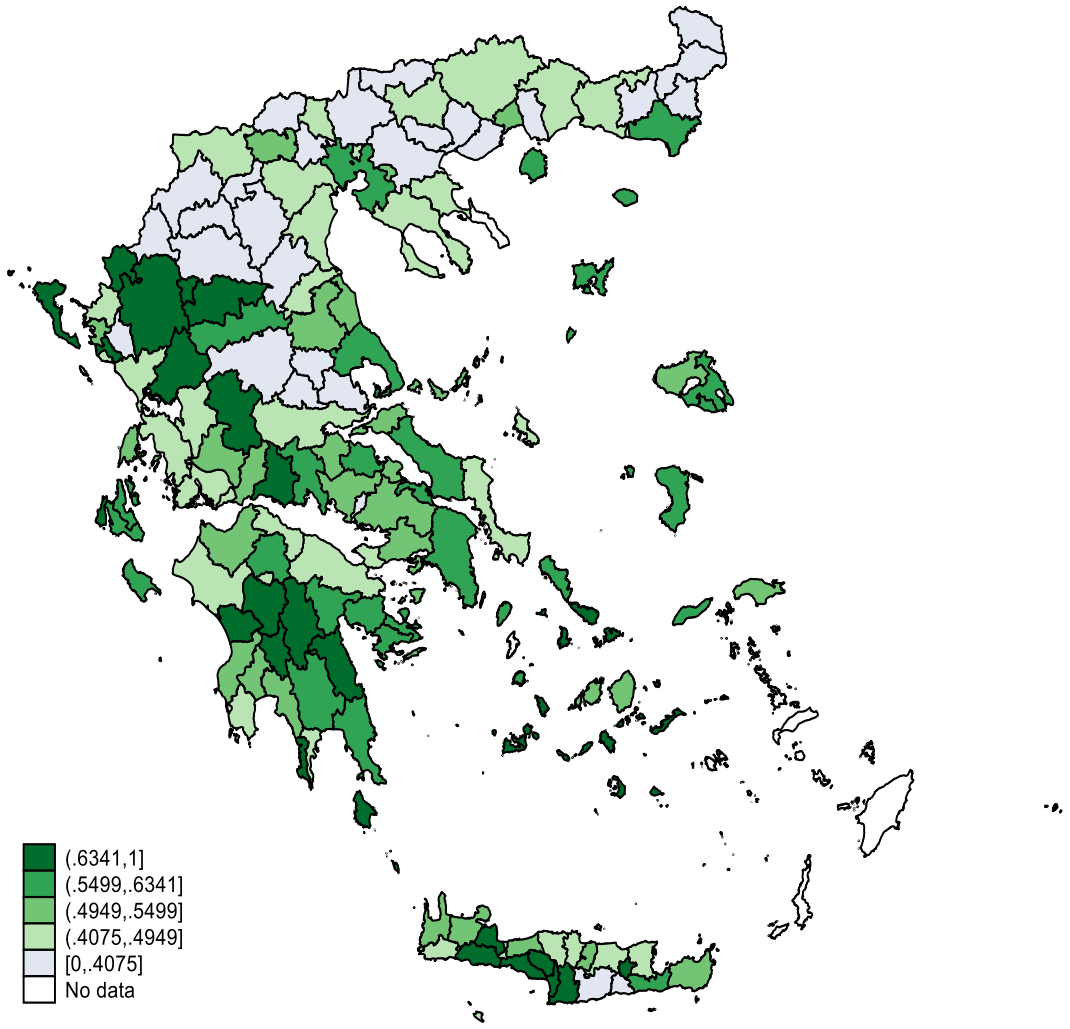
Female Refugee Literacy Rate



Lowest literacy rates: Domokou(Central Greece/Euboea), Konitsis(Epirus), Grevenon(Macedonia), Elassonos(Thessaly). **Highest literacy rates:** Tinou(Cyclades), Agiou Vasileiou(Crete), Evrytanias(Central Greece/Euboea), Apokoronou(Crete).

Map 9

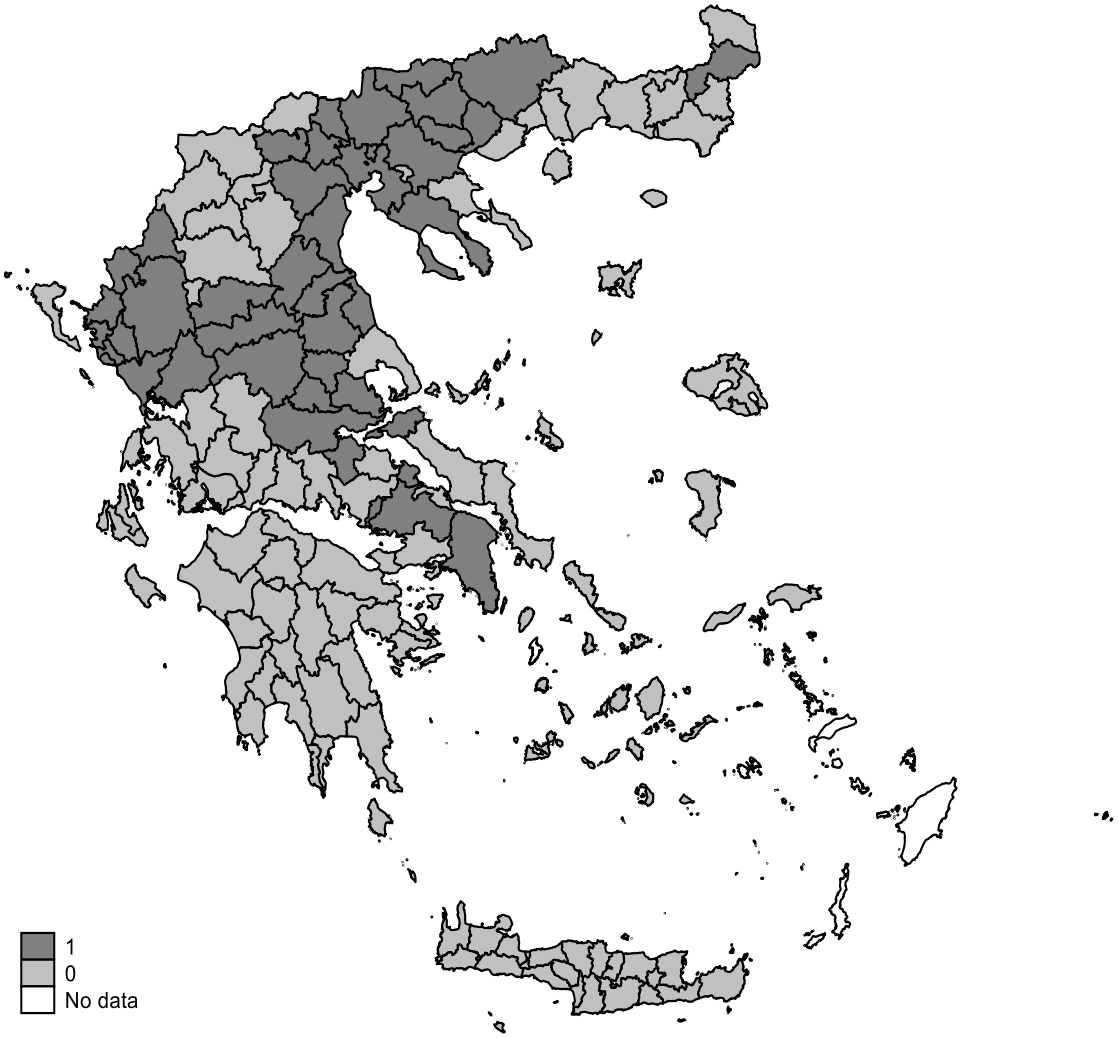
Total Refugee Literacy Rate 1928



Lowest literacy rates: Domokou(Central Greece/Euboea), Didymoteichou(Western Thrace), Grevenon(Macedonia), Paramythias(Epirus). **Highest literacy rates**: Agiou Vasileiou(Crete), Apokoronou(Crete), Sfakion(Crete), Metsovou(Epirus).

Map 10

Land Ownership in 19th century



Number ¨0¨ corresponds to provinces with small land ownership and number ¨1¨ corresponds to provinces with large land ownership. Provinces with dark gray color characterized by large landholdings.

Tables

Table 1

Definition of Variables

Variable	Description	Period	Source
Human Capital	Literacy rates (native and refugee)	1928	Census (1929), Hellenic Statistical Authority
Land Ownership	Dummy variable ("0" for regions with small land ownership and "1" for regions with large land ownership)	late 19th century	Petmezas (2003,2006,2012) Vergopoulos (1975)
Urbanization	Urbanization rate (% of population of province living in settlements with population>10,000 in 1928)	1928	Census (1929), Hellenic Statistical Authority
Population density	The quotient of population with land area	1928	Census (1929), Hellenic Statistical Authority
Altitude	Altitude in meters	1951	Hellenic Statistical Authority
Ruggedness	Standard deviation of altitude	1951	Calculated from altitude
Precipitation	Median precipitation in mm	1950-2000	IPUMS Terra
Temperature	Median Temperature in ^{0}C	1950-2000	IPUMS Terra
Soil yield	Wheat, sugar, pasture, olive, rice, potato yields		IPUMS Terra
Bulgarians/Turks	The quotient of Bulgarians/Turks with population	1920	Census (1929), Hellenic Statistical Authority

Table 2

Descriptive Statistics of Greek Provinces in 1928

Variable	Mean	Standard Deviation	Min	Max	Observations
native male literacy rate	.6095	.0902	.1837	.7712	141
native female literacy rate	.3222	.1063	.0763	.5723	141
native total literacy rate	.4647	.0856	.1297	.6626	141
male literacy rate	.6104	.0814	.2632	.7702	141
female literacy rate	.3230	.1032	.1139	.5724	141
total literacy rate	.4647	.0799	.1884	.6530	141
refugee male literacy rate	.6398	.1197	.3333	1	140
refugee female literacy rate	.4144	.1591	.0560	.8667	138
refugee total literacy rate	.5305	.1379	.2389	1	140
land area	921.0504	671.2756	31	3494	141
population density	51.4635	48.7827	9.3051	445.6735	141
land ownership	.2766	.4489	0	1	141
urbanization	.1103	.1838	0	.8193	141
altitude	280.5245	223.5981	28	1045	141
ruggedness	176.0746	74.5371	13.8914	363.9166	139
olive yield	.0716	.1048	0	.5164	137
pasture yield	7352337	7871879	0	5.26e+07	136
potato yield	.0038	.0041	4.67e-06	.0198	137
precipitation	114.4734	41.3183	47	196.5	141
rice yield	.0012	.0035	0	.0188	137
temperature	151.2589	21.7937	79.5	185	141
wheat yield	.0456	.0643	.0004	.2760	137
sugar yield	.0023	.0042	0	.0233	137
Bulgarians/Turks	.0833	.1795	0	.9147	142

Table 3

Correlation of Geographic Factors in Greek Provinces

	altitude	ruggedness	olive yield	pasture yield	potato yield	precipitation	rice yield	temperature	wheat yield	sugar yield
altitude	1.0000									
ruggedness	0.2528*	1.0000								
olive	-0.0972*	-0.2115*	1.0000							
pasture yield	-0.0852*	-0.1234*	0.3221*	1.0000						
potato yield	-0.0998*	-0.1785*	0.7825*	0.4189*	1.0000					
precipitation	0.2220*	-0.0067	0.3536*	0.4430*	0.3745*	1.0000				
rice yield	-0.2019*	0.0993*	-0.1570*	0.0111	-0.1307*	-0.4684*	1.0000			
temperature	-0.6510*	-0.3112*	0.4217*	0.3192*	0.3731*	0.2255*	-0.1011*	1.0000		
wheat yield	-0.0998*	0.0554	-0.2196*	-0.1571*	-0.1221*	-0.6313*	0.7054*	-0.3307*	1.0000	
sugar yield	-0.2049*	0.0024	-0.2715*	-0.2144*	-0.1324*	-0.5639*	0.5307*	-0.3314*	0.8757*	1.0000

Table 4

Fixed Effect Estimations of Native Literacy Rates in Greek provinces in 1928

	(1)	(2)	(3)	(4)	(5)	(6)
	native male literacy	native female literacy	native total literacy	native male literacy	native female literacy	native total literacy
land ownership	-0.018	-0.001	-0.007	-0.018	-0.001	-0.007
	(1.11)	(0.04)	(0.38)	(1.21)	(0.05)	(0.48)
land area	0.000	-0.000*	-0.000	0.000	-0.000	-0.000
	(0.04)	(2.05)	(0.92)	(0.05)	(1.25)	(0.87)
pop density	0.000	0.000	0.000	0.000*	0.000	0.000
	(1.75)	(1.27)	(1.51)	(1.70)	(1.50)	(1.66)
urbanization	0.089***	0.146***	0.127***	0.089***	0.146***	0.127***
	(4.53)	(4.17)	(5.29)	(4.30)	(4.28)	(5.02)
altitude	0.000	0.000	0.000	0.000	0.000	0.000
	(0.96)	(0.38)	(0.50)	(1.23)	(0.46)	(0.62)
ruggedness	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.07)	(0.73)	(0.44)	(0.08)	(0.80)	(0.52)
precipitation	-0.000	-0.001*	-0.001	-0.000	-0.001**	-0.001
	(0.08)	(2.03)	(1.34)	(0.07)	(2.49)	(1.31)
temperature	0.001	0.001	0.001	0.001	0.001	0.001
	(0.99)	(0.76)	(0.95)	(1.16)	(0.98)	(1.19)
olive yield	0.053	0.067	0.051	0.053	0.067	0.051
	(0.53)	(0.50)	(0.48)	(0.43)	(0.51)	(0.42)
pasture yield	-0.000*	-0.000***	-0.000**	-0.000**	-0.000***	-0.000***
	(2.17)	(3.62)	(3.23)	(2.49)	(3.69)	(3.62)
potato yield	-2.429	-0.560	-0.967	-2.429	-0.560	-0.967
	(0.92)	(0.15)	(0.33)	(0.77)	(0.16)	(0.30)
rice yield	-2.429	-4.698	-3.711	-2.429	-4.698	-3.711
	(1.31)	(1.44)	(1.63)	(0.72)	(1.12)	(1.13)
wheat yield	0.085	-0.114	-0.015	0.085	-0.114	-0.015
	(0.77)	(0.36)	(0.07)	(0.45)	(0.44)	(0.07)
Bulgarians/Turks	-0.054	-0.063	-0.040	-0.054	-0.063	-0.040
	(1.75)	(1.60)	(1.25)	(1.28)	(1.22)	(0.86)
constant	0.510***	0.351	0.416**	0.510***	0.351**	0.416***
	(3.42)	(1.71)	(2.41)	(3.78)	(2.25)	(2.98)
R^2	0.63	0.60	0.60	0.63	0.60	0.60
Observations	134	134	134	134	134	134

Notes: Fixed effect estimations with native male, female and total literacy rates as dependent variables in Greece in 1928. Fixed effects at department level and cluster standard errors at department level (Columns (1), (2), (3)) and at prefecture level (Columns (4), (5), (6). Some coefficients are zero because the values are rounded off to three decimal points. *, **, *** Significant at 10%, 5%, 1% respectively.

Table 5

OLS Estimations of Native Literacy Rates on Land Ownership in Greek provinces in 1928

	(1)	(2)	(3)	(4)	(5)	(6)
	native male literacy	native female literacy	native total literacy	native male literacy	native female literacy	native total literacy
land ownership	-0.018	-0.001	-0.007	-0.018	-0.001	-0.007
	(1.11)	(0.04)	(0.38)	(1.21)	(0.05)	(0.48)
land area	0.000	-0.000*	-0.000	0.000	-0.000	-0.000
	(0.04)	(2.05)	(0.92)	(0.05)	(1.25)	(0.87)
pop density	0.000	0.000	0.000	0.000*	0.000	0.000
	(1.75)	(1.27)	(1.51)	(1.70)	(1.50)	(1.66)
urbanization	0.089***	0.146***	0.127***	0.089***	0.146***	0.127***
	(4.53)	(4.17)	(5.29)	(4.30)	(4.28)	(5.02)
altitude	0.000	0.000	0.000	0.000	0.000	0.000
	(0.96)	(0.38)	(0.50)	(1.23)	(0.46)	(0.62)
ruggedness	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.07)	(0.73)	(0.44)	(0.08)	(0.80)	(0.52)
precipitation	-0.000	-0.001*	-0.001	-0.000	-0.001**	-0.001
	(0.08)	(2.03)	(1.34)	(0.07)	(2.49)	(1.31)
temperature	0.001	0.001	0.001	0.001	0.001	0.001
	(0.99)	(0.76)	(0.95)	(1.16)	(0.98)	(1.19)
olive yield	0.053	0.067	0.051	0.053	0.067	0.051
	(0.53)	(0.50)	(0.48)	(0.43)	(0.51)	(0.42)
pasture yield	-0.000*	-0.000***	-0.000**	-0.000**	-0.000***	-0.000***
	(2.17)	(3.62)	(3.23)	(2.49)	(3.69)	(3.62)
potato yield	-2.429	-0.560	-0.967	-2.429	-0.560	-0.967
	(0.92)	(0.15)	(0.33)	(0.77)	(0.16)	(0.30)
rice yield	-2.429	-4.698	-3.711	-2.429	-4.698	-3.711
	(1.31)	(1.44)	(1.63)	(0.72)	(1.12)	(1.13)
wheat yield	0.085	-0.114	-0.015	0.085	-0.114	-0.015
	(0.77)	(0.36)	(0.07)	(0.45)	(0.44)	(0.07)
Bulgarians/Turks	-0.054	-0.063	-0.040	-0.054	-0.063	-0.040
	(1.75)	(1.60)	(1.25)	(1.28)	(1.22)	(0.86)
constant	0.272**	0.177	0.213	0.272**	0.177	0.213*
	(2.34)	(1.04)	(1.50)	(2.28)	(1.45)	(1.87)
R^2	0.63	0.60	0.60	0.63	0.60	0.60
Observations	134	134	134	134	134	134

Notes: OLS estimations with native male, female and total literacy rates as dependent variables in Greek provinces in 1928. Fixed effects at department level and cluster standard errors at department level (Columns (1), (2), (3)) and at prefecture level (Columns (4), (5), (6). Some coefficients are zero because the values are rounded off to three decimal points. *, **, *** Significant at 10%, 5%, 1% respectively.

Table 6

Fixed Effect Estimations of Literacy Rates on Geographic Factors in Greek provinces in 1928

	(1)	(2)	(3)	(4)	(5)	(6)
	male literacy	female literacy	total literacy	male literacy	female literacy	total literacy
refugee male literacy	0.094*			0.094*		
	(2.08)			(1.76)		
land area	-0.000	-0.000*	-0.000	-0.000	-0.000**	-0.000
	(0.82)	(2.09)	(1.25)	(0.90)	(2.11)	(1.48)
pop density	0.000	0.000	0.000	0.000	0.000	0.000
	(1.22)	(0.99)	(1.06)	(1.19)	(1.02)	(1.06)
urbanization	0.093***	0.172***	0.136***	0.093***	0.172***	0.136***
	(4.13)	(4.52)	(5.23)	(4.20)	(4.66)	(5.06)
altitude	0.000	0.000	0.000	0.000	0.000	0.000
	(1.21)	(1.38)	(0.83)	(1.36)	(1.32)	(0.93)
ruggedness	0.000	-0.000	-0.000	0.000	-0.000	-0.000
	(0.40)	(0.55)	(0.06)	(0.43)	(0.59)	(0.06)
precipitation	-0.000	-0.001**	-0.001**	-0.000	-0.001***	-0.001**
	(0.63)	(3.24)	(2.44)	(0.51)	(3.10)	(2.09)
temperature	0.001	0.002	0.001	0.001	0.002*	0.001*
	(1.62)	(1.60)	(1.62)	(1.51)	(1.77)	(1.78)
olive yield	-0.002	0.003	0.002	-0.002	0.003	0.002
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
pasture yield	-0.000*	-0.000***	-0.000**	-0.000**	-0.000***	-0.000***
	(2.14)	(3.78)	(3.19)	(2.32)	(3.56)	(3.35)
potato yield	-0.909	0.582	0.188	-0.909	0.582	0.188
	(0.31)	(0.18)	(0.07)	(0.29)	(0.16)	(0.06)
rice yield	-2.427	-3.482	-3.242	-2.427	-3.482	-3.242
	(1.54)	(1.17)	(1.44)	(1.06)	(1.03)	(1.24)
wheat yield	0.078	-0.109	0.004	0.078	-0.109	0.004
	(0.76)	(0.40)	(0.02)	(0.49)	(0.46)	(0.02)
Bulgarians/Turks	-0.043*	-0.048	-0.043	-0.043	-0.048	-0.043
	(2.00)	(1.00)	(1.34)	(1.29)	(0.88)	(1.07)
refugee female literacy		0.081*			0.081**	
		(1.88)			(2.04)	
refugee total literacy			0.101**			0.101***
			(2.33)			(2.93)
constant	0.465***	0.236	0.350**	0.465***	0.236	0.350***
	(4.72)	(1.20)	(2.62)	(4.74)	(1.52)	(3.14)
R^2	0.63	0.64	0.63	0.63	0.64	0.63
Observations	133	131	133	133	131	133

Notes: Fixed effect estimations with male, female and total literacy rates as dependent variables in Greek provinces in 1928 (without land ownership as independent variable). Fixed effects at department level and cluster standard errors at department level (Columns (1), (2), (3)) and at prefecture level (Columns (4), (5), (6). Some coefficients are zero because the values are rounded off to three decimal points. Refugees included. *, **, *** Significant at 10%, 5%, 1% respectively.

Table 7

OLS Estimations of Literacy Rates on Land Ownership in Greek provinces in 1928

	(1)	(2)	(3)	(4)	(5)	(6)
	male literacy	female literacy	total literacy	male literacy	female literacy	total literacy
refugee male literacy	0.085*			0.085		
	(1.91)			(1.61)		
land ownership	-0.023*	-0.008	-0.010	-0.023*	-0.008	-0.010
	(2.18)	(0.32)	(0.92)	(1.76)	(0.39)	(0.90)
land area	-0.000	-0.000*	-0.000	-0.000	-0.000*	-0.000
	(0.24)	(1.96)	(1.09)	(0.29)	(1.84)	(1.25)
pop density	0.000	0.000	0.000	0.000	0.000	0.000
	(1.54)	(0.91)	(1.05)	(1.51)	(0.99)	(1.10)
urbanization	0.088***	0.170***	0.134***	0.088***	0.170***	0.134***
	(3.97)	(4.28)	(4.93)	(3.98)	(4.40)	(4.82)
altitude	0.000	0.000	0.000	0.000	0.000	0.000
	(0.96)	(1.30)	(0.72)	(1.10)	(1.25)	(0.82)
ruggedness	0.000	-0.000	-0.000	0.000	-0.000	-0.000
	(0.42)	(0.55)	(0.05)	(0.47)	(0.60)	(0.06)
precipitation	-0.000	-0.001**	-0.001**	-0.000	-0.001***	-0.001**
	(0.65)	(3.16)	(2.42)	(0.51)	(3.08)	(2.08)
temperature	0.001	0.002	0.001	0.001	0.002*	0.001*
	(1.33)	(1.57)	(1.53)	(1.43)	(1.74)	(1.73)
olive yield	0.013	0.007	0.008	0.013	0.007	0.008
	(0.14)	(0.07)	(0.10)	(0.11)	(0.05)	(0.07)
pasture yield	-0.000*	-0.000***	-0.000**	-0.000**	-0.000***	-0.000***
	(2.24)	(3.89)	(3.24)	(2.40)	(3.54)	(3.36)
potato yield	-1.373	0.456	-0.018	-1.373	0.456	-0.018
	(0.53)	(0.15)	(0.01)	(0.45)	(0.13)	(0.01)
rice yield	-1.509	-3.137	-2.824	-1.509	-3.137	-2.824
	(0.86)	(1.18)	(1.35)	(0.63)	(0.89)	(1.08)
wheat yield	0.110	-0.100	0.017	0.110	-0.100	0.017
	(1.03)	(0.36)	(0.08)	(0.68)	(0.43)	(0.08)
Bulgarians/Turks	-0.053*	-0.052	-0.047	-0.053	-0.052	-0.047
	(2.06)	(1.00)	(1.36)	(1.51)	(0.91)	(1.14)
refugee female literacy		0.079*			0.079*	
		(1.91)			(2.01)	
refugee total literacy			0.097*			0.097***
			(2.25)			(2.76)
constant	0.283**	0.120	0.200	0.283***	0.120	0.200**
	(3.19)	(0.76)	(1.79)	(3.56)	(1.06)	(2.40)
R^2	0.64	0.64	0.63	0.64	0.64	0.63
Observations	133	131	133	133	131	133

Notes: OLS estimations with male, female and total literacy rates as dependent variables in Greek provinces in 1928. Fixed effects at department level and cluster standard errors at department level (Columns (1), (2), (3)) and at prefecture level (Columns (4), (5), (6). Some coefficients are zero because the values are rounded off to three decimal points. Refugees included. *, **, *** Significant at 10%, 5%, 1% respectively.

Table 8

Regression Adjustment Model in Greek provinces in 1928 for literacy rates

	(1)	(2)	(3)
	male literacy	female literacy	total literacy
land ownership (ATE)	-0.019	-0.056***	-0.033**
	(1.18)	(3.10)	(2.19)
land ownership (POmean)	0.615***	0.337***	0.474***
	(66.49)	(32.97)	(56.00)
Observations	140	140	140

Table 9

Nearest-Neighbor Matching Model in Greek provinces in 1928 for literacy rates

	(1)	(2)	(3)
	male literacy	female literacy	total literacy
land ownership (ATE)	-0.029	-0.063***	-0.042**
	(1.30)	(2.84)	(2.03)
Observations	140	140	140

-

Figures

Figure 1

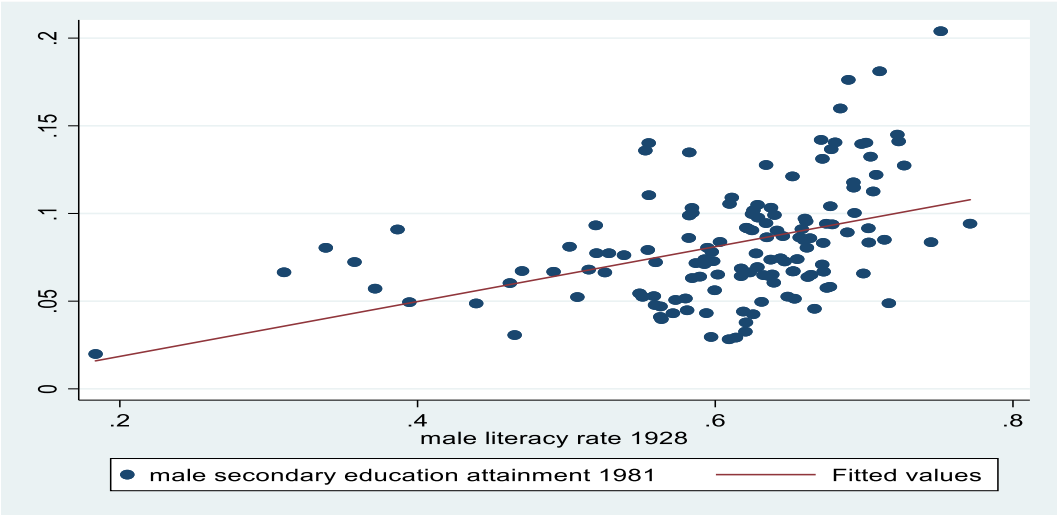


Figure 2

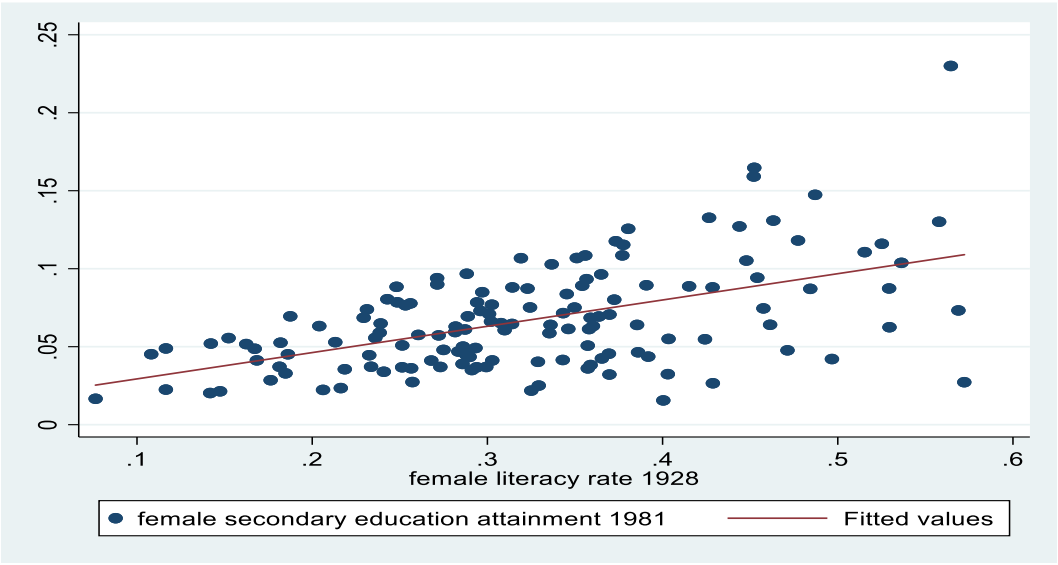


Figure 3

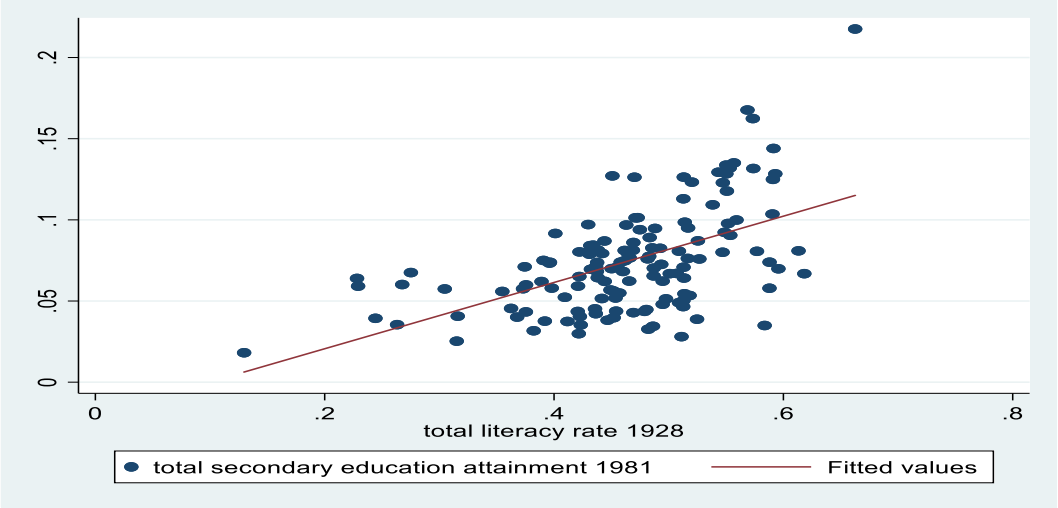


Figure 4

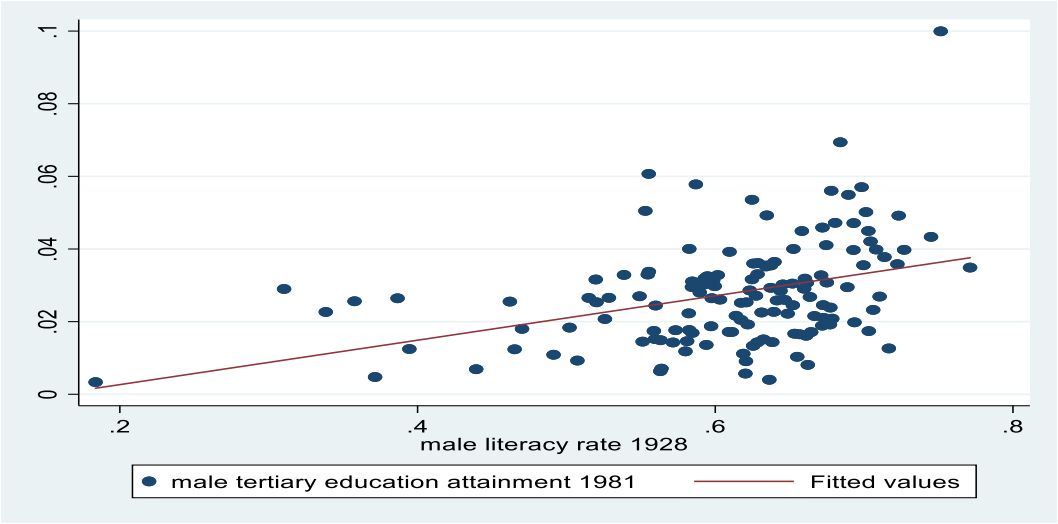


Figure 5

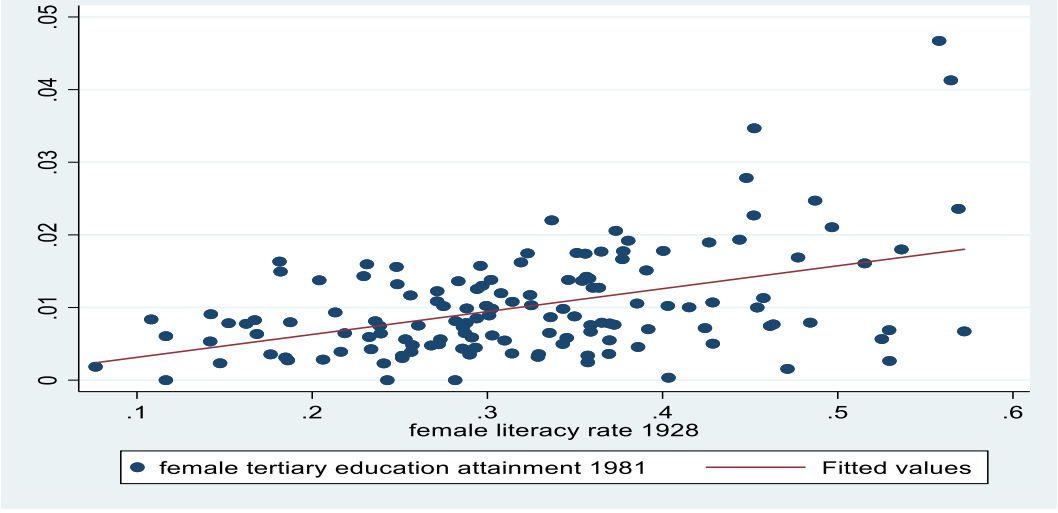


Figure 6

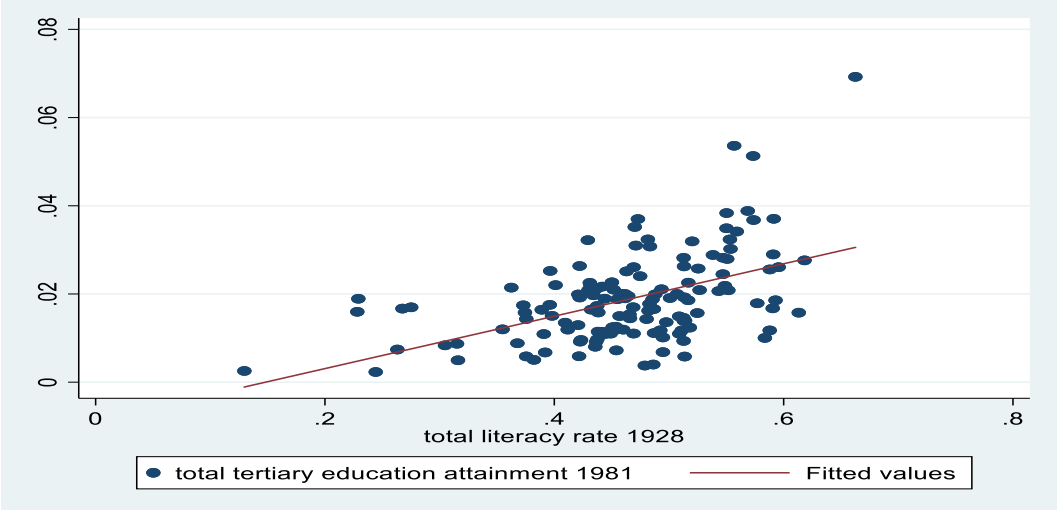


Figure 7

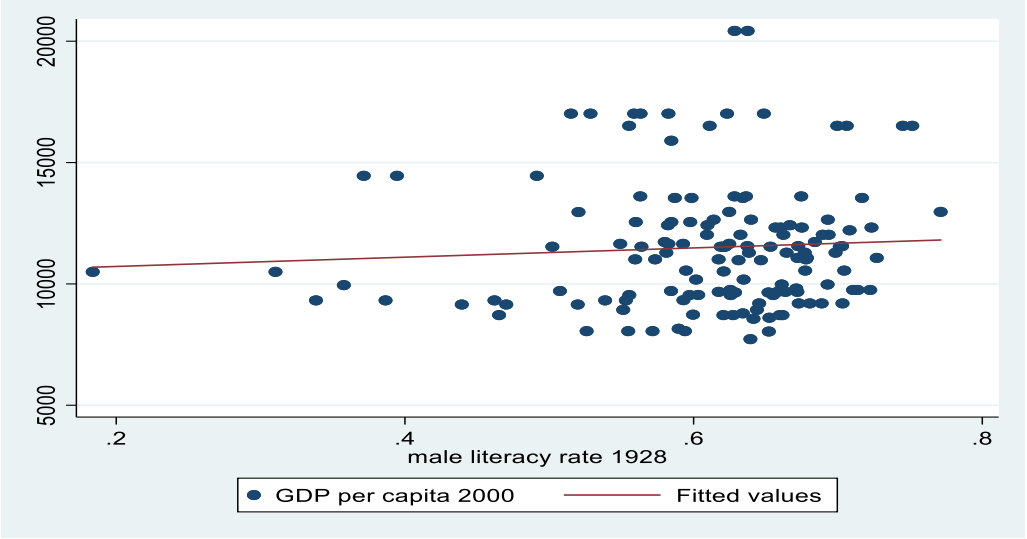


Figure 8

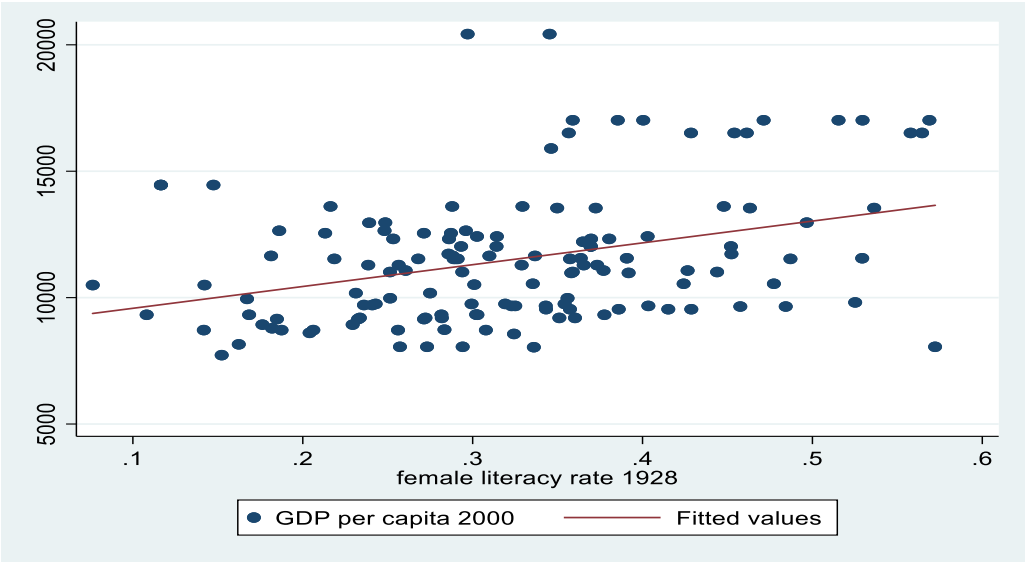


Figure 9