

Intergenerational Education Mobility and the Level of Development¹

Abdurrahman B. Aydemir and Hakki Yazici²

This version: July 2018

Abstract

This main goal of this paper is to analyze the empirical relationship between economic development and intergenerational education mobility. By exploiting large regional variation in the level of economic development across Turkey, we find that educational outcomes of daughters who grow up in more developed regions depend less on their parents' educational outcomes. We do not find evidence for a similar relationship between development and mobility for males. Moreover, the development level of place of residence during childhood has a significantly stronger positive association with intergenerational education mobility compared to development level of place of residence during adolescence. We find that regions that feature higher school availability, greater level of social capital, more favorable cultural attitudes toward women, higher level of home resources, and lower degree of educational inequalities are associated with higher intergenerational mobility.

Keywords: intergenerational mobility, education, economic development

J.E.L. Classification: J6, I2, R0

¹ This research was supported by Tübitak research grant no. 112K502 and Sabancı University Internal Research Grant. We thank conference participants at the Intergenerational Transmission of Economic Status Workshop and Özyeğin University-World Bank Workshop on Poverty, Inequality and Intergenerational Mobility, and seminar participants at Istanbul Technical University and Koc University for their helpful comments. We especially thank Mikael Lindahl, Bhash Mazumder and Jan Stuhler for their suggestions. We thank Istanbul School of Central Banking for their hospitality. Usual disclaimers apply.

² Corresponding author: Yazici: Faculty of Arts and Social Sciences, Sabancı University, Orhanlı, Tuzla 34956, Istanbul, hakkiyazici@sabanciuniv.edu, phone: 216 4839331, fax: 216 4839250
Aydemir: Faculty of Arts and Social Sciences, Sabancı University, Orhanlı, Tuzla 34956, Istanbul, IZA, CREAM-UCL, aaydemir@sabanciuniv.edu, phone: 216 4839344, fax: 216 4839250

1. Introduction

The degree to which socioeconomic performance persists from one generation to another is often interpreted as measuring a society's success in providing equality of opportunity to children from different family backgrounds. A high degree of intergenerational persistence suggests that family background plays an important role in children's later success in life while a low degree of persistence implies the opposite. Another potentially important determinant of children's opportunity sets is the broader socioeconomic environment in which they grow up. Presumably, a more developed socioeconomic environment can provide children with more opportunities, which may attenuate the effect of parental background on children's outcomes. In this sense, one may expect that higher levels of economic development may be associated with a lower degree of intergenerational persistence. In this paper, we analyze how intergenerational mobility differs across people who grow up in regions with different levels of socioeconomic development. We do so in the context of Turkey using information from a nationally representative survey that we conducted over November 2014 – February 2015 period. In addition to information on educational attainment of individuals and their parents, the survey includes unique and detailed information on people's place of residence at various stages of development during childhood.

Before we conduct our main analysis of the empirical relationship between mobility and economic development, we first provide a thorough characterization of intergenerational education mobility in Turkey at the national level. We find that the correlation between father's (resp. mother's) and offspring's years of schooling is statistically significant and 0.564 (resp. 0.532).³ A comparison of these numbers with findings from Hertz et al. (2007), who provide comparable estimates of intergenerational correlations for a large set of countries, suggests that

³ Following the existing literature, we measure intergenerational persistence in education using two closely related measures: the regression coefficient and the correlation coefficient. We discuss the measures and when it is more appropriate to use one measure rather than the other one in detail in Section 2.

the level of education mobility is lower in Turkey than it is in the rest of the world, except for South American countries. We also find that intergenerational education mobility is significantly lower for females. Our survey also includes questions on the educational attainments of the respondents' children. We use this information to analyze the persistence of educational outcomes across three generations and find that grandparents' educational attainments have no effect on children's educational attainments when we control for the years of schooling of both parents.

Next, we proceed to the characterization of how intergenerational mobility varies with the level of economic development. Turkey provides a unique case study for understanding the role of development in shaping intergenerational mobility for two reasons. First, Turkey is a large country where large disparities in the level of development exist across regions. In particular, there is substantial heterogeneity in income, educational attainment, school access and education quality. This large regional variation is what enables us to study how mobility changes with development. Second, Turkey has a relatively strong and unified state structure which imposes a fairly uniform institutional environment across the whole country. Thus, differences in institutional environments and data comparability issues, which complicate the study of the relationship between mobility and development across countries, are much less pressing in the current context.⁴

In order to investigate how intergenerational mobility changes with the level of development, we first construct a variable that measures the socioeconomic development level of the places people grew up in our sample. The Turkish Ministry of Development periodically measures various aspects of socioeconomic development throughout Turkey, and as a part of this process, the ministry has developed an index that summarizes the level of socioeconomic

⁴ Moreover, Turkey is a fairly homogenous country in which inheritability of genetic factors that affect educational attainment presumably varies less across regions relative to cross-country variation in such traits. In this regard, regional variation in intergenerational mobility within Turkey is more likely to reflect the role of development in shaping mobility.

development at the province level. Larger values of this index indicate more developed regions. In our analysis, we use the most recent index, Socio-economic Development Index for 2011 (SEGE-2011) as the proxy for the development level of each province. Our data includes separate information about the place of residence when individuals were between the ages of 0 and 6 (preschool ages), 7 and 12 (primary school ages), and 13 and 18 (secondary school ages), as well as their current place of residence. Since we are interested in people's educational attainment, we expect that the characteristics of the place where they grew up matter more than the characteristics of their current place of residence. For this reason, in our baseline analysis we assign an individual the development index value of the province he or she resided while growing up, and in particular, during ages 7-12.

We modify the basic intergenerational regression equation by adding an interaction term between parents' years of schooling and the socioeconomic development variable we have constructed. We find that the coefficient of this interaction term is statistically significantly negative (regardless of whether we set fathers' or mothers' years of schooling as the parental education variable), indicating that intergenerational education mobility is higher in regions with higher levels of development. Interestingly, when we divide our sample into male and female subsamples and estimate the modified intergenerational regression equation separately for each subsample, we find that the coefficient of the interaction term is statistically significant only for females. That is, we do not find evidence for a positive relationship between mobility and development for males.

We repeat our estimations by assigning development index values according to the place of residence during ages 0-6, 13-18, and current place of residence as well. We find that the significant positive association between the development level of place of residence and intergenerational mobility continues to hold robustly independent of how we define place of residence. Comparison of coefficient estimates of interaction terms across regressions reveals,

however, that this association is significantly stronger when we define place of residence as the place of residence during ages 0-6 and 7-12 rather than the place of residence during ages 13-18 or current place of residence. We further investigate the heterogeneity in the association between place of residence at various stages of childhood and intergenerational mobility by considering: (i) a regression specification where we include all the interaction terms between parental years of schooling and place of residence during ages 0-6, 7-12, 13-18 and current place of residence at the same time, (ii) migrant subsamples. Interestingly, we find that while the place of residence during early childhood (ages 0-12) has a significant association with intergenerational mobility, the place of residence during adolescence or current place of residence does not. The finding that the development level of the place of residence during earlier years of childhood is more important for intergenerational mobility in education resonates with the recent empirical findings on the importance of early childhood investments on adult outcomes.⁵ To our knowledge, our paper is one of the first in the literature that investigates how place of residence at different stages of childhood development affects intergenerational mobility.

We also conduct a nonparametric analysis of the relationship between development and intergenerational mobility. To do so, we partition the provinces individuals resided in Turkey while growing up into 15 development regions according to their levels of socioeconomic development. We then estimate a separate intergenerational regression coefficient for each region. We find that intergenerational education mobility is higher in regions with higher level of economic development. For instance, while one year increase in mothers' years of schooling is associated with a more than one year increase in daughters' years of schooling in the least developed region of Turkey, the same number is 0.6 in the most developed region. Moreover, when we estimate the relationship between education mobility and development for males and

⁵ See Elango et al. (2015) for an excellent review of the empirical literature on the importance of early childhood education for adult outcomes.

females separately, we find that this relationship is much stronger for the female subsample. In fact, when we conduct robustness exercises in which we partition Turkey into 5, 10, 20, and 25 regions (instead of 15), the relationship between mobility and development remains consistently significant across these exercises only for females. These results confirm our findings from the parametric case that education mobility varies significantly with economic development, and this is especially true for females.

It is important to stress that the evidence we provide regarding the positive relationship between economic development and mobility is descriptive and does not identify any causal mechanisms. Yet, we believe that the correlations we calculate here are valuable because they can guide future research that aims to identify such mechanisms.

Finally, we take a first step toward understanding why development is positively associated with intergenerational mobility. To do so, we investigate what characteristics of the environments in which children grow up are more important for intergenerational mobility. Specifically, we focus on five characteristics: availability and quality of educational facilities, social capital, cultural attitudes toward women, childhood home resources, and the extent of educational inequalities. We find that higher school availability, greater level of social capital, more favorable cultural attitudes toward women, higher level of home resources, and lower degree of educational inequalities are associated with lower intergenerational education persistence for daughters. We do not find similar evidence for an association between these characteristics and mobility for sons. This is in line with the lack of evidence for a relationship between development and mobility for sons that we present earlier. Moreover, we find that these environmental characteristics matter much more for intergenerational mobility across mother-daughter pairs relative to father-daughter pairs.

Literature review This paper contributes to an extensive literature on intergenerational mobility.⁶ Most closely related are a small collection of recent papers which analyze regional differences in mobility. Chetty et al. (2014) find that intergenerational mobility varies significantly over geographic areas in the United States. They show that high mobility areas have less residential segregation, less income inequality, better primary schools, greater social capital, and greater family stability. More recently, Guell et al. (2018) analyze intergenerational income mobility in 103 Italian provinces, and document that higher income mobility is associated with higher value added per capita, higher schooling, greater social capital, and lower income inequality. We contribute to this literature in a number of dimensions. First, we provide a partial confirmation of the earlier findings that better economic outcomes and higher levels of economic development are associated with higher intergenerational mobility. This confirmation is partial because even though we find evidence for such a relationship between mobility and development for females, we do not find similar evidence regarding mobility of males. Second, our data set has information about children's place of residence during various stages of childhood which allows us to investigate how the development level of the place of residence at different stages of childhood affects intergenerational mobility.⁷

Two recent papers by Chetty and Hendren (2018a and 2018b) are closer to the current paper in the sense that, using information on place of residence starting from age eight onward, they investigate the extent to which children's opportunities for upward economic mobility are affected by the neighborhoods they grow up in the United States. Our paper differs from theirs in at least two ways. First, while we focus on how intergenerational correlations of economic outcomes vary with development level of locations children grew up, they are primarily concerned with how upward mobility of children from economically disadvantaged families is

⁶ See Solon (1999) and Black and Devereux (2011) for excellent reviews of this literature.

⁷ Chetty et al. (2014) assign parent-child pairs into regions based on where children resided at the age of 15 in their core sample while Guell et al. (2018) assign them according to their current place of residence. Similar to Chetty et al (2014), Heidrich (2017) also analyzes regional variation in mobility in the context of Sweden. She defines childhood location as the region where a child lived for at least 6 years between the ages of 6 and 15. However, she does not compare relative importance of various places of residence during childhood for mobility.

affected by the neighborhoods they live in. Second, unlike theirs, our data set contains information on children's place of residence during early years of childhood, namely ages 0-6. The presence of this data enables us to establish that the place of residence during earlier years of childhood (ages 0-12) is more important for mobility than the place of residence during adolescence, a finding that is at odds with Chetty and Hendren (2018a), who find that neighborhood exposure effects on upward mobility do not depend on age.

By providing mobility estimates for Turkey at the national level, we also contribute to the growing literature on estimates of intergenerational mobility in developing countries.⁸ In addition, our paper also contributes to a burgeoning literature on multigenerational mobility, which analyzes the persistence of economic outcomes across three or more generations (see Solon 2015 for an excellent review). Behrman and Taubman (1985) estimate regressions of children's years of schooling on parents' and grandparents' and find grandparents' years of schooling have regression coefficients that are very small and statistically insignificant.⁹ Recently, Braun and Stuhler (2017) find a positive coefficient of grandparent educational attainment but also present evidence against its causal interpretation. Using Swedish data, Lindahl et al. (2015) estimate a statistically significant regression coefficient for grandfather's earnings. There is little work on multigenerational mobility in developing country contexts. One such example is Zeng and Xie (2014) who study the effects of grandparent's on children's schooling within the context of rural China and find that living with well-educated grandparents significantly reduces children's likelihood of school dropout.

⁸ Most studies of intergenerational education mobility in Turkey use samples that consist of children who still reside with their parents, and hence, might have not completed their educational attainments. See, for instance, Tansel (2002), Kırdar (2009), Uysal-Kolaşın and Yontar (2012). Özdural (1993) and Korkmaz (2005) are exceptions that target specific cities, Ankara and Malatya, respectively. In a concurrent paper, Tansel (2015) uses the 2007 Adult Education Survey and reports estimates of intergenerational education mobility by age cohorts. She finds some evidence of increased mobility for the younger generations of children in Turkey. For recent contributions regarding other developing countries, see Dunn (2007) on Brazil; Knight et al. (2011) on China; Azam and Bhatt (2015), Emran and Shilpi (2015) and Hnatskova et al. (2013) on India; Emran and Shilpi (2011) on Nepal and Vietnam; Piraino (2015) on South Africa.

⁹ Ridge (1974), Warren and Hauser (1997) and Kerr and Lucas (2013) confirm the findings of Behrman and Taubman (1985) using data from U.K., U.S. and Finland, respectively. Similar results are reported by Jaeger (2012) for the US and Chiang and Park (2015) for Taiwan.

2. Methodology

The main aim of the studies in intergenerational education mobility is to understand the degree to which offspring's educational outcomes depend on their parents' educational outcomes. Due to data limitations, most studies in this literature use information on the highest degree completed to measure people's educational outcomes. Another, finer, measure of educational attainment is completed years of schooling. The latter measure of educational attainment has two possible merits relative to the former. First, obviously, completed years of schooling is a more precise measure of educational attainment, and this precision might be especially important if a substantial number of people exhibit drop out behavior and drop out does not occur right after completing the highest degree attained by the individual. Second, completed years of schooling is more of a continuous variable compared to highest degree completed, and as such allows for regression analysis. Motivated by these merits of completed years of schooling as a measure of educational attainment, our survey is carefully designed to extract information on additional years of schooling completed from dropouts. As a result, we have data on completed years of schooling and that is the main educational outcome variable we use throughout the paper.

There are two well-known average measures of intergenerational education mobility: the regression coefficient and the correlation coefficient. These measures are simple statistics that inform us about the average predictive power of parental education on children's education. The regression coefficient, β , is obtained by regressing children's years of schooling S_1 on parent's years of schooling S_0 :

$$S_{1i} = \alpha + \beta S_{0i} + \varepsilon_i. \quad (1)$$

This measure has the merit of being easily interpretable: it represents the average increase in children's years of schooling that occurs when parental years of schooling increases by one

year. As such, the regression coefficient tells us how strong the association is between parent's and offspring's education.

The correlation coefficient, on the other hand, is a mobility measure that is informative about the association between a parent's rank in the parental education distribution and a child's rank in their own education distribution. In order to compute the correlation coefficient, which we denote by β^a throughout the paper, we first adjust the parents' and children's distributions of years of schooling to equate the standard deviations of these distributions. We do so by dividing the years of schooling variable for children by the standard deviation of children's years of schooling distribution. This gives us an adjusted years of schooling variable for children, denoted by S_1^a , which is distributed with a standard deviation of one. We repeat the same procedure and create adjusted years of schooling variable for parents, denoted by S_0^a , whose distribution also has a standard deviation of one. Then, we conduct the following regression analysis:

$$S_{1i}^a = \alpha^a + \beta^a S_{0i}^a + \varepsilon_i. \quad (2)$$

The regression coefficient β^a represents the standard deviation increase in children's years of schooling rank that occurs when parents' years of schooling rank increases by one standard deviation. The correlation coefficient is related to the regression coefficient as follows:

$$\beta^a = \beta \frac{\sigma_0}{\sigma_1}, \quad (3)$$

where σ_0 and σ_1 denote the standard deviations of parents' and children's unadjusted years of schooling distributions, respectively.

Notice that the correlation coefficient can be especially useful as a measure of intergenerational mobility if we want to make comparisons across countries since both parents'

and children's distributions can vary significantly across different countries. Similarly, the correlation coefficient might be a better measure if we are analyzing the evolution of mobility over time. A change in the regression coefficient might result from a change in relative distributions of parents' and offspring's generations over time whereas a change in the correlation coefficient, by definition, has to come from a change in how strong children's education is related to that of parents'.

3. Data

There is no data set in the Turkish context that provides the key information required for our analysis such as place of residence during various stages of childhood. Thus, we conducted a household survey across Turkey to construct the data set needed for our investigation. The survey questionnaire asks a total of about two hundred questions regarding the demographics, educational status, labor market status, and cultural attitudes of households.

Data collection The survey targets adult population, defined as individuals who are 18 years old or older. The number and the geographical distribution of the sampling units are determined so as to achieve representation of Turkey both geographically as well as across the rural/urban divide within each region.¹⁰ The resulting sampling frame used in our analysis includes addresses from 74 of 81 provinces. Within each household, the interviewee is also determined randomly.¹¹ If the person chosen for the interview is not available on the first visit, the same address is visited two more times. If the interview cannot be conducted after three

¹⁰ The random sample of household addresses is provided by Turkish Statistical Institute (TUIK) based on the national registers (ADNKS 2013). The sampling procedure used by TUIK partitions Turkey into 26 regions, called NUTS2 regions. There are 81 provinces in Turkey and a NUTS2 region typically involves a number of provinces (up to 5). For sampling purposes TUIK divides each NUTS2 region into blocks of 400 households and these blocks are further classified as either an urban or a rural area of residence. The random sample of households is drawn in two steps. In the first step, a random sample of blocks is drawn from each NUTS2 region. In the second step, a random sample of households is drawn from each block. The two-step procedure ensures representativeness of population across NUTS2 regions as well as across the urban/rural divide within each NUTS2 region.

¹¹ For a given household the interviewer first produces a list of birth date information of all household members who are 18 and above. Then, the interviewer chooses as interviewee that household member whose day of birth (e.g. 20th of month X) is closest to the day of the month the interviewer first visits this address (e.g. 18th of month Y).

visits, the address is dropped from the sampling frame. The survey was conducted over November 2014 - February 2015 period and resulted in 4703 completed interviews.¹²

Sample weights Household Labour Force Survey (HLFS), conducted by Turkish Statistical Institute (TUIK), is a representative sample of households in Turkey with around half a million observations in each cross-section. We construct sample weights to ensure that the unconditional distributions of age, gender and education that result from our survey match the corresponding distributions computed using HLFS 2013 data. We chose the 2013 survey because our sampling frame is based on the national registers data for that year (ADNKS 2013). Specifically, the sample weights ensure that the distributions of individuals in our sample over 13 age groups, five education categories, and two gender categories match those in HLFS. The use of sample weights is important because significant deviations in the distributions of these characteristics from a nationally representative sample such as HLFS may lead to biases in the nationwide estimates of intergenerational mobility if there are significant differences in mobility rates across population subgroups. Below we also provide evidence, using HLFS as well as another nationally representative data set, that our weighted data matches other important population characteristics (in particular geographic distribution of individuals and parental educational distribution) and yields mobility estimates at the national level that are very similar to estimates from alternative data.

¹² The field work for our survey was conducted by a private polling company. The response rate was 53% which is common for surveys conducted through polling companies within the Turkish context (as there is no obligation for respondents to complete the survey). Among the addresses that refused to respond, 96.2% rejected outright to talk to a polling company while for the remaining 3.8% randomly chosen individual within the household rejected to respond. Thus, once the interview process reached the stage of randomly choosing the person to be interviewed the successful completion rate of the interview was 96.6%. Being aware of low response rates for surveys conducted by polling companies, we have asked for a sample size of random addresses from Turkish Statistical Institute that is much larger than the target number of completed surveys. If the interview at a randomly chosen address could not be conducted after three attempts we did not allow for what is called "snow balling", where the interviewer chooses an address nearby, as this technique may jeopardize randomness. The interviewers had to strictly stick to the random sample of household addresses provided to them and continue polling until all such addresses were visited.

Our primary analysis sample includes all individuals in the base data set who are 25 years or older. We restrict our sample in this way to focus on individuals who finished their education. This sample, which we call the main sample, has 3500 observations.

Our main educational attainment variable is years of schooling. For all agents, this variable is defined as completed years of schooling. For example, if an individual reports that she dropped out after two years in high school, then her years of schooling is computed as $5+3+2=10$ years. (In Turkey, primary school and secondary school take five and three years to complete, respectively.) In addition to their own education level, survey respondents report their parents' and (all of) their children's educational attainments as well. Based on the retrospective questions about parental education, we compute parents' years of schooling.

There is another data set, namely Adult Education Survey (Yetiskin Egitimi Anketi), which provides information on educational attainment of children and parents in Turkey.¹³ Compared to AES, our survey data has three key advantages. First, our survey data contains much more detail on educational attainment of individuals. Second, our data includes province level information on the region of residence during various periods of childhood development (including current residence) while AES surveys only include information on current residence at a level of 21 broadly defined geographic regions. Finally, our data includes information on educational attainment across multiple generations while such data is not available in AES. Therefore, unlike our data, AES is adequate neither for analyzing the relationship between development and mobility nor for conducting multigenerational mobility analysis. Nevertheless, information on parental education from AES is useful to us as it provides a benchmark to compare the educational distribution of parents from our survey. The comparison yields a close match between the two surveys regarding parental education distributions. The only significant difference is observed at the bottom of the distribution of mothers' education

¹³ AES surveys were conducted by TUIK in 2007 and 2012 to collect information on participation in adult learning activities.

where in our sample the fraction of those with no education is relatively lower and the fraction with primary school education is higher; similar fractions are found for the other education categories.¹⁴

Descriptive statistics Table 1 reports summary statistics for the main sample. Panel A of the table reports basic statistics for various demographic and educational characteristics. The mean age among survey respondents (whom we refer to as children in the two-generation analysis) is 45.4 years. The mean years of schooling among children and parents in this panel indicate a substantial increase in educational attainment across generations. For instance, while mean years of schooling is only about 4 years among fathers and it is 7 years among children.

Panel B of Table 1 provides further information on educational attainment by generation. The substantial increase in mean years of schooling across generations is also reflected in terms of educational attainment. For example, while the fraction of those with less than primary education is 38.5 % among fathers, it decreases to 8.7 % among male children. There is a parallel increase in the fraction of those with college or university degree. This fraction increases from 2.3 % among fathers to 16.5 % among male children. Similar improvements in educational attainment are observed among females although females in general have lower educational attainment than their male counterparts in the same generation. The comparison of male and female education by gender also indicates that the improvement in

¹⁴ Parents' education categories in the public use micro data files of AES surveys are provided in three very broad categories (less than high school, high school, and more than high school). However, Tansel (2015) reports detailed education distributions by age cohorts for parents using confidential files of AES 2007. We compute the education distributions for fathers and mothers for the same birth cohorts using our data and compare them to those in AES 2007. Five educational attainment levels are available from the confidential AES survey: no degree holders, primary school (5 years of education), secondary school (8 years of education), high school, and above high school. The fraction of fathers in the AES survey (our survey) for these five education categories are 40.9 % (38.5 %), 46.2 % (47.9 %), 5.1 % (5.5 %), 4.8 % (5.8 %), and 2.9 % (2.3 %) respectively. The fraction of mothers in the AES survey (our survey) for these five education categories are 62.6 % (56 %), 31.9 % (37.2 %), 2.4 % (2.5 %), 2.1 % (3.3 %), and 0.9 % (1.1 %) respectively. There are also detailed tabulations provided in a publication by TUIK for AES 2012 survey (Yetişkin Eğitimi Araştırması 2012, TUIK, accessed on November 2015, www.tuik.gov.tr/IcerikGetir.do?istab_id=218). There are, however, some clear errors in the reported numbers. For example, among fathers, the fraction that completed primary education (5 years of education) is reported as 13.7 % while the fraction that completed secondary education (8 years of education) is 52.7 %. These figures are at odds with the earlier AES 2007 survey, with the Household and Labour Force Surveys and our survey. Therefore, we don't provide any comparisons to figures reported in that publication.

educational attainment is accompanied by an increase in education gender gap. For example, while the difference between females and males in the fraction of those with less than primary education (i.e. females-males) is 17.5 percentage points (pp.) among parents, it becomes 20.5 pp. among the children. This is also reflected at higher education levels. While the fraction of mothers with high school education or more is 3.8 pp. lower than fathers, this difference becomes 12.6 among children. To sum up, there has been a significant rise in educational attainment across two generations in Turkey; however, the underlying factors that led to these improvements have not equally benefitted males and females.

Table 1 [about here]

Given the importance of geography for our main analysis, we compare the geographical distribution of our sample across Turkey to the geographical distribution in HLFS. TUIK partitions Turkey into 12 broad regions called NUTS1 regions. Table 2 compares the fractions of people living in each region in our survey and HLFS 2013. Despite the much smaller size of our sample (about 1% of the corresponding sample size in HLFS 2013 for individuals 18 and over) and the fact that our weights do not correct for the spatial distribution, we find that the two surveys feature very similar geographic distributions across Turkey.

Table 2 [about here]

4. Intergenerational Education Mobility at the National Level

4.1 Baseline Estimates

In this section, we present intergenerational mobility estimates at the national level. Figure 1 Panel A plots mean years of schooling of children as a function of fathers' years of schooling. The long-dashed line refers to the 45-degree line while the short-dashed vertical and horizontal lines mark the mean schooling of children and fathers. This nonparametric representation of the relationship between children's and fathers' years of schooling shows that the conditional

expectation of children's years of schooling given fathers' years of schooling is approximately linear in fathers' years of schooling.¹⁵

Figure 1 [about here]

Partly motivated by the approximate linearity of the relationship in Figure 1, we estimate a regression of children's years of schooling on parents'. Column (1) of Table 3 presents the regression coefficient associated with fathers' years of schooling in the regression where we regress children's years of schooling on fathers' years of schooling using our main sample. The regression coefficient is 0.734 and is highly significant. This implies that an increase of one year in father's years of schooling is associated with about 0.7 years increase in child's years of schooling. Column (2) reports the regression coefficient for the specification where we regress children's years of schooling on mothers'. The regression coefficient of mothers' years of schooling, 0.767, is similar in magnitude to the estimated coefficient for fathers' years of schooling, and is also highly significant. Finally, Column (3), which reports results for the regression analysis where both parents' education levels are present as independent variables, shows that the education levels of both parents are positively related to children's education. Just to provide a comparison, we also conduct the same mobility regressions using information from AES and find that our national level estimates are broadly in line with AES based estimates.¹⁶

Table 3 [about here]

¹⁵ Formally, we also tried to fit a quadratic specification between mean child years of schooling and father's years of schooling; the coefficient of the quadratic term turned out to be insignificant.

¹⁶ For estimation purposes, we pool 2007 and 2012 AES data and restrict the sample to individuals who are 25 years old or older as in our analysis. This results in around 56,900 observations. An important restriction of the AES data is that parent's education categories are very broadly defined in the public use files. In order to partly overcome this limitation, we use information on detailed education distributions from Tansel (2015), who uses confidential files of AES 2007, and assign years of schooling to parents. The regression coefficient of father's years of schooling is 0.724 (0.007) when we regress children's years of schooling on father's using AES data. This coefficient estimate is very similar to the estimate from our survey data (0.734), but more precisely estimated given the fact that AES sample sizes are much higher. The regression coefficient of mother's years of schooling is 0.837 (0.011). This estimate is somewhat higher than the estimate obtained from our data (0.767).

It is interesting to compare Turkey to other countries in terms of education mobility. However, as explained in Section 2, using regression coefficients for such a comparison might be misleading. For this reason, we also compute correlation coefficients. Figure 1 Panel B presents a nonparametric representation of the relationship between mean adjusted years of schooling of children and fathers' adjusted years of schooling. Panel B shows that the conditional expectation of children years of schooling given fathers' years of schooling is approximately linear on fathers' years of schooling, a finding that supports the linear regression specification that delivers the correlation coefficient. Columns (4) and (5) in Table 3 report that the correlation coefficients between children's and fathers' and children's and mothers' years of schooling are 0.564 and 0.532, respectively.

According to Hertz et al. (2007), South American countries feature the lowest education mobility in the world with a regional average intergenerational correlation value of 0.6. The intergenerational correlations are typically about 0.4 in Western Europe, with the exception of Italy where the correlation is 0.54. The lowest estimates in Europe belong to the Nordic countries. The U.S. estimate is 0.46. These values imply that the mobility in Turkey is somewhat higher than the South American average but it is significantly below the developed countries of Western Europe and the United States.¹⁷

4.2 Three Generation Education Mobility

In our survey, we gather information on the educational outcomes of people's children as well as their parents. This information allows us to study intergenerational mobility across three generations. We first drop the observations in which the respondent has no children. We set the respondent's parents as the first generation, the respondent as the second generation, and the

¹⁷ The country-level correlation values reported here are taken from Table 2 on page 15 in Hertz et al. (2007). This paper provides correlations of average parental years of schooling with children's years of schooling. As such, these values are not readily comparable to our estimates of 0.564 and 0.532. Just for comparability, we also compute the correlation between average parental years of schooling and children's years of schooling and find it to be around 0.6, which is closer to estimates for Latin American countries.

respondent's children as the third generation. If the respondent has multiple children, we treat each child with the corresponding parents and grandparents as a separate observation and cluster standard errors at the family level.¹⁸ We drop the observations in which the children are less than 25 years old as they might not have completed their education by that age. Moreover, we focus on grandfathers as the first generation and exclude grandmothers completely from our regressions. We do so mainly because there is too little variation in the years of schooling of grandmothers.¹⁹

We summarize our results in Table 4. The first and second columns of the table report results of regressions in which the only independent variables are father's years of schooling and paternal grandfather's years of schooling, respectively.²⁰ The second column shows that grandfather's years of schooling has a positive and significant relationship with children's years of schooling. The third column displays results of the analysis in which we regress children's years of schooling on father's and grandfather's years of schooling at the same time. The regression coefficient associated with grandfather's years of schooling loses its significance when we control for father's years of schooling. This means that grandfathers' educational outcomes are associated with children's educational outcomes only through fathers'. Columns (5) – (8) of Table 4 report the results of the three-generation regressions in which the parent is a mother. Column (7) shows that the regression coefficient of the maternal grandfather's years of schooling remains significant even after we control for mothers' years of schooling. Following Braun and Stuhler (2017), we test for a causal effect of maternal

¹⁸ Note that for a representative sample of third generation it is necessary in our case to use information on each child of survey respondent. If instead, for example, we randomly choose one of the children of each survey respondent (when the respondent has multiple children) this would lead to a discrepancy in family characteristics of third generation between our sample and that in the population –such as in terms of the distribution of the number of siblings.

¹⁹ The average age of respondents with a randomly chosen child of at least 25 years old is 61 in our sample which implies an average age of around 85 for grandmothers for this restricted sample. The educational attainment of women in that age cohort is very low and shows little variation in Turkey. For example, 95 % of all the grandmothers in the sample of the three-generation analysis have either 0 or 5 years of schooling.

²⁰ The father-child regression coefficient reported in the first column of Table 4 is different from the corresponding coefficient in the first column of Table 3. This is because the sample we use for three generation mobility analysis is quite different from our main sample. A similar difference across estimates is also present in mother-child regressions for the same reason.

grandparent's educational attainment on children's educational attainment by adding father's years of schooling to the maternal grandfather-mother-child regression.²¹ The results are reported in Column (8). The inclusion of father's education turns the maternal grandfather's education insignificant, implying that the effect of grandfather's education on children is more likely to be spurious than causal.

Table 4 [about here]

4.3 Intergenerational Education Mobility by Gender

In this section, we assess whether intergenerational persistence in education differs by gender. To do so, we conduct the following regression analysis on our main sample:

$$S_{1i} = \alpha + \beta_0 S_{0i} + \beta_1 (S_{0i} * G_i) + \beta_2 E_i + \varepsilon_i. \quad (4)$$

Here, S_1 and S_0 represent individual's and parent's years of schooling as before. G is a dummy variable that takes value 1 if the individual is female and 0 otherwise, and $E=1-G$. In this regression, β_0 measures the increase in children's years of schooling when parents' years of schooling increases by one year while β_1 measures the additional change in the years of schooling of the offspring associated with one year increase in parental education if the offspring is female.

Column (1) of Table 5 reports the results of the regression analysis given by Equation (4) for the case where we take father's years of schooling as the parental education input. We find that the coefficient of the interaction term is positive and statistically significant, implying that the regression coefficient between fathers' and daughters' years of schooling is

²¹ Warren and Hauser (1997) is the seminal study that includes both parents' educational outcomes in three generation regressions to test for causal effects of grandparent's educational attainments. As Braun and Stuhler (2017) explain, this test only allows us to reject direct grandparent effects. If we continue to find a positive maternal grandparent effect in regressions that condition on the status of both parents, we still cannot rule out that other omitted parental characteristics are driving the result.

significantly higher than the coefficient between fathers' and sons'. Column (2) reports the same qualitative result for the case in which parental input is mother's years of schooling. Finally, we run the same exercise this time including both father's and mother's years of schooling as parental inputs.²² Interestingly, as Column (3) displays, this time we find that among the interaction terms only the variable that we create by interacting mother's years of schooling and the female dummy variables is statistically significant. Namely, one year increase in mother's years of schooling is associated with about 0.2 years increase in son's education whereas it is associated with almost 0.5 years increase in daughter's education. This suggests that it is indeed the case that intergenerational education mobility is lower for females than males and the channel that creates the extra persistence in educational attainments of parents and daughters is mainly working through mothers.²³

Table 5 [about here]

5. Education Mobility and Socioeconomic Development

There is substantial heterogeneity across regions in Turkey regarding various aspects of socioeconomic development. For instance, GDP per capita in the richest province, Istanbul, is more than three times as large as the average GDP per capita in a group of poorest provinces. Similarly, average years of schooling is more than twice in Istanbul relative to the same set of provinces.²⁴ This large regional variation enables us to study how intergenerational mobility in education changes with the level of socioeconomic development. To do so, we conduct two separate exercises. The results from both show that education mobility is higher among females who grow up in more developed regions. That is, educational outcomes of daughters are related to their parents' educational outcomes to a lesser degree if they grow up in more

²² To be precise, we estimate the equation $S_{it} = \alpha + \beta_0 F_i + \beta_1 F_i * G_i + \beta_2 M_i + \beta_3 M_i * G_i + \beta_4 E_i + e_i$, where F, M, G, and E stand for father's years of schooling, mother's years of schooling, female dummy and male dummy, respectively.

²³ We also run three-generation mobility regressions by gender and find that the coefficient of grandparents' years of schooling is statistically significantly higher for daughters than sons.

²⁴ See Table A1 and Table A2 in the Appendix for more information on how various development characteristics compare across regions in Turkey.

developed regions. Interestingly, we do not find evidence for a similar relationship between development and mobility for males. We also find that the development level of place of residence during earlier years of childhood (ages 0-12) has much stronger association with mobility than the development level of place of residence during adolescence.

5.1 Socioeconomic Development

We begin by explaining how we measure socioeconomic development level of the places where individuals grew up in our sample. Turkey is divided into 81 administrative areas called provinces. The Turkish Ministry of Development periodically measures the level of socioeconomic development at the province level. As part of the measurement process, the ministry has developed an index that aims to summarize the level of development in each province. This index combines 61 distinct measures of socioeconomic development that can be classified under eight categories: demographics, employment, education, health, competitiveness and innovation, financial development, transportation and telecommunication availability, and quality of life. Larger values of this index indicate higher level of socioeconomic development.²⁵ In our analysis, we use the most recent index, SEGE-2011, which stands for Socio-economic Development Index for 2011, as a proxy for the development level of each province.

Our data includes separate information about the province of residence when individuals were between the ages of 0 and 6, 7 and 12, and 13 and 18, as well as, the current place of residence.²⁶ Since we are interested in people's educational attainment, the characteristics of the place where they grew up matter more than the characteristics of the current place of residence. For this reason, we assign each individual to the province he or she resided while growing up (from birth till age 18). Individuals who lived in multiple provinces

²⁵ SEGE-2011 index has a mean value of zero and a standard deviation of one. See Kalkinma Bakanligi (2013) for detailed information about how the index is constructed.

²⁶ If a respondent lived in more than one province during a certain age interval, then the survey asks them to single out the one they lived the longest. That is the province of residence we use for that age interval.

during this time are assigned to the province they resided between the ages of 7 and 12 in our baseline exercise. Later on, we also conduct exercises in which we assign individuals to their province of residence during ages 0-6 and 13-18, and their current place of residence. In all the estimation in Section 5, we include all individuals in the base data set who are 20 years or older. The main rationale for expanding the lower age limit is to increase the precision of estimates reported for each development region in Section 5.3.²⁷ The resulting sample of survey respondents has 4308 observations.

5.2 Parametric Specification

Our first analysis of the relationship between mobility and development takes a parametric approach. We extend the basic intergenerational regression equation (1) by adding an independent variable: an interaction term between parents' years of schooling and the socioeconomic development level of the province where children grew up. We also add development level of the province where children grew up itself into the regression equation as a control. The equation that we estimate for the parent-child years of schooling relationship becomes:

$$S_{1i} = \alpha + \beta_0 S_{0i} + \beta_1 (S_{0i} * DevInd_i) + \gamma DevInd_i + \varepsilon_i, \quad (5)$$

where $DevInd_i$ is the development index score for individual i 's place of residence between the ages of 7 and 12 and S_{1i} and S_{0i} are child's and parent's years of schooling as before. Notice that using SEGE-2011 to determine the socioeconomic development level of provinces where people grew up assumes that the development rank of a given province in 2011 provides a good proxy for that province's development rank when individuals were between the ages of 7

²⁷ The primary analysis sample we have used until this section includes individuals who are 25 years or older. When we replicate the analysis of Section 5 using our primary analysis sample, we obtain very similar results to those reported in this section. Similarly, when we replicate our main mobility estimates reported in Section 4 with the sample age limit set to 20 as we do in the current section, we obtain results very similar to those reported in Section 4. For example, Table 2 reports regression coefficient estimates of 0.734 (0.018) for father's education and 0.767 (0.021) for mother's education in columns 1 and 2. The corresponding estimates are 0.765 (0.016) and 0.799 (0.019) when the sample includes individuals 20 years of age and older.

and 12. In our context, this is not a strong assumption because the development ranks of Turkish provinces are very persistent over time.²⁸ Since development index varies at the province level we cluster standard errors at province level (81 provinces) in regressions that we report in this section.

The estimation results for equation (5) are presented in Table 6. The first column named Father-all refers to the regression specification in which parental education variable is fathers' years of schooling and the sample includes all observations in our main sample. The next five columns are defined similarly. The third row of Table 6 reports the coefficient estimates for $S_{0i} * DevInd_i$ interaction term, which is our main variable of interest. Looking at the third row of first and fourth columns, we see that the coefficient estimates are negative and significant, indicating that the relationship between parents' years of schooling and children's years of schooling is weaker in places with higher levels of economic development. In other words, educational outcomes of children who grew up in more developed provinces depend less on their parents'.

A closer look into Table 6 reveals an interesting caveat, however. The coefficients of the interaction terms in the regressions where the sample is restricted to males (Columns 2 and 5) are close to zero and insignificant whereas those for the female sample are significant and negative, and large in absolute value (Columns 3 and 6). These results imply that the level of development has a positive relationship with the intergenerational education mobility of daughters whereas evidence for such a relationship does not exist for sons.²⁹ As a side note,

²⁸ The development ranks and indices of provinces are also available for the years 1985, 1996 and 2003. We compute the correlation of province development ranks across years and find the correlation to be 0.94 between 1985 and 2011, 0.97 between 1996 and 2011, and 0.98 between 2003 and 2011. The corresponding correlations for development index values are 0.91, 0.96 and 0.97 respectively. We prefer to use the 2011 index because the number of provinces increased from 67 to 81 during the 1980-2000 period due to newly established provinces and the survey captures the place of residence information based on the new list of provinces. There have also been several improvements in the methodology of generating the development index values reflected in SEGE-2011. See Kalkinma Bakanligi (2013) for details.

²⁹ For robustness purposes, we replicate the analysis of Table 6 with SEGE-2003 development index. An advantage of using SEGE-2003 is that, in addition to the main development index, the Turkish Ministry of

notice that the estimate of γ which is the coefficient of the *DevInd* variable is much higher in father-daughter and mother-daughter regressions than it is in father-son and mother-son regressions, indicating that a rise in development increases daughters' years of schooling much more than sons'.

Table 6 [about here]

Recall that in our baseline analysis we assign individuals who live in multiple provinces while growing up to the province they resided between the ages of 7 and 12. We also conduct exercises in which we assign such individuals to their province of residence during ages 0-6 and 13-18, and their current place of residence. Table 7 summarizes our findings. We do not report the results of Father-son and Mother-son pairs because, perhaps as expected from the findings in Table 6, the relationship between development and mobility for sons is insignificant independent of which definition of area of residence we use.³⁰ The third row of Table 7 reports the coefficient estimates for $S_{0i} * DevInd_i$ interaction term, which is again our main variable of interest. The first four columns report the results for the Father-daughter specification. The third column displays the benchmark analysis in which the area of residence is defined as the province children resided during ages 7-12.

A comparison of the first four columns reveals that the development level of place of residence has much stronger association with mobility when we define place of residence as the place of residence during ages 0-6 and 7-12 rather than the place of residence during ages 13-18 (-0.058 and -0.056 vs. -0.033). The development level of area of residence has the weakest association with mobility when place of residence is taken to be the individual's

Development also publishes three subindices of the main index, which aim to measure the development level of provinces in the specific fields of manufacturing, education, and health. We confirm the finding that there is a significant association between mobility and development for daughters while there is no such relationship for males is robust using SEGE-2003 development index and each one of these subindices of development. See Table A4 in Appendix.

³⁰ In fact, for the same reason, we only report the results of Father-daughter and Mother-daughter pairs in the remainder of this section.

current area of residence (-0.029). An identical pattern is observed when one instead considers the results regarding the Mother-daughter specification, which are reported in Columns (5) – (8) of Table 7. These findings indicate that the place of residence during earlier years of childhood, ages 0-12, is more significant for daughters' mobility than their areas of residence in later years.

Table 7 [about here]

One may suspect that the significant negative association between the development level of the current place of residence and intergenerational mobility might be an artifact of the fact that many of the individuals in our sample have been residing in the same province all their lives. That is, higher development level of the current place of residence may be associated with higher mobility because it is highly correlated with place of residence at earlier ages. To see whether this is indeed the case, we restrict our sample to those whose current place of residence is different from place of residence during ages 0-6 and 7-12. When we run the regression given by equation (5) for this restricted sample, we find that the interaction term between the development level of current place of residence and parents' years of schooling becomes insignificant, which implies that there is no statistically significant association between the development level of current place of residence and intergenerational mobility. The results of this analysis are reported in Table 8. The association between development and mobility remains strong in the cases in which the place of residence is taken to be the individual's place of residence during ages 0-6 and 7-12. Interestingly, the association between development level of area of residence during adolescence (ages 13-18) and mobility also becomes insignificant in the migrant sample. These findings suggest that the area of residence at earlier ages of childhood is more important for intergenerational mobility of daughters than their area of residence in adolescence. This conclusion is in line with a broad set of empirical findings that support the view that there are "sensitive" periods in childhood during which

parental and environmental inputs into children have especially high marginal returns (Cunha and Heckman 2008).³¹ This finding is in contrast with Chetty and Hendren (2018a) who find that neighborhood exposure effects on upward mobility do not depend on the age of exposure. There may be several reasons why the two papers reach different conclusions. First, Chetty and Hendren (2018a) conduct their analysis using data from a developed economy, the United States, whereas we use data from a developing economy, Turkey. Second, while their data set is much richer than ours in terms of sample size and in the sense that they can follow children's movements across neighborhoods at an annual rate, it does not allow them to observe children's place of residence prior to age 8. Therefore, unlike the present study, they cannot test the relative importance of place of residence during earlier years of childhood.

Table 8 [about here]

Another, and perhaps a more standard, way of investigating whether there is a distinct association between the development level of area of residence during later stages of ones life and mobility is to keep the whole sample and consider a regression specification in which we modify equation (5) to include all the interaction terms between parents' years of schooling and the development level of the province children resided during ages 0-6, 7-12, 13-18 and the development level of their current place of residence. The equation that we estimate is the following:

$$S_{1i} = \alpha + \beta_0 S_{0i} + \sum_{j=1}^4 \{ \beta_{j,1} (S_{0i} * DevInd_{j,i}) + \gamma_j DevInd_{j,i} \} + \varepsilon_i, \quad (6)$$

where $j = 1, 2, 3, 4$ refer to the place of residence of children during ages 0-6, 7-12, 13-18, and current place of residence, respectively, and $DevInd_{j,i}$ is the development index score of individual i 's place of residence in period j . The results of this regression analysis are reported in Table 9. The table supports the finding from the migrant sample analysis that place of

³¹ For more detailed information on the importance of early childhood education in determining adult outcomes, see the review article of Elango et al. (2016).

residence during ages 13-18 and current place of residence do not have significant association with intergenerational education mobility.³² As expected, the place of residence during ages of 0-6 has a significant and negative association with intergenerational education persistence. Contrary to our earlier findings, the place of residence during ages 7-12 does not have a significant effect on mobility according to this regression analysis. It is important to note, however, that the correlations between development index values of place of residence during various stages of childhood are very high. For this reason, the results of regression given by equation (6) may suffer from multicollinearity, which decreases the credibility of the results given in Table 9.³³

To avoid this issue, we also run two alternative versions of the estimation specification given by equation (6). In the first one, we interact parental years of schooling with development levels of area of residence during ages 0-6 and current residence. In the second one, we interact parental years of schooling with development levels of area of residence during ages 7-12 and current residence. These regressions do not suffer from multicollinearity problem because correlations between development index values of current place of residence and place of residence during ages 0-6 and 7-12 are relatively low.³⁴ In both these regressions, the coefficient of the interaction term between place of residence during early childhood (be it ages 0-6 or 7-12) and parent's years of schooling is significant and negative while the coefficient of the interaction term between current place of residence and parent's years of

32 In fact, in the regression analyses reported in Table 9, the interaction terms between development level of *current* place of residence and mother's and father's years of schooling are marginally significant (the p-values for these interaction terms are 0.09 and 0.08, respectively.) The significance of these interaction terms completely disappear and the coefficient estimates become virtually zero when we replicate the analysis of Table 9 with the migrant sample defined for the analysis in Table 8.

33 For example, when we consider interaction terms between father's years of schooling and the development level of the province children resided during the ages of 0-6, 7-12, 13-18, pairwise correlations between these interaction terms corresponding to pairs 0-6 and 13-18, 7-12 and 13-18, and 0-6 and 7-12 are equal to 0.75, 0.82 and 0.91 respectively. In the estimation of equation (6) this leads to high values of Variance Inflation Factors (VIF). For example, for the father-daughter regression, five of the right hand side variables have VIF values that range between 11.9 and 28.3, which indicates high degree of collinearity.

34 In these regressions, all VIFs are below 4, and hence, there is no indication of high collinearity. The results of these estimations are available upon request.

schooling is not. This is further evidence that place of residence during earlier years of childhood are more important for intergenerational education mobility.

Table 9 [about here]

5.3 Nonparametric Specification

The analysis carried out in equation (5) assumes a specific parametric relationship between mobility and development. In this section, we provide further support to the relationship between economic development and education mobility by conducting a nonparametric analysis in which we estimate intergenerational regression coefficients at a regional level. Due to limitations regarding the size of our sample, we do not run a separate regression for each province. Instead, we group provinces into 15 regions that have similar socioeconomic development index values, and estimate education mobility at the regional level. By construction, each region consists roughly 1/15 of all the individuals in the sample. We have on average 290 observations in each development region. Appendix explains the procedure we use to group provinces into development regions in detail. It also includes Table A1 that presents the list of provinces within each development region.³⁵

Figure 2 presents mobility estimates for each region. The top left panel depicts how the level of persistence between fathers' and children's years of schooling changes with economic development. Each blue dot represents a development region. The horizontal axis indicates the mean development score of the region whereas the vertical axis indicates the regression coefficient of fathers' years of schooling estimated for that region. The red-green band provides 95% confidence for the corresponding estimate of the regression coefficient. The orange line provides a linear fit to the estimated regression coefficients. The bottom left panel

³⁵ It is important to note that while this procedure collects provinces that are similar in terms of their levels of development in a group, it does not necessarily imply geographic proximity between provinces within a group. For example, for development group 10, while Adana is located in Southwestern Turkey, another province in this group, Sakarya, is located in Northwestern Turkey.

of Figure 2 displays the same graph taking mothers' years of schooling as the independent variable in the intergenerational education regression. Both graphs show that there is a significant and negative association between the level of development and the intergenerational persistence in education, confirming our earlier findings. For instance, a closer look at the graph on the top left panel of Figure 2 reveals that one year increase in fathers' years of schooling increases children's years of schooling by around 0.9 years in the least developed region whereas the same number in the most developed region (Istanbul) is about 0.6.³⁶

The top right panel displays how the level of persistence between fathers' and daughters' years of schooling changes with economic development. The bottom right panel shows the same relationship for the mother-daughter specification. We see that there is a significant negative association between development and intergenerational persistence in education for daughters. The top middle and bottom middle panels display the relationship between development and intergenerational persistence for sons. Interestingly, we find that for male children the association between mobility and economic development is less significant than that for female children. In fact, when we conduct alternative regional exercises in which we group provinces in Turkey into 5, 10, 20, and 25 regions (instead of 15 as it is in the benchmark analysis), we find that the positive association between development and mobility remains consistently statistically significant across all exercises only for females. This provides further support to our finding in the parametric analysis that the relationship between development and mobility is primarily important for females.

Figure 2 [about here]

An international comparison In this section, we compare the mobility estimates in our 15 development regions to those in the rest of the world as reported in Hertz et al. (2007). For

³⁶ In the section entitled "An international comparison", we show that a similar negative association holds between intergenerational correlation coefficient and the development index, which implies that the intergenerational regression coefficient does not vary with development only because the distribution of educational attainment varies with development.

purposes of this comparison, we follow the methodology of Hertz et al. (2007) and set average years of schooling of parents as the parental education variable. We also restrict the sample to individuals who are between the ages of 20 and 69 following Hertz et al. (2007). We estimate the correlation between the average years of schooling of parents and children's years of schooling for each region. Table A3 in Appendix lists these correlations for all regions.³⁷

Figure 3 presents a heat map of the correlations across Turkey. Lighter colors represent higher levels of mobility. A quick glance at the figure reveals that the more socioeconomically developed Western regions of Turkey feature lighter colors, and hence, higher mobility. Moreover, we see that there is significant heterogeneity in intergenerational mobility across the country. On the one hand, intergenerational education correlation is higher than 0.75 in the least developed region (that consist of the seven provinces in darkest color). This is a degree of persistence in education that is much higher than the least mobile country as it is reported in Table 2 of Hertz et al. (2007), Peru, with correlation of 0.66. On the other hand, the intergenerational correlation is as low as 0.35 in the most mobile region, which is comparable in magnitude to the relatively higher mobility countries of Europe such as Sweden or Belgium. It is important to note that the only goal of this international comparison is to provide a sense of the magnitude of regional differences in education mobility across Turkey; we do not mean to compare the magnitudes of education mobility between regions in Turkey and other countries as the mechanisms determining mobility in a region of a country vs. a whole country might be quite different.

Figure 3 [about here]

The results in this section indicate that there are large differences across regions in intergenerational mobility. These differences in intergenerational mobility across regions have

³⁷ We compute the correlation coefficient between average parental years of schooling and children's years of schooling after adjusting each of these variables by their respective standard deviations. Similar to our results in Figure 2 there is a strong relationship between the level of development and this alternative measure of mobility. A regression of the correlation coefficient on mean development index yields an estimated slope coefficient of -0.0524 with a standard deviation of 0.012.

an important implication for the intergenerational mobility literature. There has been a long debate on the origins of intergenerational correlations (see for example Sacerdote, 2002). Some studies emphasize the role of nature in the transmission of economic outcomes across generations while others focus on the importance of nurture (family investments, institutions, social environments, among others). Given that this study is conducted within a single country where the inheritability of genes is unlikely to vary significantly across regions, the regional variation in the estimated intergenerational mobility coefficients provides evidence for the role of nurture in shaping mobility.³⁸

Several mechanisms may lead to lower mobility in less developed regions. These regions may be characterized by lower educational opportunities due to, for example, lower supply of schools and higher costs of access to schooling. Therefore, in the context of less developed regions, parental investments may play a more significant role in shaping children's educational outcomes in comparison to more developed regions. Also, in less developed regions gender culture may be less favorable for female education and employment. Parental education may, therefore, be playing a more important role for overcoming these challenges in less developed regions. Although we cannot identify such causal mechanisms, in the next section, we investigate several important factors that may give rise to the correlation between the level of development and intergenerational mobility established in this section.

³⁸ Turkey consists of two main racial groups: Turks and Kurds. There is no official statistic on the fraction of Kurdish population in Turkey. According to the three surveys by KONDA, a public opinion research and consultancy company conducted over the 2006-2013 period, the fraction of Kurdish people in the entire population ranges from 15.7% to 18.3%. These estimates are based on the reported mother tongue. (<http://www.radikal.com.tr/yazarlar/tarhan-erdem/turkiyede-kurtler-ne-kadardir-2-1130993/>). Using the same definition, the fraction of Kurds in our sample is about 15.3%. Moreover, Kurds dominantly live in the Eastern provinces of Turkey. To investigate whether the lower intergenerational mobility found in these regions are related to the racial composition of these regions, we also estimate mobility for the two racial groups separately. We find no statistically significant difference between the mobility of Turks and Kurds, which strengthens the case for the role of nurture in shaping mobility. We have also replicated our parametric and non-parametric analysis reported in Sections 5.2 and 5.3 by excluding individuals of Kurdish ethnic origin. All results remain qualitatively the same for this restricted sample. We exclude a detailed discussion of the analysis of mobility by ethnicity from this paper for the sake of brevity.

5.4 A First Step toward Understanding Why Development Matters

In this section, we aim to take a first step toward understanding why development is positively associated with intergenerational mobility. To do so, we study what characteristics of the environments in which children grow up are more important for mobility. Specifically, we focus on five groups of characteristics: availability and quality of educational facilities, social capital, cultural attitudes toward women, childhood home resources, and educational inequalities. Regarding each characteristic, we construct a variable which corresponds to the regional average for that characteristic at the level of 15 development regions introduced in Section 5.3. We standardize each of these variables so that they are all distributed with mean zero and variance one across regions. We do this standardization in order to ease comparison of magnitudes of coefficient estimates across different regressions. We then estimate the following relationship:

$$S_{1i} = \alpha + \beta_0 S_{0i} + \beta_1 (S_{0i} * Var_i) + \gamma Var_i + \varepsilon_i, \quad (7)$$

where Var_i is the regional mean of the characteristic of interest in region where individual i resided between the ages of 7 and 12.³⁹ Assigning individuals regional means (rather than the value of variables for each individual) serves two purposes. First, quite a few of our variables are binary at the individual level (such as whether there was a high school present in their area of residence while growing up, which takes value 1 or 0 depending on if the answer is yes or no), and using regional averages converts these into continuous variables, which then enables us to standardize them. Second, using regional averages in the estimation of equation (7) is consistent with the estimation of equation (5) in Section 5.2 since the development index value is also a regional variable.

³⁹ In equation (7), each variable Var_i has 15 distinct values, and all the individuals that reside in the same region share the same value. This leads to few cluster problem which leads to a bias in standard errors (Cameron et al. 2008). Angrist and Pischke (2008) suggest that 42 clusters as large enough to achieve accurate statistical inference with cluster robust variance estimator. We also construct regional characteristics at the province level, which leads to 81 clusters. Using variables constructed at the province level in estimating equation (7), we get qualitatively same results. Due to data limitations discussed in Section 5.3, we prefer to present results based on 15 regions.

We use two sets of data sources to compute regional averages of these characteristics. First, our data set includes individual level information on many characteristics of childhood environments between the ages of 7 and 12. For each characteristic, we aggregate individual level information to construct a regional average of that childhood environment characteristic. Second, regarding a minority of characteristics for which we do not have information from our data set, we use regional statistics from other sources.

We find that higher school availability, higher level of social capital, more favorable cultural attitudes toward women, higher level of home resources, and lower degree of educational inequalities are associated with lower intergenerational education persistence for daughters. We do not find similar evidence for an association between these characteristics and mobility for sons. For this reason, we only present regression results for daughters in the main text.⁴⁰ This is in line with the lack of evidence for a relationship between development and mobility for sons that we present earlier. Moreover, we find that these environmental characteristics matter much more for intergenerational mobility across mother-daughter pairs relative to father-daughter pairs. This is again in line with the finding from Table 6 that development has a much stronger association with mother-daughter education persistence than father-daughter education persistence. Finally, comparison of the magnitudes of interaction terms seems to indicate that social capital and childhood home environment characteristics have a stronger association with mobility of daughters than cultural attitudes toward women.

Availability and quality of educational facilities Our survey asks individuals whether there was a high school present in their area of residence when they were between the ages of seven and twelve. For a given development region, we construct the value of our availability of educational facilities variable as the mean of the dummy variable that equals 1 if the individual reports the presence of a high school and 0 otherwise. Our regional quality variable is the

⁴⁰ Table 5A in the Appendix displays the relationship between regional characteristics and mobility for sons.

student to teacher ratio in all the schools in a given region and is taken from the Turkish Ministry of Education National Education Statistics. The second column of Panel A of Table 10 presents the results of the regression analysis given by equation (7) in which Var_i equals the availability of educational facilities variable while the third column presents the findings for the same regression equation when Var_i equals the quality variable, both for mother-daughter pairs. In each regression, our main interest is the coefficient estimate of the interaction term, which is given by the third row in Table 6. As the third row displays, this term is statistically significant and negative for school availability variable, which means that the persistence in years of schooling across mothers and daughters is significantly lower in regions in which schooling is more available. The coefficient estimate of the interaction term between mother's years of schooling and school quality variable is insignificant and positive, which indicates that lower school quality in childhood environment is associated with higher persistence in mothers' and daughters' educational outcomes. Panel B reports regression results regarding father-daughter pairs. A comparison of the coefficient estimates in the two panels reveals that both availability and quantity of schooling matter more for mother-daughter education persistence than father-daughter persistence.

A mechanism that would give rise to the significant negative association between intergenerational persistence in education and school availability for daughters might be the following. In regions where nearby schooling facilities is a rare event, parents have to either send their children to distant places or arrange for housing around available schools. This would mean there are additional costs of sending children to schools in areas that are deprived of schooling. These additional costs might be prohibitively high in terms of sending all children to school in cases of families with low education and, hence, income (especially in the presence of borrowing constraints). In such circumstances, it is very likely that the educational attainment of sons would be a priority over that of daughters (see for ex. Kingdon (2005) for evidence of bias in educational resource allocation in the Indian context). In more educated

families with possibly higher financial resources, both sons and daughters may attend schools. This introduces an additional channel through which educational outcomes of parents would covary with educational outcomes of daughters in areas with lower availability of schooling. This can also provide a partial explanation to why intergenerational persistence is higher for daughters in less developed regions. As columns 7 and 8 of Table A2 in Appendix reveals, unavailability of schooling facilities is more prominent in less developed regions, indicating this additional channel for intergenerational persistence might be more prevalent in those regions.

Social capital The presence of individuals with high levels of education may constitute an important characteristic of childhood social environments (see, for instance, Bricheno and Thornton; 2007). Relatives and family friends (or other people in families' social networks) with high education may impact the educational attainments of children in at least two ways. First, by providing role models, they may increase educational aspirations of children, possibly leading to higher educational attainment. Second, they may affect parental decisions regarding children's education by shaping parents' opinions toward children's education. We have three variables that aim to measure the level of social capital in children's social environments. Using information from 2000 Census data on years of schooling and degree completed, we compute (i) the ratio of people with college degrees to those with primary school degrees and (ii) average years of schooling for each region. Our survey also includes a question that asks the respondents the educational attainment of the relative with highest education level when they were between the ages of 7 and 12. For each region, we compute the mean of this variable across observations in that region. Columns (4)-(6) of Table 10 report the results of three regression analyses in which Var_i equals each one of these variables. We find that intergenerational education persistence is lower for daughters in regions where the average education level (be it measured by the ratio of college graduates to primary school graduate or by average years of schooling) and the years of schooling of the relative with highest education

is higher. Again, a comparison of panels A and B indicate that greater social capital in childhood environments is more important for persistence of educational outcomes between mothers and daughters than fathers and daughters.

Gender culture In a developing country context gender culture and social norms against female education and employment may have effects on the educational attainments of women (King and Hill, 1993). We construct three gender culture variables using our survey data that aims to measure such norms. First, the survey asks respondents how much say they had in choosing their spouses. The set of answers can be grouped into two categories: the first category includes answers that indicate that the individuals themselves had the most say whereas the second category includes answers that indicate it was the family's opinion that counted the most. We construct a dummy variable that takes the value 1 if the answer is in the first category and 0 if it is in the second. Then, for each region, we compute the mean of this variable. Thus, the higher this gender culture variable is for a region, the larger is the say of people in their own marital decisions (as opposed to their families). The underlying assumption is that this variable is correlated with cultural attitudes against female education; that is, places in which parents let their children choose their spouses also feature more favorable climates for female education and employment. Second, the survey asks respondents whether they approve female employment. For a given development region, we construct the value of our approval of female employment variable as the mean of the dummy variable that equals 1 if the individuals report they approve female employment and 0 otherwise. Third, the survey asks respondents whether males should have a priority to receive education over females assuming they earn same income with the received education. We construct our third gender norm variable for a development region by computing the fraction of those who refused to give men priority over those who gave such priority. Columns (7)-(9) of Table 10 report the results of three regression analyses in which Var_i equals each one of these variables. The third row of Panel A shows that the coefficient estimates regarding the interaction term between all the three gender norm

variables and mother's education are all negative, implying that intergenerational education persistence between mothers and daughters is lower in regions where gender norms are more favorable toward women. Interestingly, the choosing spouse variable – which is a somewhat indirect measure of gender norms toward female education - has a significant association with mobility while the direct measures of approval of female employment and priority of female education do not. This is possibly due to the fact that the last two variables reflect gender norms of respondents (children) while the choosing husband variable is informative about gender norms of parents as well. Finally, a quick glance at columns (7)-(9) of Panel B reveals that gender norms are in fact have no significant effect on intergenerational education persistence between fathers and daughters.

Why is it the case that regions that feature more favorable gender culture also feature higher mobility for daughters? One explanation might be that social norms that are against female education and employment introduce additional non-pecuniary costs for parents to send their daughters to school. Under the assumption that these costs are higher in families where parents are less educated, gender culture introduces a further channel of intergenerational persistence between educational outcomes of daughters and their parents. One may expect this additional channel of persistence to be stronger in regions where gender culture is less favorable toward females. This may also provide a partial explanation to why intergenerational persistence is higher for daughters in less developed regions. Column 9 of Table A2 shows that our gender culture variable, the fraction of people who choose their spouses, is higher in regions that are more developed, implying this additional channel might be stronger in less developed regions.

Childhood home environment There is a large body of empirical evidence which documents that parental (home) investments during childhood are an important factor in

determining adult outcomes.⁴¹ As such, differences in home environments is one reason suggested for the high correlations of economic outcomes generations (Solon 1992). To investigate importance of home environments for mobility, we first construct four home environment variables using our survey data. The survey asks respondents if there was an encyclopedia or a musical instrument at home while growing up. For a given development region, we construct the value of our presence of encyclopedia variable as the mean of the dummy variable that equals 1 if the individuals report there was encyclopedia and 0 otherwise. We construct the presence of our musical instrument variable exactly the same way. The survey also asks people how often they went to a theatre or a cinema while growing up and how many siblings they have. We average the responses to these questions across respondents in a given region to get two more regional average variables. We estimate the regression specification given by equation (7) separately for each home environment variable. In each regression, Var_i equals one of the home environment variables. The results of these estimations are given in columns (10)-(13) of Table 10. The coefficient estimate of the interaction term between parent's education and all the home environment variables are significant and negative, except for number of siblings variable. This indicates that intergenerational education persistence is lower for daughters who grew up in regions where home environments are more likely to possess encyclopedias and musical instruments and where children are more likely to be taken to cultural events like movies or theatres. The coefficient estimate of the interaction term between parent's education and average number of siblings is significant and positive, indicating that in regions where families tend to have more children, educational outcomes persist to a higher degree across generations. A comparison of coefficient estimates of interaction terms in panels A and B reveal that home environments matter more for mobility across mother-daughter pairs than father-daughter pairs.

⁴¹ Todd and Wolpin (2007) provides evidence on the importance of parental investments in explaining children's cognitive test score gaps. For more direct support of the effect of parental investments in determining future outcomes see Cunha and Heckman (2008).

Educational inequalities We have two measures of educational inequality. The first one is the Gini index of years of schooling among the parents of survey respondents computed at province level using our survey data. The second one is the Gini index of years of schooling among adults (individuals aged 24+) in each province using 2000 Census data. In order to analyze the relationship between intergenerational mobility in education and educational inequalities, we estimate the regression specification given by equation (7) for the cases in which Var_i equals regional mean of parental years of schooling Gini and years of schooling Gini among adults (as before, averages taken at the level of 15 development regions introduced in Section 5.3). The last two columns of Table 10 report the results. Panel A displays that the interaction term between both inequality measures and mother's years of schooling is significant and positive, implying that in regions where educational outcomes are distributed more unequally in either parents' generation or all adults, intergenerational mobility in education is also lower across mothers and daughters. This provides evidence for the existence of a Great Gatsby curve in education. This finding is in line with the evidence presented by Krueger (2012) and Corak (2013) who show that countries with more inequality as measured by income Gini coefficients also display less intergenerational income mobility. Guell et al. (2015) shows a similar result in the context of Italian provinces.

Table 10 [about here]

5.5 Absolute Mobility and Development

In this section, we investigate the relationship between development and two measures of absolute mobility. The first one is the bottom-to-top probability, which measures the probability that a child has at least a two-year college degree when the parent has less than primary school education.⁴² Figure 4 presents the results for bottom-to-top probability. We observe that probability of going from the bottom of the education distribution to its top

⁴² Less than primary school education category includes those who attended primary school but did not obtain a diploma, those who never attended formal school but are literate, and those who are illiterate.

increases significantly with socioeconomic development. For instance, the top left panel of the figure shows that only 5% of the children whose fathers have less than primary school education receive a college/university degree in the least developed region whereas the same number is 15% in the most developed region. Children born into parents with lowest education have three times more chance of going to college in the most vs. least developed regions. The other mobility measure is the probability that a child has less than primary school education when the parent also has less than primary school education. This measure, which we call bottom-to-bottom probability, summarizes the degree of persistence at the bottom of the education distribution. Figure 5 shows that the probability of being stuck at the bottom of the education distribution is much higher in less developed regions. For example, the top left panel in Figure 5 indicates that children who grew up in the provinces that are in the least developed region and whose fathers have primary school education or less have a 90% chance of having primary school education or less whereas the probability of getting stuck at the bottom of the education distribution is 60% in the most developed region.⁴³

Figure 4 [about here]

Figure 5 [about here]

⁴³ We investigate the relationship between development and bottom-to-top and bottom-to-bottom probabilities using a parametric approach as well. Our findings here echo our findings in the non-parametric approach. To investigate the effect of development on bottom-to-top probability, we restrict the sample to observations in which parents have less than primary school education and estimate:

$$Y_i = \alpha + \gamma DevInd_i + \varepsilon_i, \quad (8)$$

where Y_i is a variable that takes the value 1 if the children's education level is college or university and 0 otherwise. We estimate equation (8) separately for Father-all, Mother-all, Father-son, Mother-son, Father-daughter, and Mother-daughter pairs. We do not present the results of these estimations in the current paper for the sake of brevity, but instead briefly discuss them in this footnote. We find that the estimates of γ are highly significant and positive in all pairs, showing that the probability of reaching the top of the education distribution for children coming from families where the parents are at the bottom of the education distribution is higher in regions with higher level of economic development. We also investigate how bottom-to-bottom probability changes with economic development using specification (8) for the same restricted sample with Y_i taking the value 1 if the children have less than primary school education and 0 otherwise. The results show that the bottom-to-bottom probability is lower in regions with higher economic development. The results of the bottom-to-top and bottom-to-bottom estimations are available upon request.

6. Conclusion

The main contribution of this paper is to establish a relationship between economic development and intergenerational mobility. Both our parametric and nonparametric analyses show that educational outcomes of daughters who grow up in more developed regions depend less on their parents' educational outcomes. Interestingly, we do not find evidence for a similar relationship between development and mobility for males. Moreover, the development level of place of residence during early childhood (ages 0-12) has a significantly stronger positive association with intergenerational education mobility compared to development level of place of residence during adolescence. This finding is in line with a large body of empirical evidence that supports the view that early childhood environments are especially important for economic outcomes later in life.

We also take a first step toward understanding why development is positively associated with intergenerational mobility. To do so, we investigate what characteristics of the environments in which children grow up are more important for intergenerational mobility. We find that regions that feature higher school availability, greater level of social capital, more favorable cultural attitudes toward women, higher level of home resources, and lower degree of educational inequalities are associated with higher intergenerational mobility of daughters. We do not find similar evidence for an association between these regional characteristics and mobility of sons, a finding that is in line with the lack of evidence for an association between development and mobility of sons. While these exercises cannot establish causal relationships, they provide useful hints about the underlying causal mechanisms that lead to the association between development and intergenerational education mobility.

7. References

- ANGRIST, J.D, Pischke, J., Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press, (2008).
- AZAM, M., and V. Bhatt, Like Father, Like Son? Intergenerational Education Mobility in India, *Demography*, 52 (6), 1929–1959 (2015).
- BEHRMAN, J. R., Taubman P., Intergenerational Earnings Mobility in the United States: Some Estimates and a Test of Becker's Intergenerational Endowments Model, *Review of Economics and Statistics*, 67, 144-151, (1985).
- BLACK S. E., Devereux P. J., Recent Developments in Intergenerational Mobility, *Handbook of Labor Economics*, ed: Ashenfelter O., Card D., Elsevier, 1487-1541, (2011).
- BRAUN S., Stuhler J., The Transmission of Inequality Across Generations: Testing Recent Theories with Evidence from Germany, *The Economic Journal*, 128 (609), 576-611, (2018).
- BRICHENO, P., Thornton, J., Role model, hero or champion? Children's views concerning role models, *Educational Research*, 49 (4), 383-396, (2007).
- CAMERON, A.C., Gelbach, J.B. and Miller, D.L, Bootstrap-Based Improvements for Inferences with Clustered Errors, *Review of Economics and Statistics*, 90 (3), 414-427, (2008).
- CHIANG, Y. and H. Park, Do grandparents matter? A multigenerational perspective on educational attainment in Taiwan, *Social Science Research* 51, 163–173, (2015).

- CORAK, M., Income Inequality, Equality of Opportunity, and Intergenerational Mobility. *Journal of Economic Perspectives*, 27 (3), 79–102, (2013).
- CHETTY R., Hendren N., Kline P., Saez E., Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States, *Quarterly Journal of Economics* 129 (3), 1553-1623, (2014).
- CHETTY R. and Hendren N., The Effects of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects, *Quarterly Journal of Economics*, 133 (3), 1107-1162, (2018a).
- CHETTY R. and Hendren N., The Effects of Neighborhoods on Intergenerational Mobility II: County Level Estimates, *Quarterly Journal of Economics*, 133 (3), 1163-1228, (2018b).
- CUNHA F. and Heckman, J., Formulating, Identifying and Estimating Technology of Cognitive and Noncognitive Skill Formation, *The Journal of Human Resources*, 43 (4), 738-782, (2008).
- DICKENS, W.T., Error components in grouped data: is it ever worth weighting? *Review of Economics and Statistics* 72 (2), 328–333 (1990).
- DUNN, C. E., The Intergenerational Transmission of Lifetime Earnings: Evidence from Brazil, *The B.E. Journal of Economic Analysis & Policy*, 7 (2), 1-42 (2007).
- ELANGO, S., Garcia, J., Heckman, J., and Hojman, A., Early Childhood Education, in *Economics of Means-Tested Transfer Programs in the United States*, 2, 235-297, National Bureau of Economic Research, Inc, (2015).

- EMRAN, M. S. and F. Shilpi, Intergenerational Occupational Mobility in Rural Economy: Evidence from Nepal and Vietnam, *Journal of Human Resources*, 46 (2), 427-458 (2011).
- EMRAN, M. S., and F. Shilpi, Gender, Geography and Generations: Intergenerational Educational Mobility in Post-Reform India, *World Development*, 72 (C): 362–380 (2015).
- GUELL M., Pellizzari M., Pica G., Rodriguez Mora J. V., Correlating Social Mobility and Economic Outcomes, *The Economic Journal*, 128 (612), 353-403 (2018).
- HEIDRICH, S., Intergenerational Mobility in Sweden: a Regional Perspective, *Journal of Population Economics*, 30 (4), 1241-1280, (2017).
- HERTZ T., Jayasundera T., Piraino P., Selcuk S., Smith N., Verashchagina A., The inheritance of educational inequality: International comparisons and fifty-year trends, *The B.E. Journal of Economic Analysis & Policy*, 7 (2), Article 10, (2007).
- HNATKOVSKA, V., Lahiri A, and Paul S. B., Breaking the Caste Barrier: Intergenerational Mobility in India, *Journal of Human Resources*, 48 (2): 435-73 (2013).
- JAEGER, M. The Extended Family and Children's Educational Success, *American Sociological Review* 77(6) 903–922, (2012).
- KALKINMA BAKANLIĞI (KB), İllerin ve Bölgelerin Sosyo-ekonomik Gelişmişlik Sıralaması Araştırması, Ankara, (2013).
- KERR S. P., Lucas, R. E. B., Intergenerational Income Immobility in Finland: Contrasting Roles for Parental Earnings and Family Income, *Journal of Population Economics*, (26), 1057-1094, (2013).

- KIRDAR M., Explaining Ethnic Disparities in School Enrollment in Turkey, *Economic Development and Cultural Change* 57(2) pp. 297-333 (2009).
- KING, E. M. and M. A. Hill, *Women's Education in Developing Countries: Barriers, Benefits, and Policies*, John Hopkins University Press, USA, (1993).
- KINGDON, Geeta Gandhi, "Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure," *Economic Development and Cultural Change* 53(2): 409–51, (2005).
- KNIGHT, J., T. Sicular, X. Yue, *Educational Inequality in China: The Intergenerational Dimension*, University of Western Ontario, Centre for Human Capital and Productivity (CHCP) Working Papers 201113, University of Western Ontario, Centre for Human Capital and Productivity (CHCP) (2013)
- KRUEGER, A., "The Rise and Consequences of Inequality in the United States," Speech at the Center for American Progress, Washington, DC, January 12, 2012.
- KORKMAZ A., Sosyal Hareketlilik: Eğitim ve Mesleğin Sosyal Hareketliliğe Etkisi, *Sosyoloji Konferansları*, 30, İstanbul, (2005).
- LINDAHL, M., Palme M., Massih S. S., Sjögren A., Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations, *Journal of Human Resources*, 50(1), 1-33, (2015).
- ÖZDURAL S., Intergenerational Mobility: A comparative study between Turkey and the United States, *Economics Letters*, 43, 221-230, (1993).
- PIRAINO, P., Intergenerational Earnings Mobility and Equality of Opportunity in South Africa, *World Development*, 67(C), 396-405 (2015).

RIDGE, J. M., Three Generations, Mobility in Britain Reconsidered, ed: J. M. Ridge, Oxford, Clarendon Press, (1974).

SACERDOTE, B., The Nature and Nurture of Economic Outcomes”, American Economic Review, American Economic Association, 92 (2), 344-348, (2002).

SOLOMON G. Intergenerational Income Mobility in the United States, American Economic Review 82 (3), 393-408, (1992).

SOLOMON G., Haider, S.J., Wooldridge, J., What are we weighting for? NBER working paper 18859, (2013).

SOLOMON G., What Do We Know So Far about Multigenerational Mobility? NBER Working Paper No: 21053, (2015).

TANSEL A., Determinants of school attainment of boys and girls in Turkey: individual, household and community factors, Economics of Education Review, 21, 455–470, (2002).

TANSEL, A., Intergenerational Educational Mobility in Turkey, IZA DP No. 9590. (2015).

TODD, P.E., and Kenneth Wolpin, The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps, Journal of Human Capital, 1 (1), 91-136, (2007).

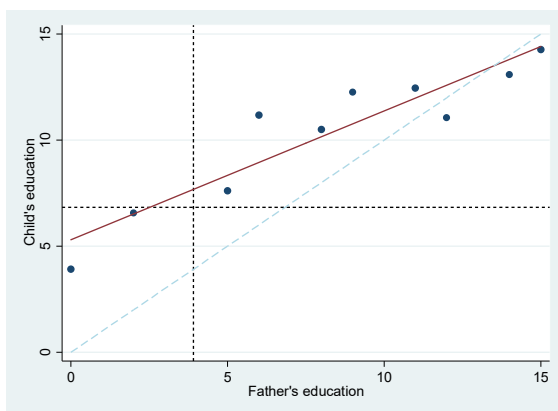
UYSAL-KOLAŞIN G., Yontar M. Y., Babanın Eğitimi Çocuğun Ücretini Etkiliyor, Betam Araştırma Notu 12/128, 1-6, (2012).

WARREN JR and RM Hauser. Social stratification across three generations: New evidence from the Wisconsin longitudinal study, *American Sociological Review*, 62 (4), 561–572, (1997).

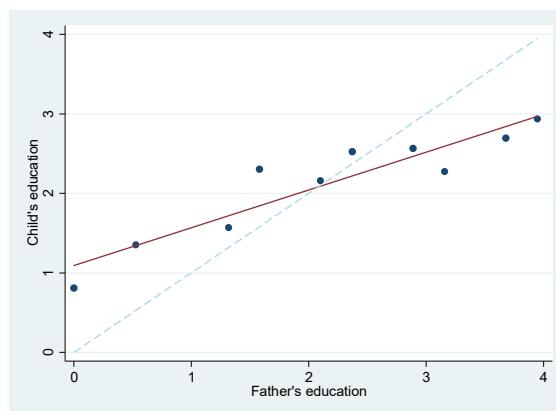
ZENG, Z. and Y. Xie, The Effects of Grandparents on Children's Schooling: Evidence from Rural China, *Demography*, 51 (2), 599–617, (2014).

Figure 1: Mean child years of schooling by father's years of schooling

Panel A: Years of schooling

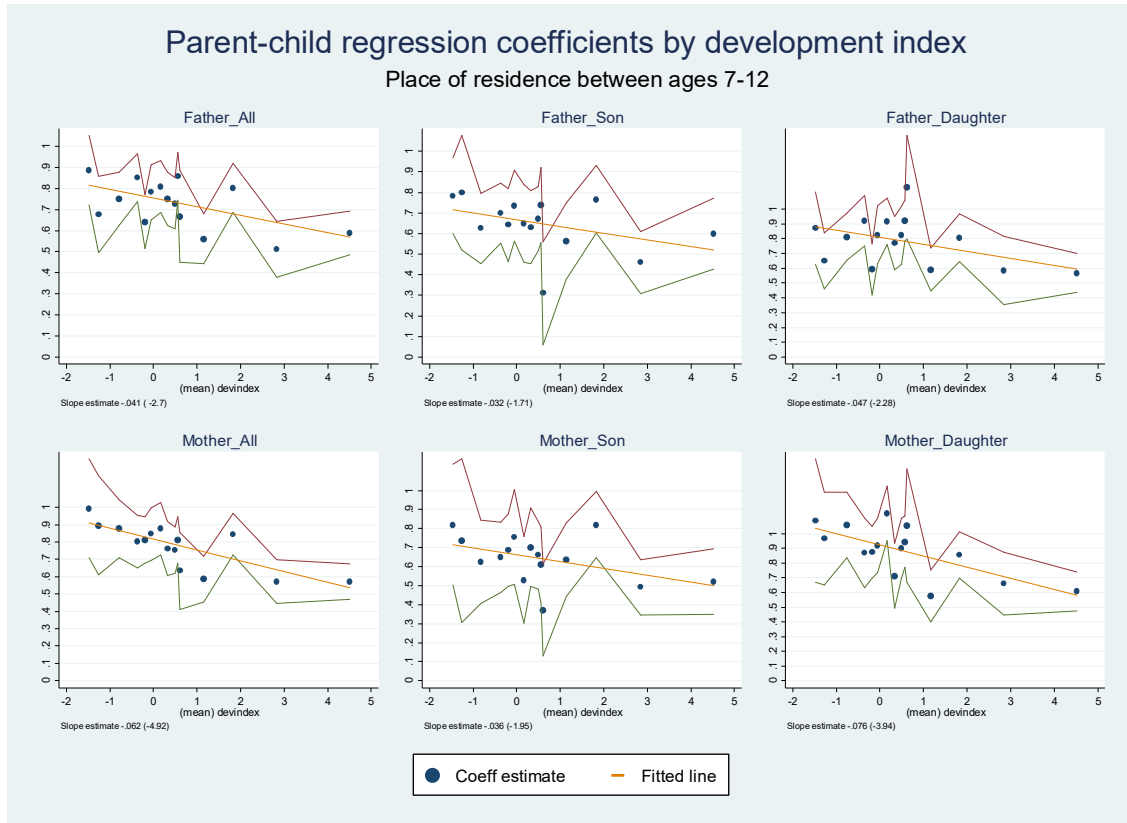


Panel B: Adjusted years of schooling



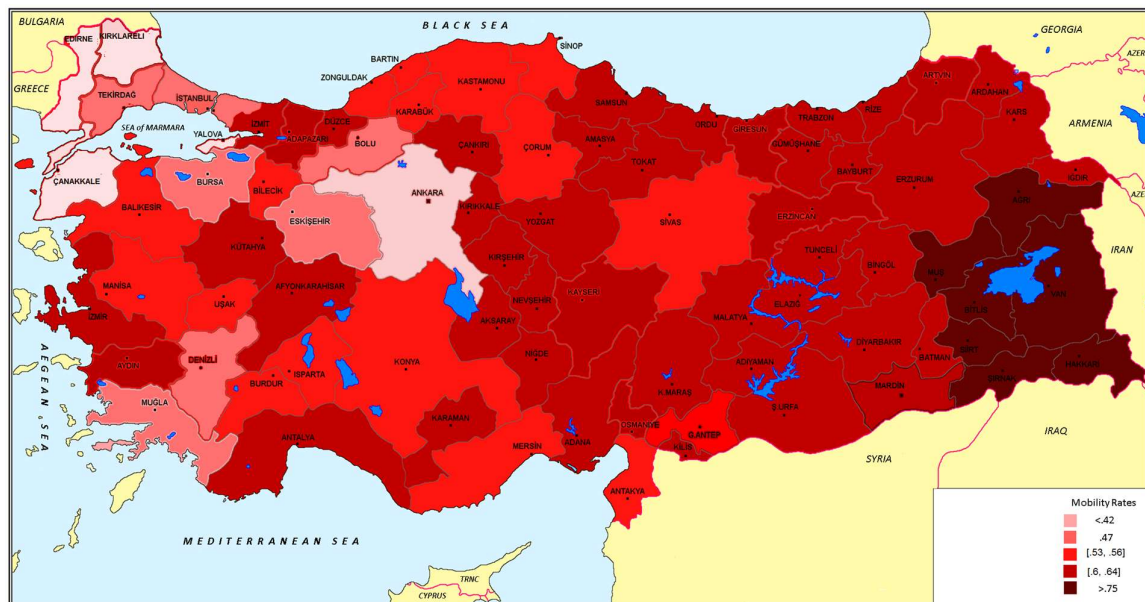
Notes: Panel A plots mean years of schooling of children as a function of fathers' years of schooling. Panel B presents the same plot taking adjusted years of schooling as education variable.

Figure 2: Education Mobility by Level of Development



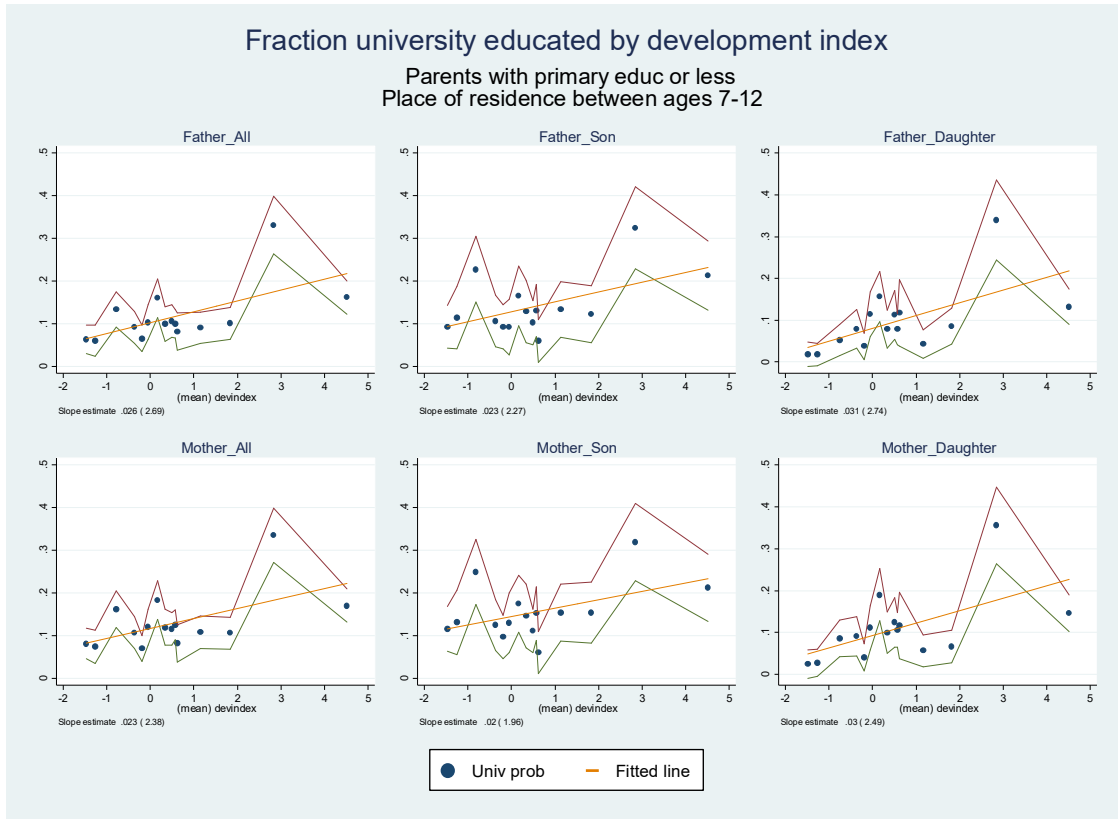
Notes: The top left panel presents estimates of the regression coefficient of fathers' years of schooling in the regression where children's years of schooling is the dependent variable and fathers' years of schooling is the independent variable for all 15 development regions. Each blue dot represents a development region. The horizontal axis indicates the mean development score of the region whereas the vertical axis indicates the regression coefficient of fathers' years of schooling estimated for that region. The red-green band provides 95% confidence for the corresponding estimate of the regression coefficient. The orange line provides a linear fit to the estimated regression coefficients. The number below each graph refers to the estimated slope of the fitted line. The number in parenthesis is the associated t-stat value. The bottom left panel displays the same graph taking mothers' years of schooling as the independent variable in the intergenerational education regression. The top middle and bottom middle panels display intergenerational mobility estimates in all 15 regions for the male subsample while the top right and bottom right panels display regional mobility estimates for the female subsample.

Figure 3: Heat Map of Intergenerational Mobility in Turkey



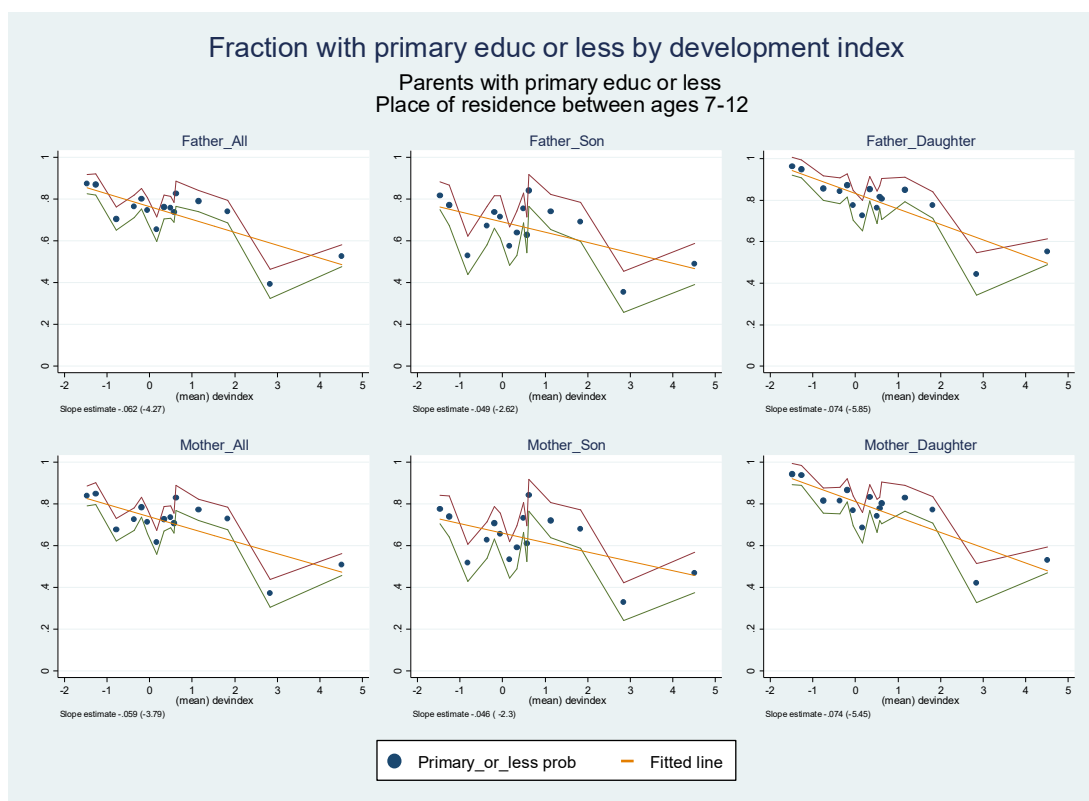
Notes: This figure provides a heat map of intergenerational mobility in Turkey. For each one of our 15 development regions, we estimate the correlation between the average years of schooling of parents and children's years of schooling in that region. Lighter colors represent lower correlation values. All provinces within a development region have the same color. The sample used to compute correlations is restricted to include individuals who are between the ages of 20 and 69.

Figure 4: Bottom-to-Top Probability by Level of Development



Notes: The top left panel presents estimates of the bottom-to-top probability – the probability that a child receives college or university education when the father has less than primary school education – for all 15 development regions. Each blue dot represents a development region. The horizontal axis indicates the mean development score of the region whereas the vertical axis indicates the bottom-to-top probability estimated for that region. The red-green band provides 95% confidence for the corresponding estimate. The orange line provides a linear fit to the estimated bottom-to-top probabilities. The number below each graph refers to the estimated slope of the fitted line. The number in parenthesis is the associated t-stat value. The bottom left panel displays the same probability for children whose mothers have less than primary school education. The top middle and bottom middle panels display bottom-to-top probabilities in all 15 regions for the male subsample while the top right and bottom right panels display regional bottom-to-top probability estimates for the female subsample.

Figure 5: Bottom-to-Bottom Probability by Level of Development



Notes: The top left panel presents estimates of the bottom-to-bottom probability – the probability that a child receives less than primary school education when the father has less than primary school education – for all 15 development regions. Each blue dot represents a development region. The horizontal axis indicates the mean development score of the region whereas the vertical axis indicates the bottom-to-bottom probability estimated for that region. The red-green band provides 95% confidence for the corresponding estimate. The orange line provides a linear fit to the estimated bottom-to-bottom probabilities. The number below each graph refers to the estimated slope of the fitted line. The number in parenthesis is the associated t-stat value. The bottom left panel displays the same probability for children whose mothers have less than primary school education. The top middle and bottom middle panels display bottom-to-bottom probabilities in all 15 regions for the male subsample while the top right and bottom right panels display regional bottom-to-bottom probability estimates for the female subsample.

Table 1: Descriptive Statistics

Panel A - Demographic characteristics					
	Obs	Mean	Std. Dev.	Min	Max
Age	3500	45.4	14.62	25	87
Male	3500	0.49	0.5	0	1
Children's education	3500	6.84	4.69	0	21
Father's education	3420	3.91	3.6	0	21
Mother's education	3405	2.65	3.25	0	21
Panel B - Educational attainment by generation					
	Less than primary	Primary/lower sec sch	HS	College or univ	
Male children	8.7	55.7	19.1	16.5	
Female children	29.2	47.8	11.6	11.3	
Fathers	38.5	53.4	5.8	2.3	
Mothers	56.0	39.7	3.3	1.1	

Notes: This table provides descriptive statistics for the main sample, which includes respondents who are age 25 or older. Panel A reports basic statistics for various demographic and educational characteristics. Panel B of Table 1 provides further information on educational attainment by generation.

Table 2 – Spatial Distribution in the Survey vs. HLFS 2013

NUTS1 Region code	NUTS1 Region name	Fraction in our sample	Fraction in HLFS 2013
TR1	İstanbul	17.21	18.23
TR2	Batı Marmara	6.71	5.12
TR3	Ege	9.94	14.33
TR4	Doğu Marmara	11.49	10.49
TR5	Batı Anadolu	10.44	9.78
TR6	Akdeniz	12.81	12.76
TR7	Orta Anadolu	4.90	4.96
TR8	Batı Karadeniz	5.68	6.31
TR9	Doğu Karadeniz	3.14	3.72
TRA	Kuzeydoğu Anadolu	2.61	2.29
TRB	Ortadoğu Anadolu	5.15	4.00
TRC	Güneydoğu Anadolu	9.91	8.00

Notes: This table provides a comparison of the geographical distribution of our sample across Turkey to the geographical distribution in the Household Labor Force Survey (HLFS) conducted by TUIK. TUIK partitions Turkey into 12 broad regions called NUTS1 regions. Column entitled “Fraction in our sample” reports the fractions of people living in each region in our survey while the column entitled “Fraction in HLFS 2013” reports corresponding fractions in HLFS 2013.

Table 3: Education Mobility at the National Level

VARIABLES	Regression coefficient			Correlation coefficient		
	(1)	(2)	(3)	(4)	(5)	(6)
Father's education	0.734*** [0.018]		0.493*** [0.026]	0.564*** [0.014]		0.379*** [0.020]
Mother's education		0.767*** [0.021]	0.364*** [0.029]		0.532*** [0.015]	0.252*** [0.020]
Observations	3,420	3,405	3,388	3,420	3,405	3,388
R-squared	0.319	0.283	0.348	0.319	0.283	0.348

Notes: Column (1) reports the regression coefficient associated with fathers' years of schooling in the regression where we regress children's years of schooling on fathers' years of schooling using our main sample. Column (2) reports results for the regression where mothers' years of schooling is the independent variable. Column (3) reports results for the regression where both parents' years of schooling are included as independent variables. Columns (4)-(6) repeat the exercises in columns (1)-(3) using adjusted years of schooling as the education measure. Standard errors in brackets *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4: Three-generation Education Mobility

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father's education	0.509*** [0.026]		0.488*** [0.032]	0.426*** [0.038]				0.410*** [0.037]
Paternal grandfather's education		0.350*** [0.064]	0.076 [0.056]	0.037 [0.055]				
Mother's education				0.147*** [0.048]	0.465*** [0.037]		0.436*** [0.046]	0.190*** [0.046]
Maternal grandfather's education						0.338*** [0.053]	0.115* [0.058]	0.025 [0.047]
Observations	2,079	1,884	1,871	1,868	2,096	1,926	1,926	1,903
R-squared	0.233	0.057	0.237	0.246	0.147	0.055	0.161	0.261

Notes: Column (1) reports the results of a regression for the subsample in which respondents are male and have at least one child of at least 25 years of age. Father's education refers to years of schooling of the respondent while children years of schooling refers to that of the randomly selected child of at least 25 years of age. Column (2) reports results of the regression where we regress children's years of schooling on paternal grandfathers' years of schooling. Paternal grandfather refers to respondent's father. Column (3) reports results for the regression where both father's and paternal grandparent's years of schooling are included as independent variables. Columns (4)-(6) repeat the exercises in columns (1)-(3) for the subsample in which respondents are female and have at least one child of at least 25 years of age. All standard errors are clustered at the family level since our sample contains different observations from same families. Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1

Table 5: Intergenerational Education Mobility by Gender

VARIABLES	(1)	(2)	(3)
Father's education	0.660*** [0.025]		0.519*** [0.038]
Father's education*Female	0.135*** [0.035]		-0.033 [0.051]
Mother's education		0.638*** [0.028]	0.202*** [0.042]
Mother's education*Female		0.242*** [0.040]	0.295*** [0.057]
Male	2.505*** [0.188]	2.571*** [0.170]	2.586*** [0.184]
Observations	3,420	3,405	3,388
R-squared	0.367	0.332	0.399

Notes: Column (1) reports results of the regression equation $S_{1i} = \alpha + \beta_0 S_{0i} + \beta_1 S_{0i} * G_i + \beta_2 E_i + e_i$, where S_1 and S_0 represent individual's and father's years of schooling, G is a dummy variable that takes value 1 if the individual is female and 0 otherwise, and E is a dummy variable that takes the value 1 if the individual is male and 0 if female. Column (2) conducts the same regression for the case in which parental input is mother's years of schooling. Column (3) reports the results of the same exercise this time including both father's and mother's years of schooling as parental inputs. Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1.

Table 6 – Relationship between Mobility and Development, Parametric Estimates

VARIABLES	Father-all (1)	Father-son (2)	Father-daughter (3)	Mother-all (4)	Mother-son (5)	Mother-daughter (6)
devindex	0.485*** [0.121]	0.225 [0.169]	0.781*** [0.095]	0.451*** [0.103]	0.186 [0.153]	0.750*** [0.108]
Parent's education	0.769*** [0.023]	0.683*** [0.028]	0.833*** [0.034]	0.825*** [0.029]	0.662*** [0.029]	0.961*** [0.043]
Parent's education* devindex	-0.036*** [0.011]	-0.015 [0.014]	-0.056*** [0.013]	-0.051*** [0.012]	-0.009 [0.018]	-0.091*** [0.016]
Constant	4.033*** [0.154]	5.426*** [0.208]	2.709*** [0.148]	4.901*** [0.145]	6.339*** [0.201]	3.552*** [0.162]
Observations	4,154	1,854	2,300	4,144	1,846	2,298
R-squared	0.368	0.340	0.423	0.322	0.262	0.403

Notes: This table reports the results of the regression equation $S_{1i} = \alpha + \beta_0 S_{0i} + \beta_1 (S_{0i} * DevInd_i) + \gamma DevInd_i + \varepsilon_i$

where $DevInd_i$ is the development index score for individual i 's place of residence between ages of 7 and 12. The first column named Father-all refers to the regression specification in which parental education variable is fathers' years of schooling and the sample includes all observations in our main sample. The next five columns are defined similarly. Standard errors are clustered at the province level. Standard errors in brackets *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7 – Relationship between Mobility and Development, Parametric Estimates

Different Definitions of Area of Residence

VARIABLES	Father-daughter				Mother-daughter			
	Current res.	13-18	7-12	0-6	Current res.	13-18	7-12	0-6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Devindex	0.364*** [0.129]	0.661*** [0.100]	0.781*** [0.095]	0.839*** [0.101]	0.337** [0.128]	0.651*** [0.094]	0.750*** [0.108]	0.756*** [0.120]
Parent's education	0.877*** [0.038]	0.827*** [0.035]	0.833*** [0.034]	0.828*** [0.034]	1.005*** [0.054]	0.963*** [0.045]	0.961*** [0.043]	0.945*** [0.042]
Parent's education*devindex	-0.029** [0.015]	-0.033** [0.016]	-0.056*** [0.013]	-0.058*** [0.015]	-0.051*** [0.018]	-0.070*** [0.019]	-0.091*** [0.016]	-0.084*** [0.019]
Constant	2.518*** [0.197]	2.628*** [0.157]	2.709*** [0.148]	2.743*** [0.151]	3.416*** [0.229]	3.455*** [0.165]	3.552*** [0.162]	3.602*** [0.163]
Observations	2,300	2,297	2,300	2,291	2,298	2,295	2,298	2,289
R-squared	0.404	0.422	0.423	0.423	0.384	0.404	0.403	0.400

Notes: Column (1) reports the estimates of the regression equation $S_{1i} = \alpha + \beta_0 S_{0i} + \beta_1 (S_{0i} * DevInd_i) + \gamma DevInd_i + \varepsilon_i$

for the female subsample where parent's education refers to father's years of schooling and $DevInd_i$ is the development index score for individual i 's current place of residence. Columns (2)-(4) report estimates of the same equation taking $DevInd_i$ to be the development index score for individual i 's place of residence during the ages of 13-18, 7-12 and 0-6. Columns (5)-(8) repeat the exercises in columns (1)-(4) taking mother's years of schooling as parental education input. Standard errors are clustered at the province level. Standard errors in brackets ***

p<0.01, ** p<0.05, * p<0.1.

Table 8 – Relationship between Mobility and Development, Parametric Estimates

Different Definitions of Area of Residence among Migrants

VARIABLES	Father-daughter				Mother-daughter			
	Current res.	13-18	7-12	0-6	Current res.	13-18	7-12	0-6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Devindex	-0.014 [0.130]	0.526** [0.206]	1.134*** [0.375]	1.367*** [0.283]	-0.073 [0.135]	0.510** [0.208]	1.085*** [0.377]	1.299*** [0.299]
Parent's education	0.799*** [0.072]	0.742*** [0.055]	0.733*** [0.054]	0.721*** [0.053]	0.885*** [0.080]	0.863*** [0.057]	0.843*** [0.060]	0.831*** [0.058]
Parent's education*devindex	-0.007 [0.026]	0.010 [0.028]	-0.070** [0.028]	-0.096*** [0.032]	0.004 [0.026]	-0.009 [0.039]	-0.096*** [0.036]	-0.126*** [0.032]
Constant	3.167*** [0.482]	2.988*** [0.282]	3.361*** [0.280]	3.458*** [0.255]	4.137*** [0.484]	3.789*** [0.253]	4.162*** [0.254]	4.243*** [0.234]
Observations	710	708	710	710	707	705	707	707
R-squared	0.346	0.377	0.387	0.396	0.329	0.359	0.368	0.378

Notes: This table reports the same regressions reported in Table 7 for the migrant subsample where a person is defined as migrant if her current area of residence is different from her area of residence during the ages of 0-6 and 7-12. Standard errors are clustered at the province level. Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1.

Table 9 – Relationship between Mobility and Development, Parametric Estimates

Full Set of Area of Residence Controls

	Father-daughter	Mother-daughter
VARIABLES	(1)	(2)
Devindex_current residence	0.081 [0.080]	0.056 [0.090]
Devindex_13-18	0.180 [0.220]	0.188 [0.242]
Devindex_7-12	0.269 [0.213]	0.458* [0.262]
Devindex_0-6	0.372*** [0.103]	0.110 [0.129]
Parent's education	0.844*** [0.039]	0.984*** [0.051]
Parent's educ*devindex_current residence	-0.029* [0.017]	-0.035* [0.020]
Parent's educ*devindex_13-18	0.048 [0.034]	0.042 [0.047]
Parent's educ*devindex_7-12	-0.070** [0.029]	-0.125*** [0.045]
Parent's educ*devindex_0-6	-0.015 [0.018]	0.020 [0.022]
Constant	2.605*** [0.172]	3.448*** [0.190]
Observations	2,289	2,287
R-squared	0.430	0.410

Notes: The table reports estimation results of $S_{1i} = \alpha + \beta_0 S_{0i} + \sum_{j=1}^4 \{\beta_{j,1}(S_{0i} * DevInd_{j,i}) + \gamma_j DevInd_{j,i}\} + \varepsilon_i$ where $j = 1, 2, 3, 4$ refer to the place of residence of children during ages 0-6, 7-12, 13-18, and current place of residence, respectively, and $DevInd_{j,i}$ is the development index score of individual i 's place of residence in period j . The first column named Father-daughter refers to the regression specification in which we restrict our sample to daughters and parental education variable is fathers' years of schooling. The mother-daughter column is defined similarly. Standard errors are clustered at the province level. Standard errors in brackets *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10 – Regional Characteristics and Intergenerational Mobility in Education

Panel A: Mother-Daughter															
	School avail/quality			Social capital			Gender culture			Childhood home environment				Educ inequality	
VARIABLES	(1) Baseline	(2) HS avail	(3) St/teach	(4) Univ/pri	(5) Avg sch	(6) Rel sch	(7) Cho spou	(8) App fem	(9) Fem prior	(10) Encycl	(11) Music ins	(12) Theatre	(13) Sibship	(14) Psch Gini	(15) Sch Gini
Mother's education	0.951*** [0.040]	0.937*** [0.040]	0.907*** [0.037]	0.912*** [0.039]	0.899*** [0.039]	0.907*** [0.036]	0.926*** [0.040]	0.905*** [0.044]	0.946*** [0.045]	0.899*** [0.039]	0.943*** [0.039]	0.931*** [0.035]	0.903*** [0.039]	0.889*** [0.038]	0.906*** [0.039]
Variable		1.137*** [0.178]	-0.880*** [0.175]	1.215*** [0.203]	1.256*** [0.117]	1.400*** [0.195]	0.719*** [0.246]	0.915*** [0.167]	0.182 [0.553]	1.295*** [0.148]	0.800*** [0.187]	1.254*** [0.153]	-1.093*** [0.204]	-1.131*** [0.171]	-1.100*** [0.161]
Mother's education*variable		-0.125*** [0.030]	0.064 [0.046]	-0.133*** [0.025]	-0.131*** [0.026]	-0.137*** [0.039]	-0.083** [0.033]	-0.058 [0.062]	-0.035 [0.107]	-0.130*** [0.023]	-0.078** [0.037]	-0.160*** [0.026]	0.126*** [0.042]	0.098** [0.040]	0.125*** [0.036]
Constant	3.759*** [0.258]	3.783*** [0.178]	3.934*** [0.228]	3.935*** [0.161]	3.973*** [0.145]	3.920*** [0.157]	3.892*** [0.226]	3.906*** [0.214]	3.791*** [0.292]	3.951*** [0.147]	3.726*** [0.199]	3.853*** [0.152]	3.999*** [0.209]	3.989*** [0.183]	3.963*** [0.180]
Observations	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298	2,298
R-squared	0.375	0.397	0.395	0.402	0.409	0.405	0.386	0.397	0.376	0.410	0.388	0.406	0.401	0.404	0.403

Panel B: Father-Daughter															
	School avail/quality			Social capital			Gender culture			Childhood home environment				Educ inequality	
VARIABLES	(1) Baseline	(2) HS avail	(3) St/teach	(4) Univ/pri	(5) Avg sch	(6) Rel sch	(7) Cho spou	(8) App fem	(9) Fem prior	(10) Encycl	(11) Music ins	(12) Theatre	(13) Sibship	(14) Psch Gini	(15) Sch Gini
Father's education	0.861*** [0.028]	0.832*** [0.033]	0.811*** [0.028]	0.809*** [0.029]	0.792*** [0.031]	0.806*** [0.029]	0.826*** [0.035]	0.802*** [0.033]	0.871*** [0.032]	0.794*** [0.030]	0.846*** [0.033]	0.816*** [0.030]	0.794*** [0.033]	0.786*** [0.033]	0.793*** [0.034]
Variable		0.976*** [0.251]	-0.832*** [0.166]	1.304*** [0.201]	1.235*** [0.102]	1.380*** [0.227]	0.794*** [0.261]	0.784*** [0.162]	0.093 [0.370]	1.246*** [0.109]	0.685** [0.272]	1.154*** [0.125]	-1.020*** [0.203]	-1.046*** [0.187]	-1.013*** [0.171]
Father's education*variable		-0.065* [0.034]	-0.004 [0.029]	-0.101*** [0.016]	-0.067** [0.027]	-0.101*** [0.027]	-0.042 [0.037]	0.037 [0.035]	0.090 [0.077]	-0.076*** [0.023]	-0.044 [0.038]	-0.083*** [0.024]	0.028 [0.028]	0.026 [0.026]	0.014 [0.035]
Constant	2.872*** [0.255]	2.950*** [0.188]	3.083*** [0.240]	3.116*** [0.156]	3.155*** [0.150]	3.129*** [0.160]	3.026*** [0.229]	3.057*** [0.231]	2.870*** [0.272]	3.147*** [0.145]	2.873*** [0.199]	3.025*** [0.149]	3.160*** [0.226]	3.171*** [0.206]	3.116*** [0.200]
Observations	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300	2,300
R-squared	0.394	0.411	0.416	0.422	0.428	0.420	0.409	0.416	0.398	0.427	0.404	0.420	0.421	0.422	0.424

Notes: Each column in the table above refers to a separate regression analysis in which we regress children's years of schooling on parent's years of schooling, a regional variable of

interest, and a term that interacts parent's years of schooling and the regional variable of interest. In Panel A, parent refers to mother while in Panel B it refers to father. Column (1) considers the baseline intergenerational regression without any additional variable. In columns (2)-(15), the variables of interest are regional high school availability (2), student-to-teacher ratio (3), ratio of university graduates to primary school graduates in the region (4), average years of schooling in the region (5), average years of schooling of relative with highest education for each region (6), fraction who choose their spouses themselves (7), fraction who approve female employment (8), fraction who value daughter's education at least as much as son's education (9), fraction who had encyclopedia at home growing up (10), fraction who had a musical instrument at home growing up (11), average number of cinema or theatre attendance growing up (12), average number of siblings (13), Gini coefficient of parental years of schooling (14), Gini coefficient of years of schooling (15). Standard errors in brackets *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Appendix

Constructing Development Regions

In this Appendix, we explain how we group provinces in Turkey into 15 development regions. To do so, we first rank all the individuals in our sample according to the development index value of the province they resided between the ages of 7 and 12. We assign each individual a development percentile rank and then divide the sample into 15 development regions based on their percentile ranks. Individuals who grow up in the same province and, hence, share the same development index value, are randomly assigned a different percentile rank according to this procedure. To prevent this randomness, for individuals who grew up in the same province, the percentile rank is replaced by the mean of the percentile ranks of all the individuals in that region.¹ This construction ensures that individuals in a region grew up in provinces with similar socioeconomic development levels. We can compute a measure of the level of development for each region by taking an average of the development index values of individuals in a given region. Table A1 provides the resulting list of provinces within each development region while Table A2 presents several characteristics of the 15 development regions.

¹This procedure is also adopted by Chetty et al. (2014).

Table A1 - List of Provinces in Each Development Region

Region	Provinces	Region	Provinces	Region	Provinces	Region	Provinces
1	Ağrı Bitlis Hakkari Muş Siirt Van Şırnak	4	Çankırı Erzurum Giresun Niğde Tokat Tunceli Aksaray Osmaniye	7	Amasya Nevşehir Rize Samsun Trabzon Karaman Düzce	11	Çanakkale Edirne Kırklareli Yalova
2	Bingöl Mardin Şanlıurfa Batman Ardahan Iğdır	5	Çorum Hatay Kastamonu Sinop Sivas Bartın	8	Bilecik Burdur Gaziantep Mersin Uşak Zonguldak Karabük	12	Bolu Bursa Denizli Eskişehir Muğla Tekirdağ
3	Adıyaman Diyarbakır Gümüşhane Kars Kahramanmaraş Ordu Yozgat Bayburt Kilis	6	Afyon Artvin Elazığ Erzincan Kırşehir Kütahya Malatya Kırıkkale	9	Balıkesir Isparta Konya Manisa	13	Antalya İzmir Kocaeli
				10	Adana Aydın Kayseri Sakarya	14	Ankara
						15	İstanbul

Notes: This table provides the list of provinces within each development region. See the discussion in this Appendix for the procedure used to construct 15 regions. By construction, each region consists roughly 1/15 of all the individuals in the sample and the provinces within a region have similar socioeconomic development index values.

Table A2 - Characteristics of Development Regions

Development region	Mean development index	GDP per capita (\$)	Average years of schooling	Grade 8 test score	Student teacher ratio	Secondary school present age7_12	High school present age7_12	Fraction who choose spouse	Parental years of sch. Gini	Highest years of sch. among relatives	N
1	-1.47	802.53	3.78	213.76	31.35	0.62	0.52	0.54	0.70	9.17	245
2	-1.27	983.67	4.16	213.05	32.34	0.65	0.61	0.37	0.61	9.37	189
3	-0.78	1139.75	5.03	256.30	26.96	0.73	0.65	0.39	0.52	9.87	309
4	-0.36	1279.00	5.65	277.55	18.86	0.71	0.61	0.35	0.44	9.92	289
5	-0.19	1608.78	5.34	285.52	21.50	0.60	0.51	0.36	0.44	10.16	304
6	-0.06	1622.25	6.06	296.07	21.06	0.68	0.57	0.43	0.50	10.23	251
7	0.17	1651.15	5.72	293.11	19.23	0.71	0.61	0.49	0.36	10.95	309
8	0.33	2016.97	5.69	272.55	24.76	0.76	0.71	0.46	0.49	9.93	244
9	0.50	1895.70	5.80	298.61	18.35	0.75	0.54	0.48	0.37	10.10	293
10	0.57	2145.49	6.22	289.45	24.45	0.84	0.79	0.45	0.44	9.83	380
11	0.61	2895.23	6.08	317.63	19.83	0.49	0.33	0.49	0.42	8.02	155
12	1.15	2822.91	6.04	312.68	20.62	0.81	0.71	0.60	0.34	10.95	290
13	1.83	3628.34	6.46	306.35	19.37	0.73	0.66	0.63	0.38	10.52	304
14	2.84	2752.00	7.81	321.42	17.73	0.98	0.95	0.68	0.28	12.97	257
15	4.52	3063.00	8.58	298.71	21.26	0.99	0.98	0.77	0.31	11.91	430

Notes: This table presents several characteristics of the 15 development regions. Development index refers to the Socio-economic Development Index for 2011 published by the Turkish Ministry of Development where larger values of this index indicate more developed regions. GDP, years of schooling and teacher student ratio figures refer to year 2000, grade 8 test score values refer to year 2007. The information on GDP and teacher student ratio is based on province level statistics published by the Turkish Statistical Institute and national education statistics published by Ministry of Education, respectively. Grade 8 test scores are based on the nationwide SBS exam (Seviye Belirleme Sınavı) conducted in 2007 that assesses competency in mathematics, science, social sciences, and language fields. The remaining information in the table is based on the survey conducted for this paper. The last column reports the weighted number of observations in data used in this paper for each region.

Table A3 - Intergenerational Correlations by Development Regions

Development region	Correlation coefficient	Standard error
1	0.770	0.076
2	0.607	0.083
3	0.616	0.051
4	0.636	0.047
5	0.533	0.047
6	0.644	0.049
7	0.622	0.050
8	0.554	0.050
9	0.526	0.046
10	0.632	0.044
11	0.349	0.081
12	0.470	0.045
13	0.596	0.042
14	0.416	0.047
15	0.474	0.038

Notes: The results are based on separate estimations of equation (2) for each region. Following Hertz (2007), average years of schooling of parents is set as the parental education variable and the sample is restricted to individuals who are between the ages of 20 and 69. For each region correlation coefficient refers to that between (adjusted) average parental education and (adjusted) child's education.

Additional Estimations for Robustness

Table A4 - Relationship between Mobility and Development, Parametric Estimates, Subindices

Panel A – Daughters

VARIABLES	Father-daughter				Mother-daughter			
	Overall	Manufacturing	Education	Health	Overall	Manufacturing	Education	Health
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parent educ	0.847*** [0.034]	0.867*** [0.032]	0.799*** [0.036]	0.824*** [0.035]	0.974*** [0.043]	0.996*** [0.041]	0.937*** [0.044]	0.942*** [0.047]
devindex	0.689*** [0.099]	0.572*** [0.112]	0.867*** [0.185]	0.889*** [0.196]	0.646*** [0.103]	0.530*** [0.113]	0.902*** [0.177]	0.845*** [0.182]
Parent educ*devindex	-0.052*** [0.012]	-0.052*** [0.012]	-0.021 [0.025]	-0.055** [0.021]	-0.081*** [0.015]	-0.081*** [0.016]	-0.102*** [0.034]	-0.086*** [0.028]
Constant	2.640*** [0.163]	2.625*** [0.177]	2.888*** [0.188]	2.748*** [0.163]	3.505*** [0.180]	3.514*** [0.196]	3.717*** [0.174]	3.580*** [0.172]
Observations	2,300	2,300	2,300	2,300	2,298	2,298	2,298	2,298
R-squared	0.418	0.410	0.423	0.422	0.397	0.390	0.401	0.401

Panel B – Sons

VARIABLES	Father-son				Mother-son			
	Overall	Manufacturing	Education	Health	Overall	Manufacturing	Education	Health
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parent educ	0.684*** [0.029]	0.692*** [0.029]	0.686*** [0.029]	0.678*** [0.029]	0.663*** [0.030]	0.673*** [0.030]	0.666*** [0.032]	0.655*** [0.031]
devindex	0.203 [0.175]	0.124 [0.140]	0.156 [0.200]	0.171 [0.251]	0.163 [0.157]	0.106 [0.150]	0.077 [0.196]	0.113 [0.230]
Parent educ*devindex	-0.012 [0.014]	-0.011 [0.013]	-0.013 [0.020]	-0.003 [0.018]	-0.007 [0.018]	-0.007 [0.019]	0.000 [0.024]	0.005 [0.021]
Constant	5.405*** [0.206]	5.413*** [0.208]	5.458*** [0.217]	5.430*** [0.204]	6.325*** [0.201]	6.332*** [0.204]	6.369*** [0.211]	6.354*** [0.198]
Observations	1,854	1,854	1,854	1,854	1,846	1,846	1,846	1,846
R-squared	0.340	0.338	0.338	0.339	0.261	0.260	0.259	0.261

Notes: Each column in this table reports the results of the regression equation

$$S_{it} = \alpha + \beta_0 S_{0i} + \beta_1 (S_{0i} * DevInd_i) + \gamma DevInd_i + \varepsilon_i, \text{ where } DevInd_i \text{ is the development index score for individual } i\text{'s}$$

place of residence between ages of 7 and 12. Panels A and B analyzes this relationship for the female and male subsamples, respectively. In Column (1), development index score is calculated using SEGE-2003 development index. In Columns (2)-(4), development index score is calculated using manufacturing, education, and health subindices. Columns (5)-(8) are defined similarly. Standard errors are clustered at the province level. Standard errors in brackets *** p<0.01, ** p<0.05, * p<0.1.

Table A5 - Regional Characteristics and Intergenerational Mobility in Education: Sons

Panel A: Mother-Son															
		School avail/quality		Social capital			Gender culture			Childhood home environment				Educ inequality	
VARIABLES	(1) Baseline	(2) HS avail	(3) St/teach	(4) Univ/pri	(5) Avg sch	(6) Rel sch	(7) Cho spou	(8) App fem	(9) Fem prior	(10) Encycl	(11) Music ins	(12) Theatre	(13) Sibship	(14) Psch Gini	(15) Sch Gini
Mother's education	0.675*** [0.027]	0.637*** [0.028]	0.681*** [0.026]	0.645*** [0.026]	0.650*** [0.027]	0.632*** [0.028]	0.664*** [0.027]	0.670*** [0.026]	0.671*** [0.027]	0.641*** [0.026]	0.669*** [0.028]	0.649*** [0.027]	0.676*** [0.026]	0.666*** [0.027]	0.681*** [0.027]
Variable		0.751*** [0.181]	-0.045 [0.192]	0.585*** [0.200]	0.347 [0.209]	0.675*** [0.177]	0.180 [0.265]	0.146 [0.147]	-0.099 [0.137]	0.517*** [0.193]	0.627*** [0.131]	0.454** [0.180]	-0.104 [0.195]	-0.192 [0.193]	-0.104 [0.183]
Mother's education*variable		-0.030 [0.023]	0.029 [0.036]	-0.034** [0.017]	-0.013 [0.026]	-0.021 [0.022]	0.004 [0.032]	-0.027 [0.035]	-0.081*** [0.018]	-0.024 [0.025]	-0.052*** [0.019]	-0.021 [0.029]	0.026 [0.031]	0.021 [0.030]	0.043 [0.028]
Constant	6.364*** [0.207]	6.523*** [0.195]	6.372*** [0.218]	6.507*** [0.195]	6.469*** [0.216]	6.516*** [0.192]	6.395*** [0.214]	6.409*** [0.221]	6.369*** [0.203]	6.507*** [0.205]	6.390*** [0.191]	6.469*** [0.204]	6.389*** [0.227]	6.417*** [0.225]	6.386*** [0.221]
Observations	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846	1,846
R-squared	0.259	0.276	0.259	0.267	0.263	0.273	0.261	0.260	0.264	0.267	0.268	0.265	0.260	0.260	0.260

Panel B: Father-Son															
		School avail/quality		Social capital			Gender culture			Childhood home environment				Educ inequality	
VARIABLES	(1) Baseline	(2) HS avail	(3) St/teach	(4) Univ/pri	(5) Avg sch	(6) Rel sch	(7) Cho spou	(8) App fem	(9) Fem prior	(10) Encycl	(11) Music ins	(12) Theatre	(13) Sibship	(14) Psch Gini	(15) Sch Gini
Father's education	0.691*** [0.025]	0.659*** [0.028]	0.694*** [0.026]	0.664*** [0.026]	0.670*** [0.026]	0.657*** [0.028]	0.682*** [0.025]	0.685*** [0.024]	0.681*** [0.025]	0.663*** [0.026]	0.682*** [0.026]	0.671*** [0.026]	0.687*** [0.025]	0.681*** [0.026]	0.689*** [0.025]
Variable		0.620** [0.251]	-0.065 [0.200]	0.599** [0.300]	0.342 [0.254]	0.592** [0.259]	0.280 [0.274]	0.143 [0.177]	0.066 [0.138]	0.461* [0.257]	0.364** [0.145]	0.392* [0.209]	-0.161 [0.208]	-0.204 [0.219]	-0.131 [0.189]
Father's education*variable		-0.016 [0.027]	0.028 [0.031]	-0.034 [0.022]	-0.014 [0.025]	-0.019 [0.027]	-0.010 [0.024]	-0.020 [0.031]	-0.071** [0.030]	-0.021 [0.025]	-0.016 [0.022]	-0.021 [0.025]	0.028 [0.028]	0.020 [0.028]	0.030 [0.029]
Constant	5.447*** [0.209]	5.618*** [0.221]	5.462*** [0.226]	5.612*** [0.224]	5.564*** [0.234]	5.618*** [0.217]	5.490*** [0.216]	5.500*** [0.225]	5.475*** [0.203]	5.602*** [0.229]	5.486*** [0.200]	5.559*** [0.222]	5.495*** [0.230]	5.515*** [0.234]	5.486*** [0.224]
Observations	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854	1,854
R-squared	0.337	0.349	0.338	0.344	0.341	0.347	0.339	0.338	0.339	0.343	0.340	0.341	0.338	0.338	0.338

Notes: This table reports the estimation results reported in Table 10 for the male subsample. Each column in the table above refers to a separate regression analysis in which we

regress children's years of schooling on parent's years of schooling, a regional variable of interest, and a term that interacts parent's years of schooling and the regional variable of interest. In Panel A, parent refers to mother while in Panel B it refers to father. Column (1) considers the baseline intergenerational regression without any additional variable. In columns (2)-(15), the variables of interest are regional high school availability (2), student-to-teacher ratio (3), ratio of university graduates to primary school graduates in the region (4), average years of schooling in the region (5), average years of schooling of relative with highest education for each region (6), fraction who choose their spouses themselves (7), fraction who approve female employment (8), fraction who value daughter's education at least as much as son's education (9), fraction who had encyclopedia at home growing up (10), fraction who had a musical instrument at home growing up (11), average number of cinema or theatre attendance growing up (12), average number of siblings (13), Gini coefficient of parental years of schooling (14), Gini coefficient of years of schooling (15). Standard errors in brackets *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.