ECONOMIC RESEARCH FORUM

منتدى البحوث الاقتصادية

2018

working paper series

ECONOMIC GEOGRAPHY, GROWTH DYNAMICS AND HUMAN CAPITAL ACCUMULATION IN TURKEY: EVIDENCE FROM REGIONAL AND MICRO DATA

Burhan Can Karahasan and Firat Bilgel

Working Paper No. 1233

ECONOMIC GEOGRAPHY, GROWTH DYNAMICS AND HUMAN CAPITAL ACCUMULATION IN TURKEY: EVIDENCE FROM REGIONAL AND MICRO DATA[1]

Burhan Can Karahasan[2] and Firat Bilgel[3]

Working Paper No. 1233

October 2018

Authors would like to thank to Nona Karalashvili for useful comments and suggestions.

Send correspondence to:
Burhan Can Karahasan
Piri Reis University
bckarahasan@pirireis.edu.tr

[1] This work has benefited from a financial grant from the Economic Research Forum (ERF) under grant no. 2017-120 of the theme "Structural Change, Resource Misallocation and Growth Dynamics in the MENA Region." Earlier version is presented at the joint workshop of ERF and World Bank in Luxor. The content and the recommendations of this research do not necessarily reflect the views of the ERF.

[2] Department of Economics and Finance, Piri Reis University, 34940 Istanbul Turkey. Email: bckarahasan@pirireis.edu.tr.

[3] Initiative for Regulation and Applied Economic Analysis, Department of Agricultural Economics and Economics, Montana State University, Bozeman MT 59715 United States. Department of Economics and Finance, Okan University, 34959 Istanbul, Turkey. Email: Firat.bilgel@okan.edu.tr.

First published in 2018 by
The Economic Research Forum (ERF)
21 Al-Sad Al-Aaly Street
Dokki, Giza
Egypt
www.erf.org.eg

Abstract

This study explores the endogenous relationship among market access, wages and human capital accumulation in Turkey. Our first set of analyses tests the impact of market access on human capital development using regional data at the NUTS III level. Results, robust to the inclusion of spatial spillovers, regional structural differences in production, possible endogeneity issues and the unobserved regional heterogeneities, validate that regions with better access to markets are the ones that accumulate more human capital in Turkey. Our second set of analyses aims to explore the background of human capital accumulation by using individual level data, which allows us to combine market accessibility, returns to education (wages) and human capital development. Remarkably, once we include wages and treat it as endogenous, we find evidence that the impact of market access on human capital development diminishes. Overall, the findings of this study validate that background of the NEG model does not work in line with the expectations. Rather the influence of geographical proximity on wages and individual's decision on human capital investment are not identical.

Keywords: Human capital, market access, Turkey
JEL Classifications: R11, R12

ملخص

تستكشف هذه الدراسة العلاقة المتجانسة بين الوصول إلى الأسواق والأجور وتراكم رأس المال البشري في تركيا. فتختبر مجموعتنا الأولى من التحليلات أثر النفاذ إلى الأسواق على تنمية رأس المال البشري باستخدام البيانات الإقليمية على مستوى تسمية الوحدات الإقليمية لأغراض الإحصاء ثلاثة . النتائج التي تصل قوتها لحد إدراج الآثار الجانبية المكانية ، والاختلافات الهيكلية الإقليمية في الإنتاج ، وقضايا التجانس المحتملة والتغايرات الإقليمية غير المرصودة ، تؤكد أن الأقاليم التركية التي تتمتع بوصول أفضل إلى الأسواق هي تلك التى يتراكم فيها المزيد من رأس المال البشري. أما مجموعتنا الثانية من التحليلات فتهدف إلى استكشاف خلفية تراكم رأس المال البشري باستخدام بيانات المستوى الفردي ، مما يتيح لنا الجمع بين الوصول إلى الأسواق ، وعائد التعليم (الأجور) وتنمية رأس المال البشري. من اللافت للنظر ، أنه بمجرد إدراج الأجور واعتبارها عنصرا أصيلا ، نجد دليلاً على أن تأثير الوصول إلى الأسواق على تنمية رأس المال البشري يتضاءل. وبشكل عام ، تؤكد نتائج هذه الدراسة أن خلفية نموذج الجغرافيا الاقتصادية الجديدة لا تتمشى مع التوقعات حيث أنه لا يوجد تطابق بين تأثير القرب الجغرافي على الأجور وقرار الفرد بشأن استثمار رأس المال البشري.

1 Introduction

Economists have investigated various dimensions of economic growth to understand between and within country differences. Technological advances, productivity improvements and resource allocation efficiency are heavily used to examine inequalities (Solow, 1956; Lucas, 1988; Romer, 1994). While contradicting theoretical insights between neoclassical and endogenous growth models dominate the economic growth literature, Krugman (1991) opens a new line by discussing the role agglomeration economies play in understanding the differences in level of economic activity. Indeed, Gallup et al. (1999) suggests that geography is important not only because increasing returns matter but also many economic and social dimensions of inequalities have a geographic pattern. Neoclassical theories consider the importance devoted to geography by specifying first nature advantage. Later on, the new economic geography (NEG) literature formalizes the way geography enters within the augmented version of production function. Redding and Venables (2004a,b) define and use the concept of market accessibility as a factor explaining cross-country variations in per capita income. Later, this approach becomes an inspiration to development economists as well as urban and regional economists.

Following the formulation of economic geography through agglomeration economies and possible externalities of market accessibility, interests of scholars shift to the relationship between the determinants of economic growth and agglomeration economies. Revisiting productivity and technology differences, economic geography literature explains the pattern of regional integration and factor accumulation by using their own formal models. Even Coe and Helpman (1995); Coe et al. (1995) and Keller (2000) discuss the negative impact of distance on technological spillovers, it is Redding and Schott (2003) to discuss the endogenous evolution of production factors (i.e. human capital) within a NEG model. The novelty of the NEG approach lies in the way geography is defined. Distance is no longer the only factor that matters but also trade, income and economic potential is embedded into an accessibility measure together with physical distance. Once accessibility is defined in an inclusive way, the remaining effort is to formally relate geography with factor accumulation. In an indirect way, NEG model asserts that the ease of accessing to market represents rising profitability at the firm and region level. This process enables firms and regions to generate extra value-added to production factors (i.e. skilled premium to educated workers). Redding and Schott (2003) considers the impact of wage premium to skilled workers and underlines that expecting higher wage premiums in central locations (with higher market access) stands as a stimulus for individuals to invest more in human capital development. Among different studies, evidence from Europe indicates that rising market access is associated with better human capital endowment and accumulation at the regional level (Faíña and Lopez-Rodriguez, 2006; López-Rodríguez et al., 2007).

While the theoretical link between geography and human capital accumulation is tested vastly, there is limited discussion on the way that returns to education and other individual characteristics influence human capital accumulation. Additionally, little is known about the comparative effect of market accessibility on wages and human capital development. Originating from this concern, the central objective of this research is to critically evaluate the link between market access and human capital accumulation for a developing country, Turkey. For this purpose, we use two strategies. First, we use a regional data set at the NUTS III level and observe the historical evolution market access and human capital accumulation relationship. While doing this we augment the traditional NEG model by considering possible endogenetiy of market access, regional heterogeneities and possible spatial externalities. Controlling for structural differences among regions and taking into account spatial dependence are two important paths to augment the traditional NEG model. As a second approach, we use micro data and focus more on the background of the NEG model. Our micro data set is representative at the NUTS I disaggregation. Our objective is to consider a number of issues which has not been discussed by the NEG literature in detail: (i) other individual characteristics that can affect human capital investment, (ii) the so-called black-box between market access and human capital accumulation (through wages) and (iii) the endogeneity of wages.

This study contributes to the existing literature from a number of points. Filiztekin (1998); Dogruel and Dogruel (2003); Gezici and Hewings (2007) underline that the geographical split between the west and the east of the country represents the historical origins of inequalities in Turkey.[1] Although the literature on Turkey considers different dimensions, attempts to implement the NEG framework are scarce.[2] Additionally, empirical studies mostly provide evidence from developing countries, leaving developing and less developed countries relatively less investigated. Moreover Duranton and Puga (2004) and Duranton and Overman (2005) underline that more attempt is needed to focus on how economic activity is localized at the micro level. However, given data concerns and the difficulty to incorporate individual level data to NEG framework, only a number of influential studies prefer the use of micro level data. To our knowledge Elbadawi et al. (2009), Fally et al. (2010) are two influential attempts to test NEG framework with the use of micro data, together with the aggregate data sources. Finally, as underlined in Redding (2010) endogenous dependence among wages, human capital accumulation and market access has not been central to applied studies testing the NEG model. We discuss that this endogenous feedback is an essential part of the model.

This study proceeds as follows: section 2 explains the theoretical background of the study, section 3 introduces the research strategy by explaining the use of regional and micro level data for the NEG model, section 4 reports the results, section 5 discusses the policy implications derived from the comparison of two

[1]See also Yildirim et al. (2009), Elveren (2010), Celebioglu and Dall-Erba (2010) for recent studies on regional inequalities in Turkey.

[2]To our knowledge Karahasan et al. (2016) is a comprehensive attempt to test the NEG framework by testing the impact of market access on wages. However, data constraints prevent the study to explore the post 2000s.

different approaches in detail and finally, section 6 concludes.

2 Theoretical Background

In light of the research on regional and urban economics, recent advances in NEG motivates development economists. Krugman (1991), Krugman and Venables (1995), Venables (1996) and Fujita et al. (1999) formalized the way agglomeration economies explain the distribution of economic activities across space. These developments motivate development economists to discuss the trade-off among clustering and inequalities at the regional level. An important contribution of the NEG framework is the importance attributed to market accessibility. Being close to demand, supply and having active networks with high income regions represent the market accessibility of locations. Naturally, locations with higher access to markets benefit from externalities generated among their geography. These locations offer higher profits for firms and create incentive mechanisms for accumulating production factors.

Redding and Venables (2004a,b) use this NEG reasoning and define market access/potential (access from now on) in order to explain inequalities across nations. Accessibility is defined in two different forms: (i) market access (MA) referring to being close to demand, (ii) supply access (SA) defines how firms may access to source of production. While theoretical model distinguishes supply and market access; given high correlation between the two, empirical specifications tend to focus only on the market accessibility. In any case, the novelty of both theoretical and empirical models lie in the way these two access measures are calculated. While Redding and Venables (2004b) measures market access from a gravity equation taking into account geographical distance and export flow across countries, Bosker and Garretsen (2010) underline the direct use of distance and regional demand potential (i.e. income, value added, population). Breinlich (2006); Boulhol and De Serres (2010); Hering and Poncet (2010); Head and Mayer (2011) compute the market access based on a gravity equation. On the contrary Mion (2004); Hanson (2005); Ottaviano and Pinelli (2006); Niebuhr (2006); Brakman et al. (2006); López-Rodríguez et al. (2007); Kosfeld and Eckey (2010) compute the distance-weighted version of market accessibility. Regardless of the method used, studies underline that firms, regions and countries that have better access to markets tend to be wealthier on average.

Redding and Schott (2003) offers a wider perspective that does not only explain the distribution of income but also depicts the geographic dispersion of production factors. This seems to be a vital turning point for NEG as the model now allows for the endogenous accumulation of production factors. Firms that locate in remote locations face higher trade costs, which disable them to accumulate more value added to be distributed to factors of production. In contrast, firms that have higher market access generate higher amount of value added. In turn, firms located in the center (with more market access) have more possibility to compensate production factors. Among different factors, human capital and specifically skilled human capital tend to benefit more as firms will be more reluctant to pay the required skill premium in central locations. Redding and Schott (2003) defines this condition by introducing a wage equation (equation 1). w_i^s and w_i^u represents wages to skilled and unskilled workers with α and β factor shares respectively. σ is elasticity of substitution and E is the consumption of manufacturing goods with a price index of G. Finally, c denotes marginal input requirement and $T_{i,j}^M$ is an iceberg-type transportation cost for manufacturing production (M).

$$(\frac{\sigma}{\sigma-1}(w_i^s)^{\alpha}(w_i^u)^{\beta}G_i^{(1-\alpha-\beta)}c_i)^{\alpha} = (\frac{1}{\bar{x}})\sum_{j=1}^{R} E_j G_j^{\sigma-1}(T_{ij}^M)^{1-\sigma} \tag{1}$$

Re-arranging equation 1 yields equation 2 which defines the maximum amount a firm in region i can afford to pay to its skilled and unskilled workers. Equation 3 shows how individuals decide human capital investment. $w_i^s - w_i^u$ is the skilled premium and $a(z)$ is a critical ability level. This defines the ease of human capital accumulation. Finally, h_i is an institutional parameter assumed to be homogenous for regions of the same country.

$$(w_i^s)^{\alpha}(w_i^u)^{\beta} = \zeta\frac{1}{c_i}(MA_i)^{\frac{1}{\sigma}}(SA_i)^{\frac{(1-\alpha-\beta)}{(\sigma-1)}} \tag{2}$$

$$w_i^s - w_i^u \geq \frac{h_i}{a(z)}w_i^u \tag{3}$$

Redding and Schott (2003) defines the condition that links market accessibility with human capital accumulation (equation 4). This condition implies that a fall in market access defines a lower relative wage rate for skilled workers if manufacturing production is skill-intensive. Revisiting equation 3 which is basically the skilled indifference condition; Faíña and Lopez-Rodriguez (2006), López-Rodríguez et al. (2007) discuss that in the new equilibrium condition a higher critical ability level will be defined over which individual becomes a skilled worker. This results in less incentive to accumulate human capital if regions are faced with diminishing market accessibility.

$$\alpha\frac{dw_i^s}{w_i^s} + \beta\frac{dw_i^u}{w_i^u} = \frac{1}{\sigma}\frac{dMA_i}{MA_i} + \frac{(1-\alpha-\beta)}{(\sigma-1)}\frac{dSA_i}{SA_i} \tag{4}$$

Overall the NEG model defines that it is indeed the relationship between market access and wages (Equation 2) that affects the individuals' human capital accumulation decision (Equation 3). That is, the impact of geography on human capital accumulation is labeled through a black box in which we identify a positive relationship between wages and human capital accumulation. Therefore, the NEG model assumes (without testing) that individuals' education decisions are nudged by the expected future skilled premium. As the model formalizes skill premium as a positive function of market access; we expect to observe a positive relationship between market access and human capital accumulation as well.

3 Research Design

Even though the augmented version of the NEG model is informative, there is still an ongoing discussion in order to improve the NEG framework for problems such as identification and missing mechanisms (Redding, 2010). For instance, Karahasan and López-Bazo (2013) remarks that the link between market accessibility and factor accumulation of human capital is affected by the spatial dimension of inequalities as well as structural differences in production. Moreover, the theoretical background of the NEG model as well as its augmented version lack a formal explanation to explain how wages and human capital are connected. Other possible factors that affect wages might also have a geographical pattern. Motellón et al. (2011) investigates the wage distribution in Spain by using micro-level survey data and highlights that not only wages are characterized by marked differences regionally, but also specific characteristics influencing the wage distribution are spatially unequal. The impact of individual characteristics of workers at the micro level and the way these individual observations vary across space seem to be unanswered questions for developing countries and stand as significant motivations for our study. More importantly, endogeneity is an important dimension of the existing NEG model. For instance, Boulhol and De Serres (2010) discusses that market access is not exogenous while explaining income and wage dispersion. Similarly, Karahasan and López-Bazo (2013) remarks that it is less likely to disregard the endogeneity issue while using market access to explain human capital differences. Nevertheless, attempts to tackle the endogeneity of wages are lacking.

Based on these concerns we construct a two-stage design. First in sub-section 3.1, we discuss various dimensions of regional evolution of market access and human capital development by controlling for structural differences, regional heterogeneities, spatial spillovers and endogeneity of market access. Here we use a regional data set allowing us to combine space (NUTS III) and time dimension (1985 to 2014) within different specifications. Next in sub-section 3.2 we focus more on the black box between market access and human capital accumulation by introducing wage distribution at the individual level. Our micro data allows us to focus more on individual characteristics this time, only at the NUTS I level for the year 2014. That said, now we are able to use human capital, market access and wages in the same framework.

3.1 First-Step: Regional Data and Testing the impact of market access on human capital accumulation

Our first attempt is to estimate the theoretical NEG model in various forms and to test the impact of market access on human capital accumulation. We use regional data at NUTS III disaggregation. One important challenge is the calculation of market access at the regional level. Among different ways of constructing the market access index, we prefer to use the Harris (1954) approach. Equation 5 is the market access index where Y_i is the per capita income of region i and $D_{i,j}$ is the motorway distance between any pair of regions i, j, retrieved from the General Directorate of Highways Republic of Turkey.[3] Regional per capita income at constant prices comes from Turkish Statistics Office (TurkStat).

$$MA_i = \sum_{j=1}^{n} \frac{Y_j}{D_{i,j}} \tag{5}$$

Redding and Schott (2003); Redding and Venables (2004a) use the gravity approach to calculate the market index. This complex two-step approach is criticized on the grounds that one has to make a set of arbitrary assumptions on the structure of trade costs. Anderson and Van Wincoop (2004); Fingleton (2008); Bosker and Garretsen (2010) discuss that region specific properties like distance, adjacency, trade barriers, language etc. can overestimate the region specific properties resulting in biased measurement of market accessibility. Based on these concerns and the lack of reliable data to implement the gravity approach for Turkey, we use distance-weighted per capita income approach rather than an auxiliary gravity model in our market access calculations. We consider the within region dynamics by calculating the internal demand which is weighted by the intra-regional distance via Head and Mayer (2006) approximation of $D_{i,j} = 0.66\sqrt{Area_i/\pi}$.

As discussed in Faíña and Lopez-Rodriguez (2006); López-Rodríguez et al. (2007), direct distance to certain locations can also be used in order to understand the ease of market access. We consider three important metropolitan areas of Turkey: Istanbul, Ankara and Izmir. While we use direct distance to these economic centers as a proxy to understand market access, we also make a comparison between direct distance measures and market access index. Our aim is to question whether distance is a broad proxy for accessibility or distance acts as a compound factor influencing the demand and supply-based linkages within a given geography.

On the side of human capital accumulation we calculate average years of schooling at the NUTS III level.[4] Human capital data comes from Population Census Data Base and Address Based Registry System

[3] Travel time can also be used to measure distance, however; we do not have reliable data on travel times across regions of Turkey.

[4] We also calculate the share of individuals with at least university education for each NUTS III region as a second human capital measure. All analyses are replicated by using this second proxy. Results are available upon request.

of TurkStat. Both census and the registry system data enable us to calculate the number of individuals with primary (5 years), secondary (8 years), high (11 years) and university (15 years) education. We consider population above 6 years of age in our analysis. For the pre-2008 period, we only have information from 1985, 1990 and 2000 population census. After 2008 with the integration of Address Based Registry System we are able to obtain annual human capital data for the 2009-2017 period. Based on data availability, we consider 1985, 2000 and 2009-2014 intervals separately.[5] It is important to remark that the duration of the compulsory education increased from 5 years to 8 years in 1996. Our data enables us to come over any possible bias that can evolve from the system change; as the census and the registry data directly reports the decomposed figures. That is, we are able to reach the exact number of individuals with certain education years both in the census and the registry data.

Before introducing the empirical estimation of the NEG model, we start by a number of spatial data analyses. Our aim is to have a preliminary idea on the level of spatial links regarding human capital accumulation. We observe both the path of global spatial auto-correlation and also the persistence of local spatial auto-correlation. Equation 6 (Moran's I) and equation 7 (Local Indicator of Spatial Association, LISA) are the global and local measures of spatial auto-correlation respectively, where n is the number of cross sections, s is the summation of the all elements in the weight matrix (w) (Anselin, 1995, 1996). We construct a contiguity weight matrix, which is a binary weight matrix assigning 1 to adjacent units and 0 otherwise.[6]

$$I = \frac{n}{s} \frac{\sum_i w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum z_i^2} \tag{6}$$

$$I_i = (x_i - \bar{x}) \sum_j w_{ij}(x_j - \bar{x}) \tag{7}$$

Next we construct the formal NEG model as offered in Redding and Schott (2003). In the initial phase we estimate models for years 1985 and 2000 cross-sections separately. Following this we also estimate the panel fixed-effects versions for the 2009-2014 period, where we also take into account time-invariant unobserved heterogeneity. Our benchmark models are non-spatial. Equation 8 is estimated for years 1985 and 2000, while equation 9 is the fixed-effects model used to estimate the 2009-2014 interval. i and t represents cross section and time dimension respectively. HK is the average years of schooling, MA is the natural log of market access, X contains information on the composition of regional labor demand, unemployment rate and first nature geographic advantages. Naturally we expect to see that the share of employment in specific

[5] Number of cross sections change from 67 to 81 provinces during the pre-2000s as some new provinces are formed from the existing ones. Note that, new provinces share more or less similar structures with their origins.

[6] We replicate all spatial analysis by using other weight matrices (i.e. inverse distance, threshold distance). These results are available upon request.

industries acts as a proxy for the regional labor demand which is assumed to affect the incentive to invest in human capital. Moreover, unemployment is also a socioeconomic indicator which affects human capital accumulation negatively. Finally, we control for the first nature advantages of regions by using the elevation of each province and a coastal dummy controlling for the impact of being located in accessible areas.

$$HK_i = \alpha + \beta MA_i + \gamma X_{i,t} + \varepsilon_i \tag{8}$$

$$HK_{i,t} = \alpha + \beta MA_{i,t} + \gamma X_{i,t} + v_i + \varepsilon_{i,t} \tag{9}$$

Next we augment our specifications by embedding spatial dependence into the NEG model (Karahasan and López-Bazo, 2013). Following Anselin (2010) and Elhorst (2010) we use four different spatial specifications. Equations 10, 11, 12 and 13 are the panel specifications for the spatial lag model (SAR), spatial error model (SEM), spatial Durbin model (SDM) and spatial Durbin error model (SDEM) respectively (we skip the representation of the cross-section models). Both for cross section and panel models; SAR assumes spillovers over dependent variable, SEM considers the common spillover of shocks thus omitted variables and finally SDM and SDEM respectively take into account the possible spatial spillovers from dependent and independent variables together with omitted variables.[7] Note that we follow the approach offered in Elhorst (2010) to compare the spatial models' specifications (See subsection 4.1 for implementations of LR-Test).

Note that, one should take into account the fact that geography enters into the realm of these models in two different forms. One, over the MA variable and second over the weight matrix (W). One may naturally think on the possible overrepresentation of the geographical dimension. However, as discussed in Karahasan et al. (2016), while the weight matrix is incorporated to define location-based networks over spatial dependence, market access variable incorporates the impact of purchasing power, weighted by the distance. Moreover, since we use a contiguity weight matrix rather than an inverse distance, we have no reason to expect that our results will be influenced from a possible relationship between use of distance in MA index and spatial dependence parameter (i.e. $Wy_{i,t}$ or $W\epsilon_{i,t}$ etc).

$$HK_{i,t} = \alpha + \beta MA_{i,t} + \rho WHK_{i,t} + \gamma X_{i,t} + v_i + \varepsilon_{i,t} \tag{10}$$

$$HK_{i,t} = \alpha + \beta MA_{i,t} + \gamma X_{i,t} + v_i + \lambda We_{i,t} + \varepsilon_{i,t} \tag{11}$$

$$HK_{i,t} = \alpha + \beta MA_{i,t} + \rho WHK_{i,t} + \gamma X_{i,t} + \delta WX_{i,t} + v_i + \varepsilon_{i,t} \tag{12}$$

[7] Recent advances in spatial econometrics allow to consider other issues like spatial mobility, heterogeneity and persistence via Geographically Weighted Regressions and Markov Chain Analyses in logit format (i.e. multinomial logit). However, at this stage we find it more informative to move into the background of the NEG model rather than to focus more on the spatial dimension. We prefer to delay this for a further study and start explaining how we attempt to integrate the use of micro data to the NEG framework.

$$HK_{i,t} = \alpha + \beta MA_{i,t} + \rho WHK_{i,t} + \gamma X_{i,t} + \delta WX_{i,t} + \lambda We_{i,t} + v_i + \varepsilon_{i,t} \tag{13}$$

Finally, we also take into account the possible endogeneity of market access. This has not been discussed in details within the NEG model, however omitting variables that are likely to explain human capital variable and the possible reverse causality running from human capital development towards market access are two crucial aspects that have to be considered. To deal with the endogeneity issue, we construct a number of cross-section and panel instrumental variables (IV) models, where we use different variants of sum of distances as instruments for market access (Boulhol and De Serres, 2010).

3.2 Second Step: Micro Data and Wage as an endogenous incentive for human capital accumulation

Our second attempt will be to focus more on the background mechanisms of the NEG model. In order to do so, we use individual level data obtained from the quadrennial Earnings Structure Survey (ESS) administered by TurkStat in 2010 and 2014. First administered in 2006, the survey aims to provide information on employee earnings and wages along with age, gender, tenure, occupation, education, and geographic region and the sphere of economic activity. The survey enables to produce estimates stratified by firm size, geographic region (NUTS I) and type of economic activity.

The sampling method of the ESS consists of two-stages. First stage involves the selection of the sample business establishments using stratified simple random sampling and the second-stage involves the selection of wage-earning respondents from within the sample business establishments. Both the 2010 and the 2014 ESS were conducted in business establishments with at least 10 employees and respectively administered in 20,155 and 17,137 establishments, of which 14,332 and 11,190 replied. From these establishments, data on a total of 198,375 (164,204) wage-earning employee were retrieved for the 2010 (2014) ESS.

Next we lay out our empirical strategy to test the black-box of the NEG model using individual data from the ESS. While the NEG assumes wages and skilled premium induce more human capital accumulation, the possibility that human capital accumulation drives up wages by the virtues of the Mincerian earnings equation can also be in effect and therefore wages can be endogenous. This problem requires the use of an instrument that can plausibly be viewed as randomly moving around human capital accumulation. Under a reverse causation scenario as we posit, unobservable confounders may be important in the determination of wages and human capital accumulation and thus, higher wages are likely to be correlated with the individual's education level. If the NEG model is empirically supported, then the processes that determine human capital accumulation and wages cannot be thought independent.

Our outcome of interest is the individual's education level, measured on an ordinal scale with J possible ordered outcomes, $j = 1, \ldots, J$. The outcome equation can be written as:

$$E_i = \begin{cases} 1 & \text{if} \quad -\infty < E_i^* \leq \mu_1 \\ 2 & \text{if} \quad \mu_1 < E_i^* \leq \mu_2 \\ \vdots & \qquad \vdots \\ J-1 & \text{if} \quad \mu_{J-1} < E_i^* \leq \mu_J \\ J & \text{if} \quad \mu_J < E_i^* \leq \infty \end{cases} \tag{14}$$

where μ_j's are the cutpoints. The latent outcome variable $\mathbf{E}^*$ is defined as:

$$\mathbf{E}^* = \mathbf{X}\beta + \delta\mathbf{W} + \eta\mathbf{K} + \varepsilon \tag{15}$$

where $\mathbf{W}$ is the individual's wage, $\mathbf{X}$ are the covariates of the outcome equation that include individual, employment and firm characteristics such as age, gender, tenure, type of employment, whether the individual works full-time, whether the individual has a permanent employment, weekly working hours, size of the firm, whether a collective bargaining agreement exists and quadratic terms of age and tenure, $\mathbf{K}$ consists of dummy variables on ISCO-08 occupations and on NACE Rev.2 branch of economic activity and regional dummy variables and ε is the idiosyncratic error term.

The wage equation is given by the following model:

$$\mathbf{W} = \mathbf{Z}\gamma + \theta\mathbf{MA_k} + \eta\mathbf{K} + \upsilon \tag{16}$$

where $\mathbf{Z}$ is a superset of $\mathbf{X}$ that additionally controls for excluded instruments, $\mathbf{MA}$ is the market access of the i^{th} region and υ is the error term.

Ignoring the ordered nature of the education variable for the moment, equation (15) is a reverse specification of the Mincer equation (Mincer, 1958, 1974). While Mincerian earnings function models the logarithm of earnings as the sum of years of schooling and linear and quadratic terms of labor market experience, here, we posit that wages provide an incentive to invest in human capital; hence, education is a function of wages and possibly of linear and quadratic terms of labor market experience, among others. The challenge with this view is that higher education levels are likely to drive up wages by the virtues of the Mincer equation and thus wage is likely to be endogenous in equation (15) due to reverse causality in the same spirit that schooling is endogenous in the Mincer equation. This implies that the unobservable factors that determine human capital are likely to be correlated with the unobservable factors that determine wages ($\mathbf{cov}\,(\varepsilon, \upsilon) \neq 0$). Failure to handle this reverse causality in equation (15), yields biased and inconsistent estimates.

If the correlation between the error terms of the education equation and the error terms of the wage equation, $\rho = 0$, then equation (14) can be estimated by a generalized ordered probit. If $\rho \neq 0$, then the unobservable determinants of wages are said to be correlated with the unobservable determinants of education, indicating that wage is endogenous. The possibility of endogeneity requires a joint estimation of equations (16) and (15) to obtain consistent and asymptotically efficient estimates.

To account for the possibility that wages may be endogenous to education, a source of exogenous variation should be found such that it might plausibly be viewed as randomly moving around education levels. In practice, this source of exogenous variation helps model identification. It should be (strongly) correlated with wages (i.e. relevant), should exhibit an impact on education through and only through wages and should not be directly related to education (i.e. excluded) or the latent errors, ε, of the model (i.e. clean). Primal candidates that might satisfy such properties are overtime and a dummy variable that indicates whether the individual has administrative duties. Individuals who work overtime or who have administrative duties should otherwise earn higher wages and that these two measures have no direct, evident relation to one's education level or to the unexplained factors of education levels.

4 Findings

4.1 NEG Model with Regional Data

Among different indicators of human capital development, we start by calculating the average years of schooling of Turkish regions. Figures 1 is the spatial distribution of average years of schooling. First remarkable finding is the spatial dichotomy of education. Regions clustered in the north-western and western Turkey together with a set of regions in the south and in the center forms the group of regions with the highest average education level. On the contrary for the eastern and specifically south-eastern regions we report very low levels of schooling. Moreover, the regional distribution of average years of schooling is highly persistent. Highly educated regions in 2014 are mostly the ones that were already highly educated in 1985. These findings are in line with the previous literature on human capital-based inequalities that underline the dual structure of human capital accumulation (Filiztekin and Karahasan, 2015; Erdem, 2016). Moreover it perfectly mimics the regional disparities of per capita income for Turkish regions (Dogruel and Dogruel, 2003; Gezici and Hewings, 2007).

Naturally this pattern makes one consider the spatial dimension of average years of schooling. If the pattern of regional average years of schooling is dispersed relatively more equally, one would expect to observe less spatial ties among the cross sections.[8] We calculate the Moran's I spatial auto-correlation measure to calculate the extent of spatial dependence for average years of schooling (Figure 2). Figure 2 also incorporates the continuous increase in the mean value of average years of schooling which identifies an overall improvement in education attainment. However, findings from spatial auto-correlation analysis indicate significant spatial dependence in average years of schooling, which becomes stronger during the post 2000 period. Note that we observe some weakening in the extent of spatial dependence which gives an inverted u-shaped figure for the path of spatial spillovers. That said, the short time dimension of the average years of schooling does not enable us to discuss this non-linearity in detail. Nevertheless, we find it noteworthy to remark that sample period witnesses an average improvement in schooling which is unevenly disturbed based on the extent of spatial dependence.

In order to better understand how overall spatial dependence affects local variation of schooling, we implement the LISA decomposition analyses for each year in our sample. After calculating the local LISA values for every NUTS III region, we construct a four-group distribution composed of two sets of clusters and two sets of outliers. Clusters of High-High and Low-Low represent the group of regions with high schooling and low schooling respectively. Outlier regions of High-Low (Low-High) represent the regions with high (low) schooling in close proximity to regions with lower (higher) schooling. Table 1 gives the overview of Turkish regions where we count the times for each region to be reported in one of the given groups.[9] An important finding of the LISA count analysis lies in the way high and low education clusters deviate from each other. Considering figure 1 this is naturally expected. Regions that have the highest education level continue to locate in the same group of regions during the 1985-2014 period (i.e. Istanbul, Izmir, Ankara). On the contrary, regions with the lowest average years of schooling are reported within the low education clusters continuously for the sample period (i.e. Siirt, Diyarbakır, Şanlıurfa). These two initial observations highlight the persistence of local spatialities in favor of rising and rigid human capital based inequalities. It is also interesting that a number of regions manages to deviate from their geography. For instance, Tunceli in the eastern Turkey is reported within the High-Low outlier for 7 out of the 8 years of the sample. Similarly, Sivas, Osmaniye, Kayseri, Samsun, Amasya, Malatya and Elazığ are significant examples where these regions are reported within the High-Low outliers for more than 6 years of the sample. These regions are mostly eastern or north-central regions and in close proximity to low educated and relatively less developed hinterland of Turkey.

[8] See Combes et al. (2008) for use of spatial statistics in inequality analyses.

[9] Note that we have 8 years in our sample. However, for some certain regions we have only 7 observations. These regions are: Aksaray, Bayburt, Karaman, Kırıkkale, Batman, Şırnak, Bartın, Ardahan, Iğdır, Yalova, Karabük, Kilis, Osmaniye and Düzce. These regions are formed during the 1985-2001 period and we have missing observations for these regions prior to their formation. That said, we do not think that this will create a problem as these regions share more or less similar fundamentals with the regions that they depart from.

Having seen the spatial dimension of average years of schooling we focus on how to measure market accessibility. We consider four different indicators in order to understand the ease of accessing to markets. First we calculate the market access index as offered in Harris (1954). Spatial distribution of Turkish regions' market access is illustrated in figure 3. The pattern clearly shows the clustering of market accessibility which is significantly agglomerated in the western Turkey. On the contrary, eastern regions suffer from low levels of market accessibility. Next we consider direct distance to specific economic centers as exercised in López-Rodríguez et al. (2007). We consider three specific economic areas of Turkey: Istanbul, Izmir and Ankara as these regions are historically the dominant economic activity areas of Turkey.[10] In order to see whether accessibility really matters for human capital development we run a series of simple non-spatial regressions. Figures 4, 5, 6 and 7 give the comparisons for the selected years of our sample. In all cases we detect a positive relationship between market access index and the average years of schooling. This expected result is also supported by the negative and significant relationship that we report between distance to specific economic activity areas and average years of schooling. Regardless of the economic center, being distant from metropolitan areas (Istanbul, Ankara and Izmir) reduces the average years of schooling. One would discuss how to use these four accessibility measures within the NEG model. We simply re-run the first set of simple non-spatial models by using all accessibility measures together. Results reported in table 2 indicate that market access is the dominant indicator suppressing the impact on these economic areas. This result is intuitive as market access already incorporates the impact of these three economic centers. Moreover, these results also pinpoint that it is not only a matter of proximity certain economic centers, rather both internal and external potentials of regions influence the overall impact of geographical proximity.

The relationship between market access and average years of schooling is tested by estimating a set of models. Table 3 gives a descriptive overview of the indicators for selected years used in the econometric analyses. We also report spatial auto-correlation test results for each variable used in the econometric models. In general, both market access and average years of schooling increases on average during the sample period. There is some fall in the standard deviation and spatial clustering of these two variables. Note that all variables other than the unemployment rate are spatially correlated throughout the sample period. In order to focus on the causal channels, we estimate a number of models. Results are supplied in table 4. We consider three time periods: for 1985 and 2000 we estimate cross sectional models, for 2009-2014 we estimate fixed-effects panel data models. First we estimate the non-spatial variants of the models and next we augment these models for the existence of spatial dependence. All models are conditioned on a set of regional control variables. Our first finding from non-spatial specifications indicates that market access is a significant factor affecting the average years of schooling. This finding is robust to the inclusion of a

[10] See Karahasan et al. (2016) for a similar attempt.

set of regional controls as well as to the unobserved time-invariant heterogeneity in panel models. Similarly, once we further control for spatial spillovers; we observe more or less a consistent pattern. Only for the cross sectional SAR models (for 1985 and 2000) we are unable to detect a significant relationship between market access and average years of schooling. However cross sectional SEM, SDM and SDEM as well as all spatial variants of panel models validate the significance of market access in order to explain regional schooling years. Among the control variables, service-based employment and unemployment rate significantly influence regional schooling. Note that for the panel models we are unable to use employment shares; rather we include the value added of industrial and service-oriented production. Results pinpoint that service-based production positively and significantly affects the schooling of regions with the exception of the SAR panel model. Finally, related with the spatial parameters; we detect that ρ and λ are significant in all of the SAR and SEM models. Interestingly even though we report significant ρ for the SDM specification, the spatial lag of the market access is mostly insignificant and has an impact on schooling that contradicts our expectations. In the case of SDEM this effect turns out to be in line with our expectations only for the panel specifications where unobserved time-invariant heterogeneities are considered. Note that it is possible to make an overall comparison among the spatial models to select the correct spatial specification. We calculate the LR test based on the log-likelihood values of each model. LR test challenges whether SDM and SDME models can be simplified into SAR and SEM models. As discussed in Elhorst (2010) the rejecting the null hypothesis pinpoints that SDM and SDEM models can be simplified into SAR and SEM models respectively. Our results indicate that for cross sectional models, the SDM can be simplified into the SEM in 1985 and into the SAR in 2000. For the panel results our findings indicate that the SDM best describes the data as both hypotheses are rejected. Considering the SDEM our results indicate that the SDEM can be simplified into the SAR and the SEM models for 1985 and 2000 cross sections but not for the panel model. Note that regardless of the specification of Durbin models, market access significantly explains the distribution of human capital development in Turkey.

A potential threat to the identification of the above models rests with the possibility that market access may be endogenous to human capital accumulation. Table 5 reports the instrumental variables estimates of the NEG model using regional data as an accommodation for the possible endogeneity of market access. Columns (1)-(3) and (4)-(6) respectively show the cross-sectional estimates for the years 1985 and 2000 and columns (7)-(9) show the panel estimates for the 2009-2014 period. Following Boulhol and De Serres (2010) and Karahasan and López-Bazo (2013) for all specifications in Table 5, the natural log of regional market access is instrumented by the sum of distance in columns (1) and (4), sum of distance and its squares in columns (2) and (5), inverse sum of distance in columns (3) and (6) and by the time-varying sum of distance in columns (7)-(9).

The first-stage F statistic, as a suggested measure to assess the explanatory power of the excluded instruments, reported at the bottom of the table, is well above 10 for all specifications except those reported in column (8), indicating that the instruments are not weak (Bound et al., 1995; Staiger and Stock, 1997). The strong correlation between market access and the instruments is further confirmed by the underidentifcation test results reported at the bottom of the table. However, the null hypothesis that market access is exogenous cannot be rejected at conventional test levels with the exception of column (8).

For the fact that the observations correspond to a collectively exhaustive set of provinces in Turkey, we perform a fixed-effects estimation in column (7)-(9). Column (7) controls for province fixed-effects, column (8) controls for province and year fixed-effects and column (9) performs a first-difference estimation in order to wipe out province fixed-effects that serves as a competing method to fixed-effects transformation. When T is larger than two, the fixed-effects and the first-difference transformation yield different results.

While for all cross-sectional models in Table 5, market access drives up human capital accumulation by a factor that ranges between 1.417 and 1.715, two of the three panel model specifications depict a different story. First, the province fixed-effects results are consistent with cross-sectional instrumental variables models, that market access drives up human capital accumulation. However, the impact of market access on human capital accumulation in the two-way fixed-effects model reported in column (8) suggests the opposite. Yet, the diagnostics in column (8) also show that market access cannot be treated exogenous and that the instruments are correlated with the unobservable determinants of human capital accumulation. Therefore, the results reported in column (8) are not admissible. Finally, the first-difference transformation results reported in column (9) suggest that market access is unrelated to human capital accumulation at conventional test levels. Given that market access is likely to exogenous, based on the endogeneity test results, the first-difference transformation should be preferred over the fixed-effects estimator in the presence of serially correlated errors since its differences are likely to be serially uncorrelated.[11]

The overall assessment of the results reported in Table 5 is that market access appears to be exogenous and both the province-level cross-sectional and longitudinal instruments are not weak although the overidentification tests indicate that the excluded instruments are valid for cross-sectional models but not valid for panel models (i.e. correlated with the unobserved determinants of human capital accumulation). Therefore, we have enough reason to believe that initial set of results from the traditional NEG model are indeed reliable and consistent.

[11] Note that so far we do not take into account the possible path dependence of human capital development. That is, historically developed regions accumulate more human capital, as they used to have more human capital previously. Not surprisingly, the reverse is valid for the less developed regions. We estimate dynamic variants of our panel models, mostly yielding comparable results with the panel specifications. However, considering the short time dimension of the date set, we do not find it informative to focus on this time-wise correlation of human capital development. Rather we continue to focus on the space dimension through spatial specifications.

4.2 NEG Model with Micro Data

Our second attempt is to incorporate returns to education (wages) into the NEG model. Panel A and B of Table 6 respectively report the descriptive statistics for the full sample and by the education level for the 2014 and the 2010 ESS. Overall, 19.6 (22.8) percent of the respondents had at most primary education, 16.6 (16.1) percent had primary and secondary school education, 27.3 (25.5) percent had high school education, 8.5 (9.0) percent had vocational school education and 28.0 (26.7) percent of the respondents had university education in the 2014 ESS (2010 ESS). The geographic distribution of the ordered education levels by quantile from the 2014 ESS are given in Figure 8a. The highest ordered education levels prevail in Istanbul region and the Central and Northeastern Anatolia, by and large corresponding to a high school education on average.

Both the 2014 and the 2010 ESS samples show a volatile and overly right-skewed wage distribution with a range of about 94,860 TL (121,840 TL) for the 2014 ESS (2010 ESS). Figure 9b shows the average gross monthly wage of the respondents during the year 2014. The highest average wage levels prevail in the regions of Istanbul, Western Anatolia and Eastern Marmara. These regions are also characterized by dense industrialization and irregular urbanization, especially the Eastern Marmara region.

The last figure shows the geographic distribution of market access at NUTS I level. Note that this pattern perfectly mimics the market access distribution that we already report at NUTS III level. Western geography of Turkey diverges from mostly the eastern and south-eastern Turkey suggesting the dual economic structure in terms of market accessibility. That said, spatial distribution of the education levels reported in figure 8a does not perfectly matches with the distribution of average years of schooling at NUTS III level. Partially this pattern is dominated by the relatively low levels of education of the respondents to ESS residing in the North-Western Marmara. Note that a similar pattern is also followed for the spatial distribution of wages. At this stage an early descriptive comment from the comparison of market access and education levels is that; although the Western Marmara region has a high market access, the region stands out as having low education levels and low wages compared to its surrounding regions. The last five columns in table 6 report the descriptive statistics by the education level. Expectedly, there is a clear and increasing gradient of wages and market access by education levels. There is also a clear gradient for large-sized firms and for those with permanent employment, being more prevalent at higher education levels. These two descriptive findings confirm our initial concenrs on the NEG model.

After the descriptive analysis of ESS, we challenge our second research design accordingly. As discussed, one important property of the ESS comes from the ability to use wage, education and market access variables together.[12] One challenge here is the use of wages. In line with our concerns on the endogeneity of wages, we

[12] Here note that we do not focus on the possible endogeneity of market access and human capital as we have already challenged this in the previous sub-section. A possible strategy can be to use sum of distance at NUTS I level as an instrument for market

first report the instrumental variables (IV) estimates of the education equation where the ordinal nature of education level can be safely ignored. Our sole aim is to assess the relevancy, cleanliness, excludability and the endogeneity of the instrument. Table 7 reports the results where wages are instrumented by the overtime and by a dummy variable that indicates whether the individual has administrative duties. At the bottom of the table, we report an exhaustive set of diagnostics on endogeneity, instrument relevance, weak identification and instrument validity. In order to test for instrument relevance and the endogeneity of wages, we report the heteroscedasticity consistent version of the Anderson canonical correlation LM statistic (Kleibergen-Paap LM statistic) and the endogeneity test. We further report weak identification-robust inference test results (Moreira, 2003).

Columns (1) and (3) of the table report the first-stage results obtained from a regression of wages on market access, employment characteristics, individual characteristics, the geographic location of the establishments classified by the NUTS I level and the excluded instruments. A suggested measure to assess the explanatory power of the excluded instruments is the first-stage F-statistic (Bound et al., 1995; Staiger and Stock, 1997). The first-stage F-statistic is greater than 10, suggesting that the excluded instruments are likely to be strongly correlated with wages. This is further confirmed by the underidentification test result given at the bottom of columns (2) and (4). For both ESS samples, the null hypothesis that the excluded instruments are irrelevant can be rejected at conventional test levels. The endogeneity test results confirm our expectation that wages are highly endogenous to education levels for both the 2010 and the 2014 EES data.

Having multiple instruments allows us to test for overidentifying restrictions. Instrument validity, assessed by the Hansen J statistic, indicates that the instruments are uncorrelated with the unobservable factors of education and that they are correctly excluded from the education equation. The IV diagnostics provide unequivocal evidence that the excluded instruments are valid and can be used to isolate the causal effect of wages on education.

After having seen that our concerns on endogeneity of wages and education are indeed valid and the strategy offered is accurate, we construct our final empirical specification via a recursive bivariate ordered probit model (See Table 8). We report the results separately by the education and by the wage equation. Both equations host the same control variables except that the wage equation additionally includes two excluded instruments (overtime and administrative duty) to help model identification. Note that, from the wage equation of Table 8, while administrative duties increase wages in both samples, overtime exerts a statistically significant impact on wages only in the 2010 ESS sample.

accessibility. However, we prefer to focus more on the endogeneity of wages and education as our central aim at this part of the research is to shed light on the black box behind the NEG model.

Regarding the control variables, increasing firm size and working hours increases wages but individuals with longer working hours tend to report lower education levels. Expectedly, wages are positively associated with ageing and market experience (tenure). The squares of the age and tenure variables, added to the model to capture possible non-linearities, indicate that the effects of age and tenure on wages follows an inverted bell shape. For the education equation, the effects of age and labor market experience follow a bell-shaped pattern. Full-time and permanent employment is associated with higher wages but full-time or permanent employees tend to report lower education levels compared to part-time and temporary employees. Meanwhile collective bargaining agreement (CBA) is intended to improve the socioeconomic status of the employees and physical and non-physical working conditions. Expectedly, employees of the firms that made a CBA earn higher wages by about 5.5-5.9 percent in both samples. We find yet stronger effects that wage workers of firms that entered into a CBA tend to have higher education levels compared to workers of firms without a CBA.

After observing the key characteristics of individuals we return back to our central concern. While NEG framework asserts that market access affects human capital development, we have enough reason to believe that this effect is mostly reflected over the returns to education. That is, it could be the case that it is returns of education that influences individuals' educational investment. Therefore, it is possible to see that market access' influence on education mostly originates from the signals that returns to education give to individuals. Now, as we have to isolate any potential confounding effect of market access from the effects of wages on education levels, a measure of market access is included in both education and wage equations. Since only the 2014 ESS includes geographical information, the market access variable does not appear in the education or the wage equation of the 2010 ESS. Our results show that; while a one percent increase in market access increases wages by about 0.09 percent, it does not exert a statistically distinguishable effect on education levels. This finding is crucial, as our micro level evidence show that impact of market access on education investment decision is negligible. However, wages continue to influence the education investment decisions.

In order to focus more on the impact of returns to education and instead of directly interpreting the parameter estimate of the wage variable, we report the average marginal effects (AME) at the bottom of Table 8. Since the education levels are measured on an ordinal scale, the AME is reported for every level of education among J possible ordered outcomes. The AME for the 2014 EES (2010 EES) indicates that in the full sample, higher wages decrease the probability of receiving no education and the probability of receiving a primary & second school education by about 4.4 and 1.1 percent (6.5 and 1.2 percent), respectively, and increases the probability of receiving high school, vocational school and university education by about 0.8, 0.7 and 4.0 percent (1.0, 1.0, and 5.7 percent), respectively. This initial evidence shows that even though we fail to detect significant impact of market access on education, wages continue to function as an incentive for human capital accumulation.

The size and the statistical significance of the error correlation reported at the bottom of Table 8 shows that the unobservables of the education equation and the unobservables of the wage equation are negatively correlated, confirming that wage is endogenous to education levels.

Next, to understand the differential impact of wage on education levels by geographic location, we divide the full sample by the 12 NUTS I regions of Turkey and report the regression results in Table 9. We report these results for the 2014 ESS since it is the only version of the survey for which geographic location is available.

The size and the statistical significance of all parameter estimates for the regional samples are consistent with the overall picture in the full sample. Strong, negative and significant cross-equation error correlation persists. While the parameter estimates of the log of average monthly wage is positive and statistically distinguishable from zero for all 12 regions except the Central Eastern Anatolia, the AME reported at the bottom of Table 9 shows that higher wages incentivize human capital accumulation in Istanbul and Aegean regions, Eastern Marmara and in the Mediterranean but does not affect education levels in the remaining regions. Strikingly, most of the regions for which the AME of wage are statistically indistinguishable from zero are clustered in the Eastern Turkey with a history of persistent underdevelopment. The geographic distribution of the statistically significant regional AME is mapped in Figure 9 for every ordered education level ($\alpha = 0.10$). Consequently, the largest drop in the probability of receiving no education and in the probability of receiving primary & secondary education due to higher wages are observed in the Aegean and Istanbul regions, respectively.

In Figure 9, the increase in the probability of receiving education either at the high school or at the vocational school level due to higher wages ranges between 0.3 and 1.2 percent. However, we find that the increase in the probability of receiving university education ranges between 3.9 and 5 percent and the largest effects are concentrated in the Istanbul region. The geographical distribution of the AME shows that higher wage prospects drive up human capital in the western regions of Turkey, albeit exceptions do exist. One potential explanation is that the job prospects in the Eastern and Southeastern regions of Turkey may not be sufficiently appealing for investing in human capital simply because returns to (higher) education in these regions is either non-existent or to small to detect. The absence of the feedback between earnings and human capital in these regions may create this duality.

5 Discussion

Recent advances in the NEG give momentum to regional scientists to explore different dimensions of regional development. Within this study we are highly inspired from the ability of the NEG model to discuss the endogenous accumulation of human capital. However, we approach the NEG model from a critical perspective. Our reasoning is the so-called black box within the original Redding and Schott (2003) model. While the central expectation of the construct works over the strong link between geographical proximity and firm profitability, the way this feedback influences individual's incentive to invest more in education seems blurry.

Given this concern, we are tempted to focus more on the interaction among geographical proximity (market access), human capital accumulation and returns to education. That is, we would like to make sure that the impact of market accessibility is visible both on returns to education as well as the level of human capital development. One important limitation is the inability to compile the variables of interest at the same aggregation level. Our study area is Turkey and we lack significant amount of regional data specifically for returns to education (wages). This make us follow a more holistic approach and change our research design by having the liberty to use micro and aggregate regional data sets in two separate setups. Certainly this brings methodological and measurement issues. For instance, the first set results are from spatial cross section and panel models. These models enable us to control for local interdependencies and regional time-invariant heterogeneity. On the contrary, our second set of results are from individual level data, where we lose the locality due to the representation level of the micro dataset. However, unless we use this individual level data, it would not be possible to focus on the interaction among market access, human capital accumulation and returns to education. Moreover, there seems to be a general tendency and flexibility as some new evidence from other developing country cases attempt to combine aggregate and individual level datasets. Among them, (Fally et al., 2010) combines micro and aggregate data sets for Brazil and highlight the applicability of both set of analyses in the same framework.

Based on this complexity of the research design, we first use the regional aggregate data which enables us to observe province-based market access and human capital development. That said, at the provincial level, we are unable to obtain information on the regional distribution of wages. In that sense, this first setup is a direct test of the NEG model offered by Redding and Schott (2003). Note that, unlike the original model we use different spatial specifications together with a detailed discussion on the possible endogeneity of the market access variable. Given various issues such as spatial dependence, regional heterogeneity and endogeneity evolving from possible omissions and reverse causality, our results strongly support the theoretical view that central provinces accumulate more human capital.

In the second part of our analyses, we use the ESS that provides information at the individual level. Education level and wages can be collected for each individual in the survey. For the regional dimension, we are only able to collect information at the NUTS I level. This brings a new challenge as the first set of analyses are from NUTS III level provinces. However, using different aggregation levels should not bring a bias or confusion because at the end, we plan to have a holistic view of the interconnection among geography, human capital accumulation and returns to education. Our results are interesting. First, we realize that wage distribution is strongly and positively related with market accessibility. This validates the sub-channel of the NEG model which we were unable to check during the first set of analyses. That said, our results highlight that this strong tie does not significantly alter individuals' decision to invest in education. Results show that market access and education level are not significantly related at the micro level. This finding somehow matches with our critical approach to the black-box of the NEG model.

It is apparent that wage distribution is an important element to influence individuals' decision on education. In that sense the NEG framework has a lot to offer to policy makers, as the model expects that regions with higher market access offer higher wages. Naturally we would expect to see higher human capital accumulation in those regions with higher market accessibility. However, we find that even there tends to be a strong co-movement among these variables, it seems difficult to identify a causal channel running from market access to human capital accumulation. It seems that the causal channel explains the distribution of the wages but not the education. We therefore underline that regional policy makers should focus more on individual characteristics of regions and people in order to understand how market accessibility can bring more incentives to individuals in order to accumulate more human capital for new generations. In a way given that causal influence varies based on education level and the aggregate geographic region (NUTS I), more policy flexibility will be essential in the future.

6 Conclusion

Our central objective is to question the impact of geography (measured by market access) on human capital development differences among the regions of Turkey. While doing this, we aim at moving a step further by discussing the black-box (the so-called missing link) within the original NEG model. As we certainly would like to understand whether people increase their human capital levels by investing more in education because they locate in central locations; or is it the returns to education that motivates their incentives to invest in human capital. As the former case is more policy insensitive we underline the need for this decomposition.

We have conducted two sets of analyses. First set of results are from a regional dataset and confirms that market access significantly influences the regional human capital development in Turkey. These results are robust to the inclusion of sectoral control (regional structural differences), regional heterogeneities (fixed effect panel models), spatial spillovers (various spatial models) and endogeneity of market access. In that sense, it would not be naive to highlight that individuals residing in central (core) areas invest more in human capital keeping aside the possible impact of migration. Even these results are vital, we are unable to visualize the true background since we have no information on the distribution of returns to education (wages) at the province level.

Our second set of results is from a unique micro dataset which enables us to construct a set of models which includes human capital development, market access and returns to education. Results show that inclusion of wages changes the whole story. Since, we were unable to control for wages at the NUTS III aggregation in the first set of analysis, we could only construct an indirect link between market access and human capital accumulation without knowing the transmission across market access and incentive to accumulate human capital (wages). That said, we have shown that wages-human capital link which is regarded as the black-box in the initial set of analyses basically wipes out the impact of market access on human capital accumulation. Remarkably we identify a positive link between market access and wages; however, we are unable to report any significant connection between market access and education. Therefore, even if wages have had an influence on education accumulation this would not necessarily imply that the background is the market accessibility.

References

Anderson, James E and Eric Van Wincoop (2004): Trade costs. *Journal of Economic Literature*, 42(3):691–751.

Anselin, L. (1995): Local indicators of spatial association-LISA. *Geographical Analysis*, 27(2):93–115.

Anselin, L. (1996): The Moran scatterplot as an ESDA tool to assess local instability in spatial association. *Spatial analytical perspectives on GIS*, 111:111–125.

Anselin, L. (2010): Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1):3–25.

Bosker, M. and H. Garretsen (2010): Trade costs in empirical new economic geography. *Papers in Regional Science*, 89(3):485–511.

Boulhol, H. and A. De Serres (2010): Have developed countries escaped the curse of distance? *Journal of Economic Geography*, 10(1):113–139.

Bound, J.; D.A. Jaeger; and R.M. Baker (1995): Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450.

Brakman, S.; H. Garretsen; and M. Schramm (2006): Putting new economic geography to the test: Free-ness of trade and agglomeration in the EU regions. *Regional Science and Urban Economics*, 36(5):613–635.

Breinlich, H. (2006): The spatial income structure in the European Union what role for Economic Geography? *Journal of Economic Geography*, 6(5):593–617.

Celebioglu, F. and S. Dall-Erba (2010): Spatial disparities across the regions of Turkey: an exploratory spatial data analysis. *The Annals of Regional Science*, 45(2):379–400.

Coe, D.T. and E. Helpman (1995): International r&d spillovers. *European Economic Review*, 39(5):859–887.

Coe, D.T.; E. Helpman; and A. Hoffmaister (1995): *North-south R&D spillovers*. Tech. rep., National Bureau of Economic Research.

Combes, P.; T. Mayer; and J. Thisse (2008): *Economic geography: The integration of regions and nations*. Princeton University Press.

Dogruel, F. and S. Dogruel (2003): *Iktisat Uzeirne Yazilar I Kuresel Duzen: Birikim, Devlet ve Siniflar Korkut Boratav'a Armagan*, Iletisim Publications, chap. Turkiye'de Bolgesel Gelir Farkliliklari ve Buyume.

Duranton, G. and H.G. Overman (2005): Testing for localization using micro-geographic data. *The Review of Economic Studies*, 72(4):1077–1106.

Duranton, G. and D. Puga (2004): *Handbook of regional and urban economics Volume:4 Cities and Geography*, Elsevier, chap. Micro-foundations of urban agglomeration economies, pp. 2063–2117.

Elbadawi, I.; T. Mengistae; T. Temesge; and A. Zeufack (2009): Economic geography and manufacturing productivity in Africa: ananalysis of firm level data. *The Journal of Developing Areas*, 42(2):223–252.

Elhorst, J.P. (2010): Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 5(1):9–28.

Elveren, A.Y. (2010): Wage inequality in Turkey: Decomposition by statistical regions, 1980–2001. *Review of Urban & Regional Development Studies*, 22(1):55–72.

Erdem, U. (2016): Regional human capital distribution and disparities in Turkey. *Review of Urban & Regional Development Studies*, 28(1):16–31.

Faíña, J A. and J. Lopez-Rodriguez (2006): Market access and human capital accumulation: the European Union case. *Applied Economics Letters*, 13(9):563–567.

Fally, T.; R. Paillacar; and C. Terra (2010): Economic geography and wages in Brazil: Evidence from micro-data. *Journal of Development Economics*, 91(1):155–168.

Filiztekin, A. (1998): Convergence across industries and provinces in Turkey. *Koc University, http://myweb. sabanciuniv. edu/alpayf/files/2010/04/turkconv981. pdf.*

Filiztekin, A. and B.C. Karahasan (2015): Mapping the Educational Attainment in Turkey. In: *Proceedings of Ekonomik Yaklaşımlar International Congress of Economics II-Growth, Inequality and Poverty.* No. 285.

Fingleton, B. (2008): Competing models of global dynamics: evidence from panel models with spatially correlated error components. *Economic Modelling*, 25(3):542–558.

Fujita, M.; P.R. Krugman; and A.J. Venables (1999): *The spatial economy: cities, regions and international trade.* MIT Press.

Gallup, J.L.; J.D. Sachs; and A.D. Mellinger (1999): Geography and economic development. *International Regional Science Review*, 22(2):179–232.

Gezici, F. and G.J.D. Hewings (2007): Spatial analysis of regional inequalities in Turkey. *European Planning Studies*, 15(3):383–403.

Hanson, G.H. (2005): Market potential, increasing returns and geographic concentration. *Journal of International Economics*, 67(1):1–24.

Harris, C.D. (1954): The, Market as a Factor in the Localization of Industry in the United States. *Annals of the association of American geographers*, 44(4):315–348.

Head, K. and T. Mayer (2006): Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics*, 36(5):573–594.

Head, K. and T. Mayer (2011): Gravity, market potential and economic development. *Journal of Economic Geography*, 11(2):281–294.

Hering, L. and S. Poncet (2010): Market access and individual wages: Evidence from China. *The Review of Economics and Statistics*, 92(1):145–159.

Karahasan, B.C.; F. Dogruel; and S. Dogruel (2016): Can Market Potential Explain Regional Disparities in Developing Countries? Evidence from Turkey. *The Developing Economies*, 54(2):162–197.

Karahasan, B.C. and E. López-Bazo (2013): The spatial distribution of human capital: can it really be explained by regional differences in market access? *International Regional Science Review*, 36(4):451–480.

Keller, W. (January 2000): *Geographic Localization of International Technology Diffusion*. Working Paper 7509, National Bureau of Economic Research.

Kosfeld, R. and H. Eckey (2010): Market access, regional price level and wage disparities: the German case. *Jahrbuch für Regionalwissenschaft*, 30(2):105–128.

Krugman, P. (1991): Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–499.

Krugman, P. and A.J. Venables (1995): Globalization and the Inequality of Nations. *The Quarterly Journal of Economics*, 110(4):857–880.

López-Rodríguez, J.; J.A. Faíña; and J. López-Rodríguez (2007): Human capital accumulation and geography: empirical evidence from the European Union. *Regional Studies*, 41(2):217–234.

Lucas, R. (1988): On the mechanisms of development planning. *Journal of Monetary Economics*, 22(1):3–42.

Mincer, J. (1958): Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4):281–302.

Mincer, J. (1974): Schooling, Experience, and Earnings. *New York: NBER Press.*

Mion, G. (2004): Spatial externalities and empirical analysis: the case of Italy. *Journal of Urban Economics*, 56(1):97–118.

Moreira, M.J. (2003): A conditional likelihood ratio test for structural models. *Econometrica*, 71(4):1027–1048.

Motellón, E.; E. López-Bazo; and M. El-Attar (2011): Regional heterogeneity in wage distributions: evidence from Spain. *Journal of Regional Science*, 51(3):558–584.

Niebuhr, A. (2006): Market access and regional disparities. *The Annals of Regional Science*, 40(2):313–334.

Ottaviano, G.I.P. and D. Pinelli (2006): Market potential and productivity: evidence from Finnish regions. *Regional Science and Urban Economics*, 36(5):636–657.

Redding, S. and P.K. Schott (2003): Distance, skill deepening and development: will peripheral countries ever get rich? *Journal of Development Economics*, 72(2):515–541.

Redding, S. and A. Venables (2004a): Geography and export performance: external market access and internal supply capacity. In: *Challenges to globalization: Analyzing the economics*, University of Chicago Press, pp. 95–130.

Redding, S. and A.J. Venables (2004b): Economic geography and international inequality. *Journal of international Economics*, 62(1):53–82.

Redding, S.J. (2010): The empirics of new economic geography. *Journal of Regional Science*, 50(1):297–311.

Romer, P.M. (1994): The origins of endogenous growth. *The journal of Economic Perspectives*, 8(1):3–22.

Solow, R.M. (1956): A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65–94.

Staiger, D. and J.H. Stock (1997): Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3):557–586.

Stock, J.H. and M. Yogo (2005): Testing for Weak Instruments in Linear IV Regression. *in: Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.*

Venables, A.J. (1996): Equilibrium locations of vertically linked industries. *International Economic Review*, pp. 341–359.

Yildirim, J.; N. Öcal; and S. Özyildirim (2009): Income inequality and economic convergence in Turkey: A spatial effect analysis. *International Regional Science Review*, 32(2):221–254.

Table 1: LISA Clusters and Persistance (1985-2014)

		High-High	**Low-Low**	**High-Low**	**Low-High**	**Total**
TR100	Istanbul	8	0	0	0	8
TR211	Tekirdag	8	0	0	0	8
TR212	Edirne	8	0	0	0	8
TR213	Kirklareli	8	0	0	0	8
TR221	Balikesir	8	0	0	0	8
TR222	Canakkale	8	0	0	0	8
TR310	Izmir	8	0	0	0	8
TR321	Aydin	8	0	0	0	8
TR322	Denizli	8	0	0	0	8
TR323	Mugla	8	0	0	0	8
TR331	Manisa	6	0	0	2	8
TR332	Afyonkarahisar	7	0	0	1	8
TR333	Kutahya	8	0	0	0	8
TR334	Usak	8	0	0	0	8
TR411	Bursa	8	0	0	0	8
TR412	Eskisehir	8	0	0	0	8
TR413	Bilecik	8	0	0	0	8
TR421	Kocaeli	8	0	0	0	8
TR422	Sakarya	8	0	0	0	8
TR423	Duzce	7	0	0	0	7
TR424	Bolu	8	0	0	0	8
TR425	Yalova	7	0	0	0	7
TR510	Ankara	8	0	0	0	8
TR521	Konya	8	0	0	0	8
TR522	Karaman	7	0	0	0	7
TR611	Antalya	8	0	0	0	8
TR612	Isparta	8	0	0	0	8
TR613	Burdur	8	0	0	0	8
TR621	Adana	3	0	5	0	8
TR622	Mersin	8	0	0	0	8
TR631	Hatay	0	7	1	0	8
TR632	Kahramanmaras	0	6	0	2	8
TR633	Osmaniye	0	2	5	0	7
TR711	Kirikkale	7	0	0	0	7
TR712	Aksaray	0	0	0	7	7
TR713	Nigde	0	0	0	8	8
TR714	Nevsehir	6	0	2	0	8
TR715	Kirsehir	8	0	0	0	8
TR721	Kayseri	0	0	8	0	8
TR722	Sivas	1	1	6	0	8
TR723	Yozgat	0	0	0	8	8
TR811	Zonguldak	7	0	1	0	8
TR812	Karabuk	7	0	0	0	7
TR813	Bartin	0	0	0	7	7
TR821	Kastamonu	0	3	0	5	8
TR822	Cankiri	8	0	0	0	8
TR823	Sinop	0	8	0	0	8
TR831	Samsun	0	2	6	0	8
TR832	Tokat	0	6	0	2	8
TR833	Corum	0	0	0	8	8
TR834	Amasya	0	0	8	0	8
TR901	Trabzon	6	0	2	0	8
TR902	Ordu	0	6	0	2	8
TR903	Giresun	2	1	0	5	8
TR904	Rize	8	0	0	0	8
TR905	Artvin	0	0	8	0	8
TR906	Gumushane	6	0	0	2	8
TRA11	Erzurum	0	8	0	0	8
TRA12	Erzincan	4	0	4	0	8
TRA13	Bayburt	5	0	0	2	7
TRA21	Agri	0	8	0	0	8
TRA22	Kars	0	8	0	0	8
TRA23	Igdir	0	7	0	0	7
TRA24	Ardahan	0	7	0	0	7
TRB11	Malatya	0	1	7	0	8
TRB12	Elazig	0	1	7	0	8
TRB13	Bingol	0	8	0	0	8
TRB14	Tunceli	0	1	7	0	8
TRB21	Van	0	8	0	0	8
TRB22	Mus	0	8	0	0	8
TRB23	Bitlis	0	8	0	0	8
TRB24	Hakkari	0	8	0	0	8
TRC11	Gaziantep	0	8	0	0	8
TRC12	Adiyaman	0	8	0	0	8
TRC13	Kilis	0	7	0	0	7
TRC21	Sanliurfa	0	8	0	0	8
TRC22	Diyarbakir	0	8	0	0	8
TRC31	Mardin	0	8	0	0	8
TRC32	Batman	0	7	0	0	7
TRC33	Sirnak	0	7	0	0	7
TRC34	Siirt	0	8	0	0	8

Table 2: Average Years of Schooling and Accessibility

	1985	2000	2009	2014
Market	0.414***	0.222*	0.3607***	0.207**
Access	(0.150)	(0.122)	(0.1324)	(0.080)
Distance to	-0.025	-0.044	-0.0207	-0.019
Istanbul	(0.056)	(0.045)	(0.040)	(0.025)
Distance to	-0.0124	-0.0189	-0.008	0.003
Ankara	(0.049)	(0.038)	(0.0379)	(0.023)
Distance to	-0.028	-0.043	-0.0471*	-0.026
Izmir	(0.035)	(0.027)	(0.0252)	(0.015)
R^2	0.57	0.46	0.52	0.47
Obs.	64	78	78	78

Notes: ***, **, * indicates significance at 1%, 5% and 10%
Standard Errors in ( )

Table 3: Descriptive Analysis for Regional Analysis

	1985				2000			
	mean (sd.)	**min**	**max**	**Moran's I (sd)**	**mean (sd.)**	**min**	**max**	**Moran's I (sd)**
Market Access (ln)	13.405 (0.329)	12.578	14.292	0.831*** (0.077)	19.435 (0.295)	18.736	20.263	0.862*** (0.071)
Average Years of Schooling	3.567 (0.757)	2.035	5.306	0.716*** (0.078)	4.753 (0.799)	3.046	6.65	0.710*** (0.072)
Manufacturing Emp. (%)	0.072 (0.059)	0.005	0.342	0.460*** (0.074)	0.079 (0.064)	0.007	0.316	0.516*** (0.071)
Service Emp. (%)	0.059 (0.035)	0.018	0.249	0.291*** (0.071)	0.083 (0.043)	0.022	0.268	0.303*** (0.069)
Unemployment Rate	0.041 (0.0167)	0.02	0.09	0.019 (0.064)	0.078 (0.031)	0.035	0.174	-0.006 (0.046)
Number of Observations	67				81			
	2009				**2014**			
	mean (sd.)	**min**	**max**	**Moran's I (sd)**	**mean (sd.)**	**min**	**max**	**Moran's I (sd)**
Market Access (ln)	21.383 (0.248)	20.754	22.012	0.840*** (0.071)	22.045 (0.245)	21.42	22.696	0.829*** (0.071)
Average Years of Schooling	5.424 (0.861)	3.092	7.217	0.758*** (0.071)	6.91 (0.702)	5.036	8.45	0.688*** (0.071)
Industrial Output (%)	0.203 (0.092)	0.056	0.475	0.463*** (0.072)	0.249 (0.090)	0.09	0.503	0.462*** (0.072)
Service Output (%)	0.523 (0.072)	0.347	0.674	0.254*** (0.072)	0.497 (0.072)	0.324	0.691	0.227*** (0.072)
Number of Observations	81				81			

Notes: *** represents significance at 1%, standard errors are in ( )

Table 4: Testing the NEG Model

	Non-spatial			Spatial Lag Model			Spatial Error Model			Spatial Durbin Model			Spatial Durbin Error Model		
	1985	2000	2009-2014	1985	2000	2009-2014	1985	2000	2009-2014	1985	2000	2009-2014	1985	2000	2009-2014
Market	0 .411***	0.199***	0.763***	0.089	0.050	0.096***	0.266**	0.229***	0.769***	0.253**	0.271**	0.463***	0.250**	0.242**	0.383***
Access	(0.070)	(0.062)	(0.026)	(0.066)	(0.052)	(0.016)	(0.108)	(0.078)	(0.095)	(0.130)	(0.123)	(0.199)	(0.123)	(0.113)	(0.143)
Manufacturing	-0.248	-0.059		0.075	-0.102		0.313	0.108		0.106	0.013		-0.065	-0.054	
Employment	(0.617)	(0.291)		(0.434)	(0.220)		(0.412)	(0.219)		(0.417)	(0.232)		(0.459)	(0.252)	
Service	2.630***	2.203***		1.698**	1.557***		1.464**	1.336***		2.138***	1.430***		2.744***	1.765***	
Employment	(0.966)	(0.419)		(0.689)	(0.331)		(0.587)	(0.340)		(0.695)	(0.340)		(0.807)	(0.358)	
Unemployment	-2.824**	-2.021***		-0.297	-0.601		0.234	-0.597		-0.570	-0.470		-1.530	-0.737	
Rate	81.208)	(0.464)		(0.913)	(0.409)		(0.924)	(0.504)		(1.076)	(0.508)		(1.216)	(0.481)	
Coastal	0.115*	0.004		0.044	0.0006		0.068*	0.025		0.080*	0.027		0.087*	0.017	
Dummy	(0.063)	(0.048)		(0.045)	(0.036)		(0.041)	(0.036)		(0.045)	(0.039)		(0.051)	(0.042)	
Elevation	0.018	0.001		0.009	0.003		0.014	0.004		0.019	0.004		0.021	0.003	
	(0.018)	(0.013)		(0.012	(0.009)		(0.011)	(0.009)		(0.012)	(0.010)		(0.015)	(0.012)	
Industry % in			0.291**			0.028			0.042			0.042			0.072
GDP			(0.124)			(0.059)			(0.065)			(0.064)			(0.066)
Service % in			0.280***			0.062			0.265***			0.204***			0.160***
GDP			(0.082)			(0.038)			(0.063)			(0.072)			(0.061)
ρ				0.705***	0.631***	0.879***				0.758***	0.592***	0.882***			
				(0.095)	(0.094)	(0.020)				(0.091)	(0.108)	(0.020)			
λ							0.850***	0 .721***	0.889***				0.767***	0.622***	0.888***
							(0.073)	(0.093)	(0.019)				(0.081)	(0.101)	(0.018)
W*Market										-.0246	-0.282*	-0.415*	-0.026	-0.138	0.351**
Access										(0.172)	(0.153)	(0.206)	(0.212)	(0.169)	(0.149)
W*rhs inc.	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.69	0.66	0.57	0.65	0.68	0.65	0.66	0.64	0.57	0.78	0.74	0.56	0.76	0.74	0.58
Obs.	67	81	486	67	81	486	67	81	486	67	81	486	67	81	486
LL	45.93462	70.42917	1028.316	61.282096	85.173126	1396.562	65.363867	84.360776	1400.751	69.937559	89.973879	1407.555	69.6499	90.3008	1137.5799
AIC	-77.86924	-126.8583	-2048.632	-104.5642	-152.3463	-2783.123	-112.7277	-150.7216	-2791.501	-109.8751	-149.9478	-2799.11	-109.2997	-150.6017	-2249.819
LR-Test										17.31***	9.6	21.99***			
SAR vs. SDM										[0.00]	[0.14]	[0.00]			
LR-Test										9.15	11.23*	13.61***			
SEM vs. SDM										[0.17]	[0.08]	[0.00]			
LR-Test													8.57	10.47	15.34***
SEM vs. SDEM													[0.1991]	[0.1061]	[0.0015]

Notes: ***, **, * represents significance at 1%, 5% and 10% respectively. Standard errors are in paranthesis, P-values are in brackets

Table 5: Instrumental variables estimates: The NEG model

Outcome variable	Average years of schooling								
	cross-section 1985			cross-section 2000			panel 2009-2014		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Market access	1.692*** (0.655)	1.644** (0.722)	1.715*** (0.666)	1.418*** (0.443)	1.417*** (0.436)	1.469*** (0.434)	4.687*** (0.106)	-3.137*** (1.162)	-0.162 (0.236)
Manufacturing employment	-2.435 (3.270)	-2.227 (3.442)	-2.497 (3.259)	-0.614 (1.643)	-0.686 (1.654)	-0.731 (1.615)	-	-	-
Service employment	10.821* (6.211)	10.906* (6.455)	10.680* (6.188)	11.355*** (2.578)	11.277*** (2.525)	11.238*** (2.327)	-	-	-
Unemployment rate	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-	-	-
Coastal dummy	0.480* (0.261)	0.480* (0.272)	0.485* (0.249)	0.242 (0.201)	0.213 (0.198)	0.248 (0.202)	-	-	-
Elevation	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-	-	-
Constant	-0.198 (1.174)	-0.117 (1.293)	-0.236 (1.196)	-0.424 (1.251)	-0.374 (1.253)	-0.564 (1.262)	-	-	-
Industry % in GDP	-	-	-	-	-	-	0.800 (0.518)	-0.241 (0.601)	-0.937** (0.367)
Service % in GDP	-	-	-	-	-	-	2.524*** (0.190)	-0.602 (0.615)	-1.352*** (0.325)
Number of observations	67	67	67	81	81	81	486	486	405
First-stage F statistic	21.11 [0.0000]	11.19 [0.0000]	23.11 [0.0000]	55.35 [0.0000]	27.91 [0.0000]	52.62 [0.0000]	378.48 [0.0000]	6.88 [0.0000]	322.02 [0.0000]
Underidentification test	12.216 [0.0005]	13.639 [0.0011]	13.253 [0.0003]	22.309 [0.0000]	22.723 [0.0000]	22.368 [0.0000]	37.347 [0.0000]	21.780 [0.0000]	68.459 [0.0000]
Redundancy test	12.216 [0.0005]	13.639 [0.0011]	13.253 [0.0003]	22.309 [0.0000]	22.723 [0.0000]	22.368 [0.0000]	37.347 [0.0000]	21.780 [0.0000]	68.459 [0.0000]
Hansen J statistic	-	0.0989 [0.7531]	-	-	0.7361 [0.3909]	-	74.768 [0.0000]	14.654 [0.0055]	72.188 [0.0000]
Endogeneity test	0.3952 [0.5296]	0.3394 [0.5602]	0.4293 [0.5123]	0.6105 [0.4346]	0.6326 [0.4264]	0.8713 [0.3506]	0.1035 [0.7476]	3.6190 [0.0571]	0.0693 [0.7924]
Weak identification test	21.110	11.191	23.108	55.351	27.909	52.622	378.48	6.8819	322.02
Centered R^2	0.7226	0.7241	0.7217	0.6464	0.6457	0.6446	0.9124	0.9694	0.0155
Error component	NA	NA	NA	NA	NA	NA	Province FE	Province & year FE	FD

Notes: The unit of observation is the province. The natural log of regional market access is instrumented by the sum of distance in columns (1) and (4), sum of distance and its squares in columns (2) and (5), inverse sum of distance in columns (3) and (6) and by the time-varying sum of distance in columns (7)-(9). The underidentification test reports the Kleibergen-Paap rk LM statistic and the p-value for the null hypothesis that the equation is underidentified (i.e. the excluded instruments are irrelevant). The weak identification test reports the Cragg–Donald Wald F-statistic and the p-value for the null hypothesis that the equation is weakly identified. Hansen J statistic reports the chi-square and the p-value for the joint null hypothesis that the instruments are uncorrelated with the error term and that the excluded instruments are correctly excluded from the estimated equation (i.e. the instruments are valid). The redundancy test reports the chi-square and the p-value for the null hypothesis that the instruments are redundant. The endogeneity test reports the chi-square and the p-value for the null hypothesis that regional market access is exogenous. Bootstrapped standard errors (500 replications) in parentheses in columns (1)-(6) are clustered at the province-level and are robust to arbitrary heteroscedasticity. Standard errors in columns (7)-(9) are clustered at the province level. *, ** and *** denote statistical significance at the 10, 5 and 1 percent level respectively. FE: Fixed-effects; FD: First-difference.

Table 6: Descriptive statistics, Earnings Structure Survey

Panel A: ESS 2014				Education level				
N = 164, 204	mean (s.d)	min	max	less than primary	Primary & Secondary	High school	Vocational school	University
Education	3.09 (1.47)	1	5	-	-	-	-	-
Average monthly wage (TL)	1759.5 (2287.1)	3.1	94862.2	1176.5 (736.6)	1089.9 (819.7)	1283.5 (1128.4)	1642.8 (1201.6)	3065.4 (3714.5)
Market access	1323.06 (624.38)	321.83	2276.53	1279.8 (602.39)	1296.91 (622.60)	1281.26 (630.18)	1354.44 (547.43)	1400.11 (648.66)
Collective bargaining agreement (=1 if yes; =0 otherwise)	0.11 (0.31)	0	1	0.09 (0.28)	0.08 (0.27)	0.08 (0.26)	0.22 (0.41)	0.13 (0.33)
Firm size	2.61 (1.45)	1	5	2.26 (1.33)	2.40 (1.39)	2.69 (1.48)	2.80 (1.42)	2.84 (1.46)
Hours/week worked	44.20 (4.17)	1	60	44.51 (3.83)	44.63 (3.15)	44.31 (4.01)	44.68 (2.87)	43.46 (5.20)
Permanent employment (=1 if permanent; =0 otherwise)	0.92 (0.27)	0	1	0.89 (0.31)	0.90 (0.29)	0.92 (0.27)	0.94 (0.23)	0.93 (0.25)
Full-time employment (=1 if full-time; =0 otherwise)	0.98 (0.13)	0	1	0.99 (0.10)	0.99 (0.09)	0.98 (0.15)	0.99 (0.09)	0.98 (0.15)
Gender (=1 if male; =0 otherwise)	0.74 (0.44)	0	1	0.81 (0.39)	0.83 (0.37)	0.72 (0.45)	0.82 (0.38)	0.63 (0.48)
Age	34.2 (9.27)	14	85	41.15 (7.97)	31.52 (9.42)	32.10 (8.56)	32.87 (8.50)	33.35 (8.53)
Tenure (years)	3.16 (4.46)	0	43	3.45 (4.69)	2.37 (3.79)	2.77 (4.07)	4.39 (5.46)	3.42 (4.55)
Administrative duty (=1 if yes; =0 otherwise)	0.13 (0.34)	0	1	0.06 (0.23)	0.06 (0.23)	0.11(0.31)	0.12 (0.33)	0.26 (0.44)
Overtime (hours)	4.89 (14.12)	0	210	5.76 (15.76)	6.16 (15.87)	4.50 (13.35)	9.03 (18.65)	2.64 (9.87)

Panel B: ESS 2010				Education level				
N = 198, 375	mean (s.d)	min	max	less than primary	Primary & Secondary	High school	Vocational school	University
Education	3.01 (1.49)	1	5	-	-	-	-	-
Average monthly wage (TL)	1215.1 (1632.1)	2	121841.1	782.9 (500.3)	739.5 (562.8)	957.3 (852.3)	1122.1 (917.7)	2148.3 (2714.2)
Collective bargaining agreement (=1 if yes; =0 otherwise)	0.08 (0.27)	0	1	0.06 (0.24)	0.05 (0.21)	0.06 (0.23)	0.14 (0.35)	0.11 (0.31)
Firm size	2.12 (1.41)	1	5	1.79 (1.17)	1.82 (1.21)	2.19 (1.44)	2.25 (1.46)	2.48 (1.55)
Hours/week worked	44.25 (3.96)	1	78	44.75 (3.13)	44.72 (3.50)	44.37 (3.90)	44.47 (3.28)	43.37 (4.88)
Permanent employment (=1 if permanent; =0 otherwise)	0.96 (0.18)	0	1	0.95 (0.22)	0.95 (0.22)	0.97 (0.18)	0.98 (0.15)	0.98 (0.13)
Full-time employment (=1 if full-time; =0 otherwise)	0.99 (0.12)	0	1	0.99 (0.10)	0.99 (0.11)	0.98 (0.12)	0.99 (0.10)	0.98 (0.14)
Gender (=1 if male; =0 otherwise)	0.75 (0.43)	0	1	0.85 (0.36)	0.84 (0.36)	0.72 (0.45)	0.80 (0.40)	0.62 (0.49)
Age	33.88 (9.00)	14	97	39.12 (7.84)	31.24 (9.24)	31.95 (8.37)	32.23 (8.28)	33.39 (8.81)
Tenure (years)	3.49 (4.99)	0	55	3.57 (4.82)	2.70 (4.38)	3.33 (5.00)	4.41 (5.75)	3.75 (5.12)
Administrative duty (=1 if yes; =0 otherwise)	0.13 (0.34)	0	1	0.05 (0.23)	0.05 (0.22)	0.11 (0.31)	0.13 (0.33)	0.28 (0.45)
Overtime (hours)	4.53 (14.17)	0	250	4.72 (14.60)	4.83 (15.20)	5.02 (14.24)	7.84 (19.12)	2.60 (10.36)

Note: Firm size is coded as an ordered variable with the following values: =1 if 10<size<49, =2 if 50<size<249, =3 if 250<size<499, =4 if 500<size<999, =5 if size>1000.

Table 7: Instrumental variables estimates, Earnings Structure Survey (ESS)

	ESS 2014		ESS 2010	
	First-stage	Second-stage	First-stage	Second-stage
Constant	4.128*** (0.165)	-3.465*** (0.401)	4.152*** (0.050)	-3.262*** (0.370)
Log (average monthly wage)	-	0.882*** (0.078)	-	1.778*** (0.087)
Log (market access)	-0.075*** (0.025)	0.558*** (0.040)	-	-
Permanent employment	0.370*** (0.021)	-0.221*** (0.037)	0.251*** (0.016)	-0.339*** (0.039)
Full-time employment	1.091*** (0.039)	-1.009*** (0.097)	0.923*** (0.031)	-1.835*** (0.100)
Gender	0.062*** (0.008)	-0.061*** (0.013)	0.033*** (0.006)	-0.103*** (0.013)
Age	0.052*** (0.003)	-0.090*** (0.006)	0.040*** (0.002)	-0.148*** (0.005)
Age squared	-0.0006*** (0.00003)	0.001*** (0.000)	-0.0004*** (0.00003)	0.001*** (0.000)
Tenure	0.159*** (0.003)	-0.127*** (0.013)	0.126*** (0.002)	-0.208*** (0.012)
Tenure squared	-0.006*** (0.0002)	0.005*** (0.000)	-0.003*** (0.0001)	0.006*** (0.000)
Administrative duty	0.219x*** (0.011)	-	0.214*** (0.009)	-
Overtime	0.002*** (0.0002)	-	0.001*** (0.0001)	-
NUTS regions				
Istanbul	0.307*** (0.045)	-1.106*** (0.073)	-	-
West Marmara	0.151*** (0.037)	-0.744*** (0.060)	-	-
Aegean	0.129*** (0.026)	-0.543*** (0.043)	-	-
East Marmara	0.194*** (0.033)	-0.545*** (0.054)	-	-
West Anatolia	0.159*** (0.025)	-0.540*** (0.041)	-	-
Mediterranean	0.140*** (0.020)	-0.383*** (0.032)	-	-
Central Anatolia	0.082*** (0.022)	-0.410*** (0.035)	-	-
West Black Sea	0.075*** (0.033)	-0.611*** (0.053)	-	-
East Mediterranean	0.107*** (0.026)	-0.373*** (0.043)	-	-
Northeastern Anatolia	-0.005*** (0.002)	0.000 (0.003)	-	-
Central Eastern Anatolia	-	-	-	-
Southeastern Anatolia	-	-	-	-
First-stage F-statistic	240.08 [0.0000]	-	286.03 [0.0000]	-
Underidentification test	-	449.46 [0.0000]	-	526.62 [0.0000]
Redundancy test	-	338.18 [0.0000]	-	462.38 [0.0000]
Hansen J statistic	-	0.8919 [0.3450]	-	0.4404 [0.5069]
Conditional LR	-	136.08 [0.0000]	-	561.07 [0.0000]
Endogeneity test	-	90.729 [0.0000]	-	420.34 [0.0000]
Weak identification test	-	240.08	-	286.03
Number of observations	164,023	164,023	198,161	198,161
Centered $\mathbf{R}^2$	0.4455	0.4105	0.4142	0.0128

Notes: The unit of observation is the individual. The outcome variable is the level of education (ordered). The natural log of average monthly wage is instrumented by overtime and whether the individual has administrative duties. Underidentification test reports the Kleibergen-Paap rk LM statistic and the p-value for the null hypothesis that the equation is underidentified (i.e. the excluded instruments are irrelevant). Weak identification test reports the Cragg–Donald Wald F-statistic and the p-value for the null hypothesis that the equation is weakly identified. Stock and Yogo (2005) weak identification test critical values for 10% and 15% maximal IV size are 5.44 and 3.87 respectively. Moreira (2003)'s conditional likelihood ratio reports the weak identification-robust inference likelihood ratio and the p-value for the null hypothesis that the coefficient of wage is zero. Hansen J statistic reports the chi- square and the p-value for the joint null hypothesis that the instruments are uncorrelated with the error term and that the excluded instruments are correctly excluded from the estimated equation (i.e. the instruments are valid). The redundancy test reports the chi- square and the p-value for the null hypothesis that the instruments are redundant. The endogeneity test reports the chi-square and the p-value for the null hypothesis that wage is exogenous. Standard errors in parentheses are clustered at the individuul level and are robust to arbitrary heteroscedasticity. All specifications use sampling weights provided by TurkStat and include $9-1$ dummy variables on ISCO-08 occupations and $17-1$ dummy variables on statistical classification of economic activities (NACE). *, ** and *** denote statistical significance at the 10, 5 and 1 percent level respectively.

Table 8: Recursive bivariate ordered probit estimates, Earnings Structure Survey

	ESS 2014	ESS 2010
Education equation		
Log (average monthly wage)	0.932*** (0.050)	1.287*** (0.016)
Log (market access)	-0.017 (0.039)	-
Collective bargaining agreement	0.235*** (0.044)	0.097*** (0.010)
Firm size	-0.016 (0.012)	-0.043*** (0.004)
Hours/week worked	-0.024*** (0.004)	-0.009*** (0.001)
Permanent employment	-0.178*** (0.050)	-0.321*** (0.019)
Full-time employment	-0.726*** (0.095)	-1.163*** (0.046)
Gender	-0.114*** (0.013)	-0.094*** (0.006)
Age	-0.082*** (0.004)	-0.107*** (0.002)
Age squared	0.001*** (0.000)	0.001*** (0.000)
Tenure	-0.134*** (0.010)	-0.172*** (0.003)
Tenure squared	0.005*** (0.000)	0.005*** (0.000)
NUTS-1 Regions		
West Marmara	0.188*** (0.009)	-
Aegean	0.121*** (0.023)	-
East Marmara	0.222*** (0.011)	-
West Anatolia	0.157*** (0.033)	-
Mediterranean	0.162*** (0.040)	-
Central Anatolia	0.129*** (0.028)	-
West Black Sea	0.222*** (0.011)	-
East Mediterranean	0.257*** (0.026)	-
Northeastern Anatolia	0.220*** (0.051)	-
Central Eastern Anatolia	0.334*** (0.048)	-
Southeastern Anatolia	-0.035 (0.065)	-
Wage equation		
Constant	3.570*** (0.136)	4.493*** (0.043)
Administrative work	0.226*** (0.012)	0.210*** (0.006)
Overtime	-0.000 (0.000)	0.000*** (0.000)
Log (market access)	0.089*** (0.008)	-
Collective bargaining agreement	0.059*** (0.020)	0.055*** (0.006)
Firm size	0.073*** (0.005)	0.090*** (0.001)
Hours/week worked	0.014*** (0.004)	-0.005*** (0.001)
Permanent employment	0.289*** (0.048)	0.332*** (0.012)
Full-time employment	0.980*** (0.119)	1.109*** (0.028)
Gender	0.086*** (0.015)	0.017*** (0.004)
Age	0.061*** (0.004)	0.048*** (0.001)
Age squared	-0.001*** (0.000)	-0.001*** (0.000)
Tenure	0.151*** (0.004)	0.133*** (0.001)
Tenure squared	-0.005*** (0.000)	-0.004*** (0.000)
NUTS-1 Regions		
West Marmara	-0.157*** (0.011)	-
Egean	-0.082*** (0.005)	-
East Marmara	-0.094*** (0.008)	-
West Anatolia	0.001 (0.005)	-
Mediterranean	-0.053*** (0.008)	-
Central Anatolia	-0.118*** (0.009)	-
West Black Sea	-0.187*** (0.013)	-
East Mediterranean	-0.133*** (0.007)	-
Northeastern Anatolia	-0.095*** (0.007)	-
Central Eastern Anatolia	-0.133*** (0.006)	-
Average marginal effects of wage		
Primary school or less	-0.044*** (0.013)	-0.065*** (0.006)
Primary & Secondary school	-0.011*** (0.003)	-0.012*** (0.001)
High school	0.008*** (0.002)	0.010*** (0.0004)
Vocational school	0.007*** (0.002)	0.010*** (0.0007)
University	0.040*** (0.012)	0.057*** (0.006)
Number of observations	164,023	198,161
Error correlation (ρ)	-0.553*** (0.040)	-0.752*** (0.013)

Notes: Standard errors in parentheses are clustered at the NUTS-1-level. All specifications include $9-1$ dummy variables on ISCO-08 occupations and and $17-1$ dummy variables on statistical classification of economic activities (NACE). Standard errors of the average marginal effects are computed via the Delta method. *, ** and *** denote statistical significance at the 10, 5 and 1 percent level respectively.

Table 9: Recursive bivariate ordered probit estimates by region, 2014 Earnings Structure Survey

	Istanbul	West Marmara	Aegean	East Marmara	West Anatolia	Mediterranean
Education equation						
Log (average monthly wage)	0.926*** (0.065)	1.027*** (0.156)	1.026*** (0.070)	0.711*** (0.103)	0.678*** (0.192)	0.753*** (0.092)
Collective bargaining agreement	0.248*** (0.028)	0.333*** (0.057)	0.100*** (0.032)	0.321*** (0.032)	0.151** (0.070)	0.274*** (0.039)
Firm size	-0.044*** (0.007)	-0.030* (0.017)	-0.011 (0.011)	0.043*** (0.010)	-0.001 (0.021)	0.070*** (0.016)
Hours/week worked	-0.017*** (0.003)	-0.029*** (0.007)	-0.037*** (0.005)	-0.053*** (0.006)	-0.024*** (0.005)	-0.020*** (0.005)
Permanent employment	-0.266*** (0.052)	-0.517*** (0.071)	-0.347*** (0.051)	-0.077 (0.072)	0.043 (0.052)	-0.060 (0.053)
Full-time employment	-0.784*** (0.112)	-0.652*** (0.219)	-0.567*** (0.183)	-0.012 (0.208)	-0.246 (0.325)	-0.435*** (0.143)
Gender	-0.156*** (0.013)	-0.100** (0.041)	-0.145*** (0.021)	-0.113*** (0.024)	-0.133*** (0.027)	-0.071*** (0.026)
Age	-0.085*** (0.005)	-0.113*** (0.009)	-0.091*** (0.007)	-0.075*** (0.008)	-0.055*** (0.014)	-0.096*** (0.009)
Age squared	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.000*** (0.000)	0.000 (0.000)	0.001*** (0.000)
Tenure	-0.138*** (0.010)	-0.138*** (0.025)	-0.156*** (0.013)	-0.083*** (0.017)	-0.112*** (0.032)	-0.104*** (0.017)
Tenure squared	0.005*** (0.000)	0.005*** (0.001)	0.006*** (0.001)	0.003*** (0.001)	0.004*** (0.001)	0.004*** (0.001)
Wage equation						
Constant	4.439*** (0.104)	3.561*** (0.178)	3.916*** (0.135)	4.410*** (0.175)	4.505*** (0.147)	3.396*** (0.159)
Administrative duty	0.217*** (0.014)	0.218*** (0.030)	0.269*** (0.018)	0.259*** (0.017)	0.181*** (0.022)	0.286*** (0.019)
Overtime	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.001** (0.000)	0.001 (0.000)	-0.000 (0.000)
Collective bargaining agreement	0.046*** (0.013)	0.040** (0.018)	0.092*** (0.016)	0.099*** (0.013)	0.195*** (0.021)	-0.099*** (0.021)
Firm size	0.075*** (0.003)	0.066*** (0.006)	0.075*** (0.004)	0.046*** (0.005)	0.081*** (0.004)	0.094*** (0.005)
Hours/week worked	0.006*** (0.002)	0.028*** (0.004)	0.014*** (0.004)	0.014*** (0.005)	0.001 (0.004)	0.029*** (0.004)
Permanent employment	0.485*** (0.027)	0.394*** (0.036)	0.394*** (0.032)	0.440*** (0.046)	0.138*** (0.030)	0.335*** (0.027)
Full-time employment	1.103*** (0.058)	0.562*** (0.135)	0.966*** (0.109)	0.832*** (0.133)	1.345*** (0.104)	0.349*** (0.095)
Gender	0.050*** (0.009)	0.149*** (0.016)	0.126*** (0.013)	0.103*** (0.014)	0.086*** (0.015)	0.111*** (0.015)
Age	0.069*** (0.003)	0.058*** (0.006)	0.057*** (0.004)	0.055*** (0.005)	0.068*** (0.005)	0.073*** (0.005)
Age squared	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Tenure	0.146*** (0.003)	0.145*** (0.005)	0.157*** (0.004)	0.148*** (0.004)	0.162*** (0.005)	0.162*** (0.005)
Tenure squared	-0.005*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)	-0.006*** (0.000)
Average marginal effects of wage						
Primary school or less	-0.043*** (0.014)	-0.042 (0.050)	-0.050** (0.021)	-0.043* (0.023)	-0.042 (0.036)	-0.045** (0.022)
Primary & Secondary school	-0.014*** (0.004)	-0.005 (0.006)	-0.009** (0.004)	-0.012** (0.006)	-0.015 (0.012)	-0.010** (0.004)
High school	0.003*** (0.0003)	0.009 (0.008)	0.010*** (0.003)	0.004*** (0.001)	0.006** (0.003)	0.007*** (0.003)
Vocational school	0.004*** (0.001)	0.012 (0.014)	0.008*** (0.003)	0.012** (0.006)	0.005 (0.004)	0.005** (0.002)
University	0.050*** (0.017)	0.026 (0.034)	0.040** (0.018)	0.039* (0.022)	0.045 (0.04)	0.043** (0.021)
Number of observations	39145	10686	18047	18565	16632	14164
Error correlation (ρ)	-0.536*** (0.056)	-0.612*** (0.118)	-0.593*** (0.055)	-0.369*** (0.079)	-0.347** (0.161)	-0.398*** (0.071)

Notes: Standard errors in parentheses are clustered at the individual level. All specifications include $9-1$ dummy variables on ISCO-08 occupations and and $17-1$ dummy variables on statistical classification of economic activities (NACE). Standard errors of the average marginal effects are computed via the Delta method. *, ** and *** denote statistical significance at the 10, 5 and 1 percent level respectively.

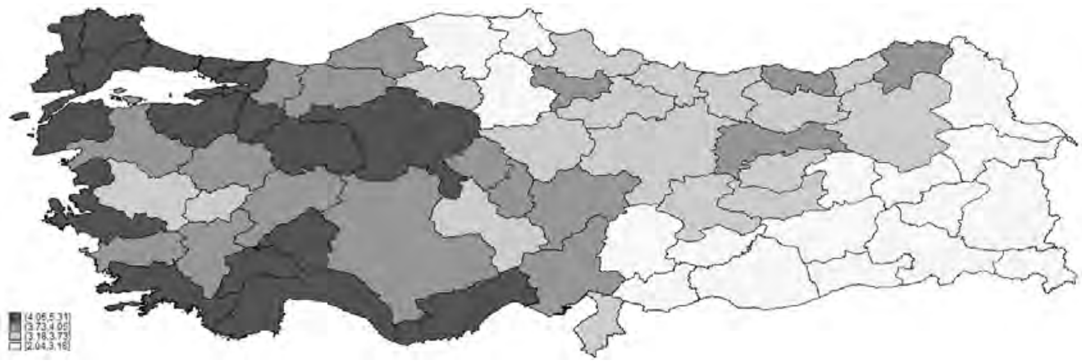
Table 9: Recursive bivariate ordered probit estimates by region, 2014 Earnings Structure Survey (continued)

	Central Anatolia	West Black Sea	East Mediterranean	Northeastern Anatolia	Central Eastern Anatolia	Southeastern Anatolia
Education equation						
Log (average monthly wage)	1.208*** (0.086)	0.413 (0.710)	0.804*** (0.143)	0.767*** (0.257)	1.101*** (0.108)	0.976*** (0.111)
Collective bargaining agreement	0.129*** (0.037)	0.396*** (0.082)	0.225*** (0.065)	0.475*** (0.084)	0.118 (0.078)	0.320*** (0.069)
Firm size	-0.022* (0.012)	0.055 (0.078)	-0.019 (0.017)	0.000 (0.037)	-0.028* (0.016)	-0.023* (0.013)
Hours/week worked	-0.046*** (0.007)	-0.005 (0.007)	-0.070*** (0.010)	-0.008 (0.008)	-0.035*** (0.007)	-0.031*** (0.006)
Permanent employment	-0.067 (0.060)	-0.185 (0.290)	-0.331*** (0.072)	-0.042 (0.100)	-0.188*** (0.053)	-0.001 (0.040)
Full-time employment	-0.405** (0.171)	-0.557 (0.709)	0.278 (0.307)	-0.796*** (0.264)	-0.858** (0.362)	-0.702*** (0.174)
Gender	-0.110*** (0.031)	0.076 (0.118)	-0.014 (0.033)	-0.056 (0.051)	-0.099** (0.039)	-0.148*** (0.034)
Age	-0.079*** (0.009)	-0.076*** (0.027)	-0.061*** (0.012)	-0.057*** (0.016)	-0.069*** (0.011)	-0.051*** (0.009)
Age squared	0.001*** (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000 (0.000)
Tenure	-0.181*** (0.018)	-0.048 (0.085)	-0.108*** (0.023)	-0.129*** (0.041)	-0.161*** (0.019)	-0.181*** (0.020)
Tenure squared	0.006*** (0.001)	0.001 (0.003)	0.004*** (0.001)	0.004*** (0.001)	0.005*** (0.001)	0.007*** (0.001)
Wage equation						
Constant	4.044*** (0.180)	4.324*** (0.202)	4.078*** (0.282)	3.932*** (0.229)	3.957*** (0.303)	4.002*** (0.147)
Administrative work	0.189*** (0.026)	0.151** (0.060)	0.301*** (0.027)	0.212*** (0.046)	0.244*** (0.025)	0.245*** (0.025)
Overtime	-0.000 (0.000)	0.003* (0.002)	0.001* (0.001)	-0.001 (0.001)	0.001** (0.001)	-0.001 (0.001)
Collective bargaining agreement	-0.009 (0.023)	0.041* (0.022)	0.043 (0.040)	-0.078* (0.042)	0.178*** (0.036)	0.126*** (0.029)
Firm size	0.054*** (0.005)	0.090*** (0.007)	0.075*** (0.007)	0.100*** (0.008)	0.059*** (0.007)	0.056*** (0.006)
Hours/week worked	0.035*** (0.004)	0.008** (0.004)	0.011 (0.008)	0.015*** (0.004)	0.011*** (0.004)	0.033*** (0.003)
Permanent employment	0.221*** (0.030)	0.399*** (0.034)	0.348*** (0.038)	0.246*** (0.041)	0.211*** (0.031)	-0.035 (0.026)
Full-time employment	0.217* (0.127)	0.979*** (0.111)	1.065*** (0.176)	0.708*** (0.147)	1.170*** (0.245)	0.669*** (0.130)
Gender	0.095*** (0.019)	0.129*** (0.021)	0.062*** (0.018)	0.055* (0.029)	0.086*** (0.023)	0.088*** (0.021)
Age	0.060*** (0.006)	0.042*** (0.006)	0.045*** (0.007)	0.052*** (0.008)	0.052*** (0.007)	0.035*** (0.006)
Age squared	-0.001*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)
Tenure	0.160*** (0.006)	0.116*** (0.007)	0.148*** (0.006)	0.156*** (0.008)	0.152*** (0.007)	0.171*** (0.006)
Tenure squared	-0.005*** (0.000)	-0.004*** (0.000)	-0.005*** (0.000)	-0.005*** (0.000)	-0.005*** (0.000)	-0.006*** (0.000)
Average marginal effects of wage						
Primary school or less	-0.034 (0.032)	-0.039 (0.164)	-0.028 (0.033)	-0.024 (0.052)	-0.023 (0.031)	-0.032 (0.033)
Primary & Secondary school	-0.009 (0.007)	-0.008 (0.034)	-0.009 (0.010)	-0.008 (0.017)	-0.007 (0.009)	-0.003 (0.004)
High school	0.011 (0.008)	0.012 (0.046)	0.002 (0.002)	0.004 (0.008)	0.009 (0.011)	0.012 (0.012)
Vocational school	0.006 (0.006)	0.009 (0.038)	0.005 (0.005)	0.004 (0.009)	0.004 (0.005)	0.003 (0.003)
University	0.025 (0.025)	0.026 (0.113)	0.029 (0.035)	0.023 (0.052)	0.017 (0.024)	0.020 (0.023)
Number of observations	10304	8603	7161	4928	6732	9056
Error correlation (ρ)	-0.759*** (0.066)	-0.175 (0.518)	-0.447*** (0.099)	-0.461** (0.195)	-0.681*** (0.077)	-0.592*** (0.081)

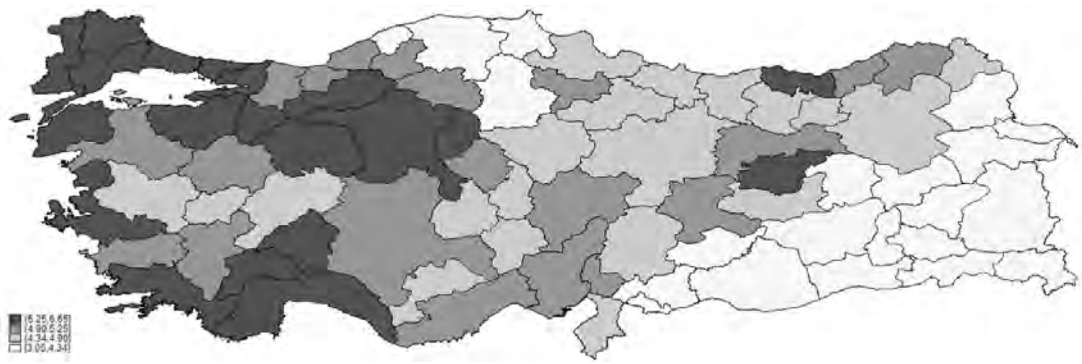
Notes: Standard errors in parentheses are clustered at the individual level. All specifications include $9-1$ dummy variables on ISCO-08 occupations and $17-1$ dummy variables on statistical classification of economic activities (NACE). Standard errors of the average marginal effects are computed via the Delta method. *, ** and *** denote statistical significance at the 10, 5 and 1 percent level respectively.

Figure 1: Average Years of Schooling

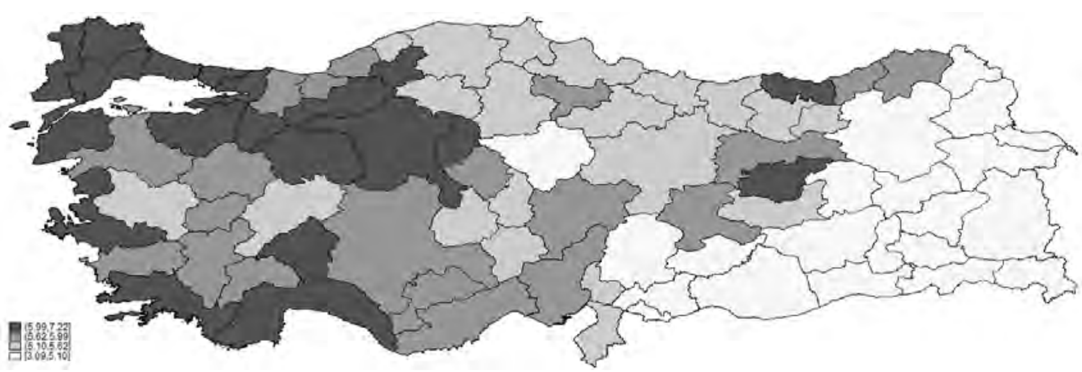
(a) 1985



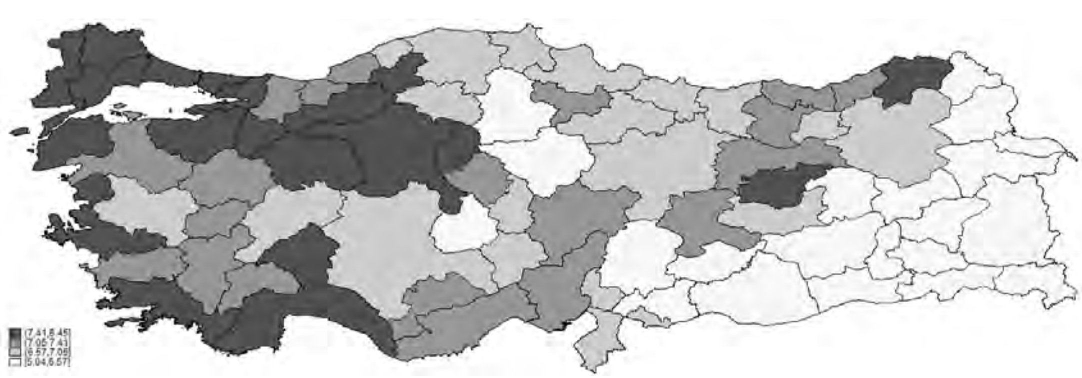
(b) 2000



(c) 2009

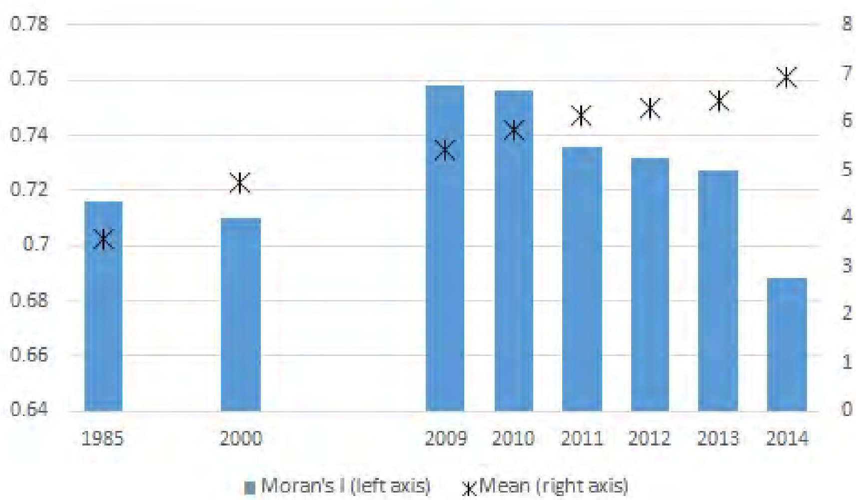


(d) 2014



Source: TurkStat, Authors' own calculations

Figure 2: Spatial Dependence and path of Average Years of Schooling



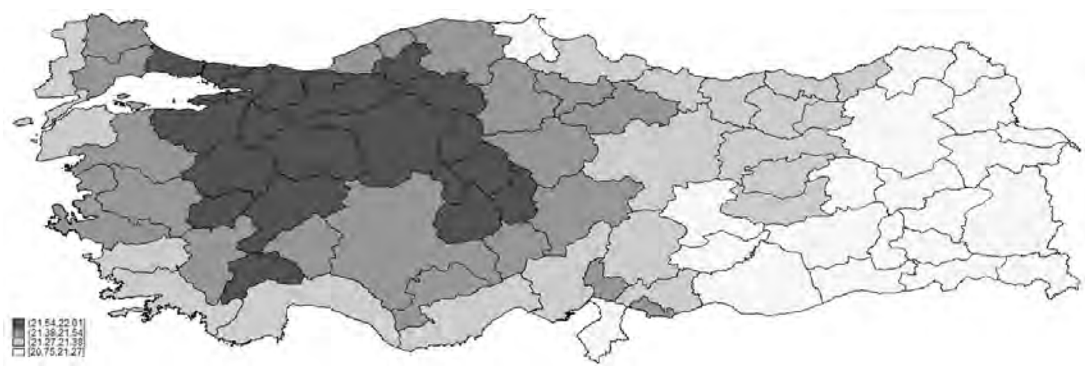
Source: TurkStat, Authors' own calculations.

Figure 3: Market Access (ln)

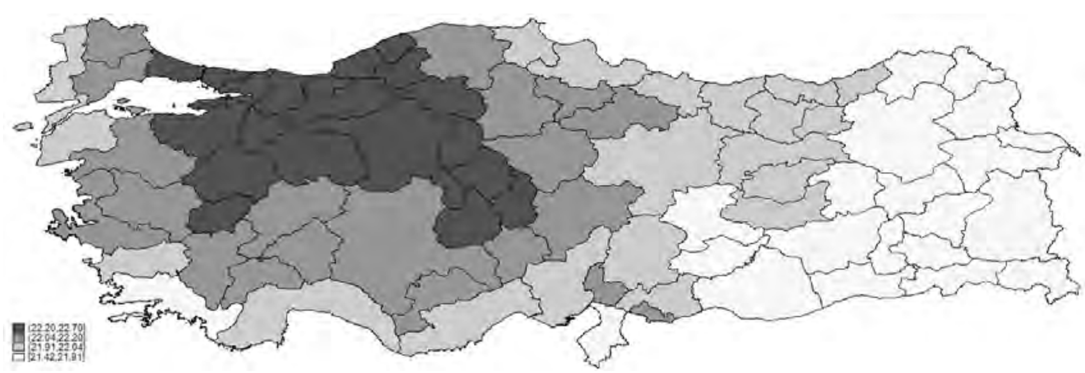
(a) 1985

(b) 2000

(c) 2009

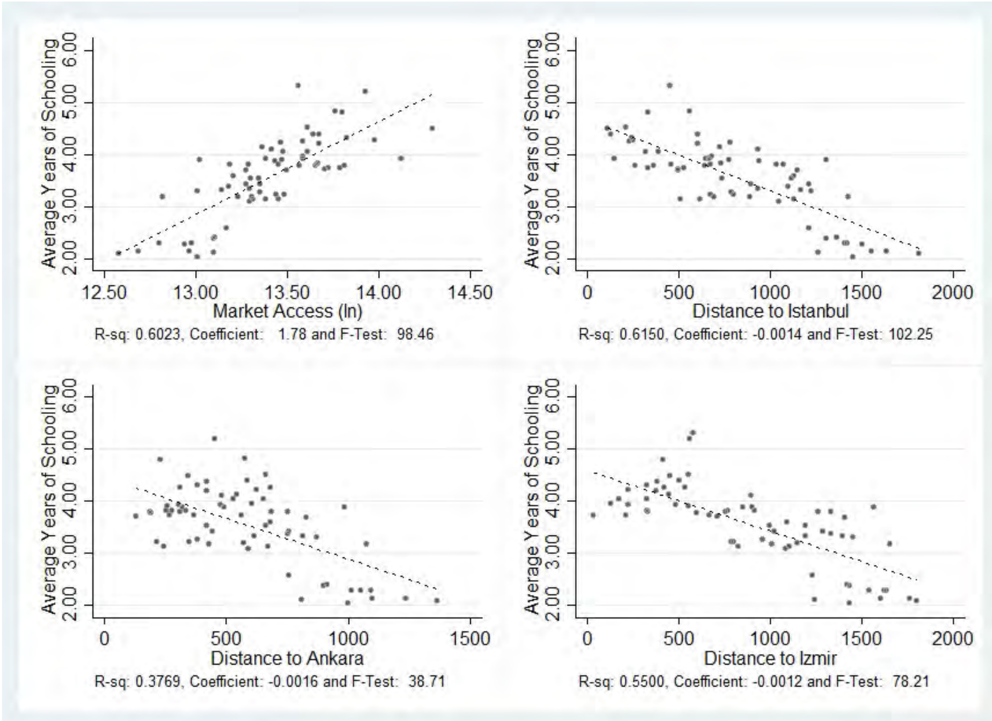


(d) 2014



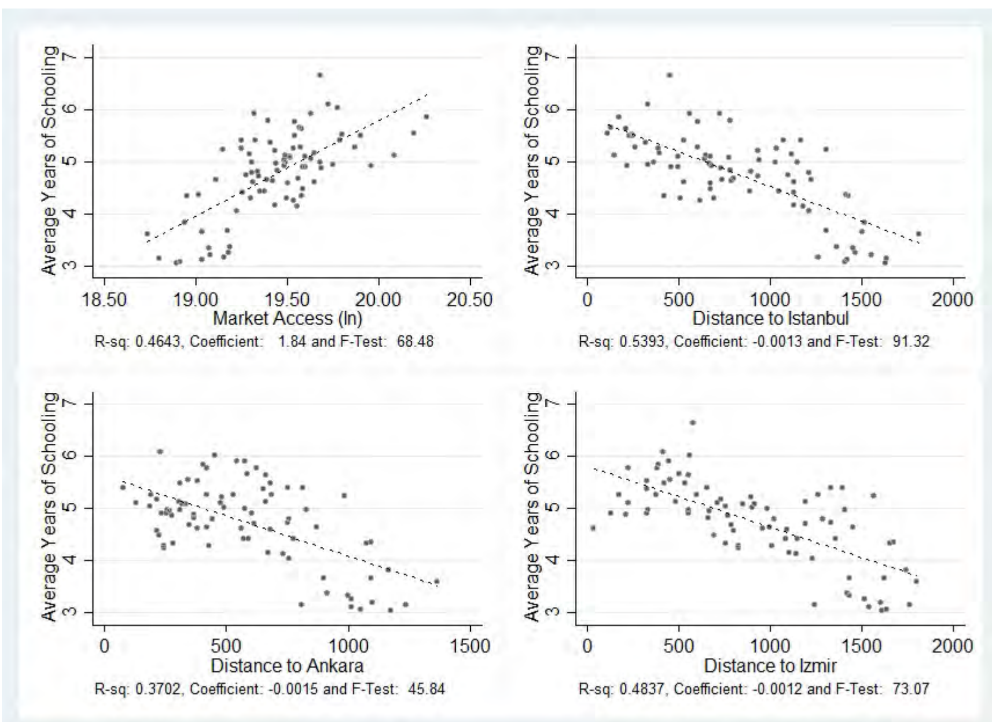
Source: TurkStat, Authors' own calculations

Figure 4: Accessibility and Human Capital Accumulation 1985



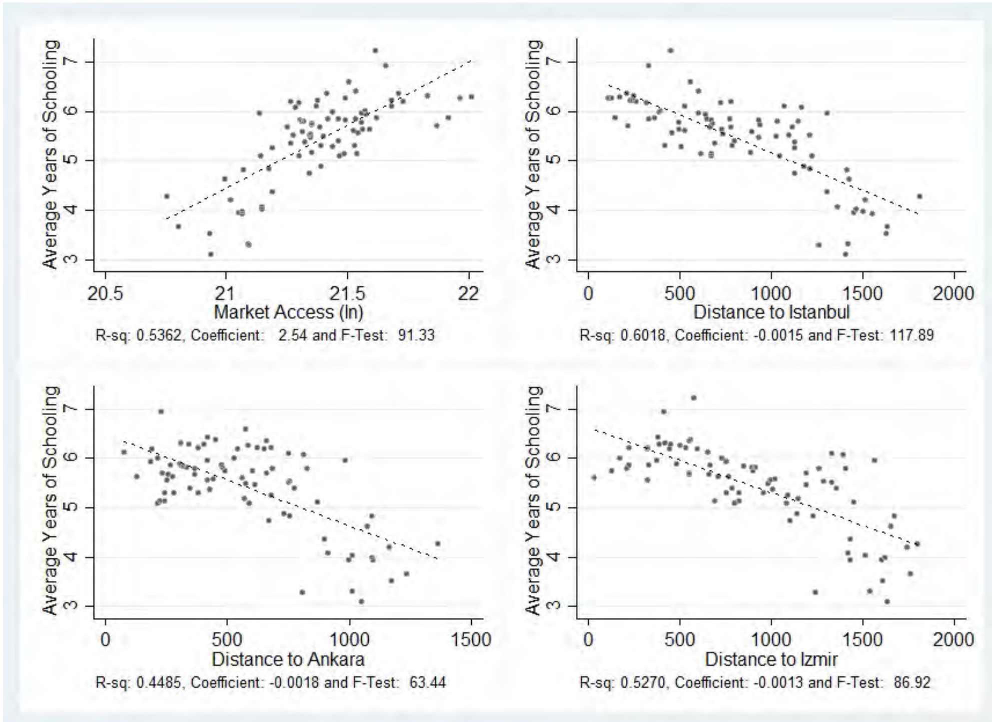
Source: Authors' own calculations.

Figure 5: Accessibility and Human Capital Accumulation 2000



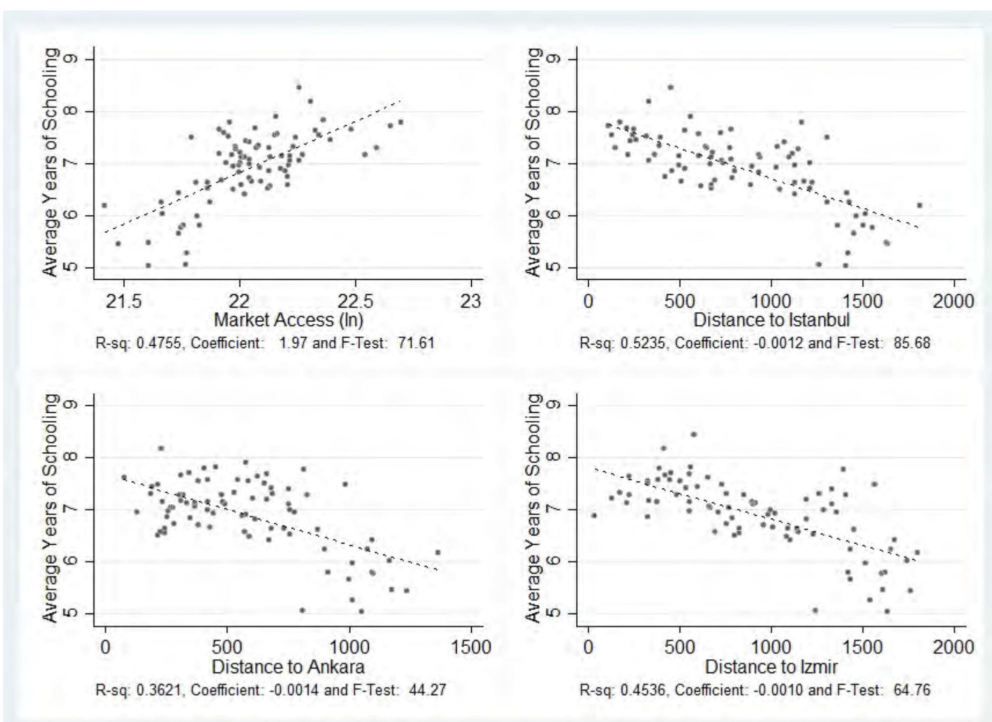
Source: Authors' own calculations.

Figure 6: Accessibility and Human Capital Accumulation 2009



Source: Authors' own calculations.

Figure 7: Accessibility and Human Capital Accumulation 2014



Source: Authors' own calculations.

Figure 8: Geographic distribution of education, wages and market access, NUTS-1, 2014 ESS

(a) Education (=1 if primary or less, ..., =5 if university)

(b) Average gross monthly wage (TL)

(c) Market Access

Source: TurkStat, Authors' own calculations

Figure 9: Geographic distribution of the average marginal effects of wage on education, NUTS-1, 2014

(a) Primary school or less

(b) Primary & Secondary school

(c) High school

(d) Vocational school

(e) University

Source: TurkStat, Authors' own calculations