Intergenerational Income Mobility in Turkey

Erkan DUMAN[*]

ABSTRACT

This study presents a thorough review of the literature on intergenerational income mobility with a focus on the issues regarding consistent estimation of mobility correlations. In the light of the discussions, the intergenerational income association in Turkey is analyzed using OLS and IV estimation strategies. The sample is extracted from Household Budget Surveys and includes father-son and father-daughter pairs with either non-zero wages or non-zero earnings living in the same household in any of the years between 2003 and 2011. OLS and IV estimation strategies are shown to bound the true population mobility correlation from below and above; hence, resulting in a consistent interval estimate of intergenerational mobility. The true mobility correlations in the population are estimated to be in the range of (0.10, 0.51) for sons and (0.17, 1.00) for daughters. The large mobility correlations may call for government intervention to break the harmful income link across generations.

Türkiye'deki Kuşaklar Arası Gelir Hareketliliği

ÖZ

Bu çalışma, hareketlilik korelasyonlarının tutarlı tahminine ilişkin konulara odaklanarak kuşaklar arası gelir hareketliliği literatürünün kapsamlı bir incelemesini sunmaktadır. Tartışmalar ışığında, Türkiye'deki kuşaklar arası gelir hareketliliği "En Küçük Kareler" ve "Araç Değişken" tahmin stratejileri kullanılarak analiz edilmiştir. Örneklem Hanehalkı Bütçe Anketlerinden alınmıştır ve 2003 ile 2011 arasındaki yılların herhangi birinde aynı hanede yaşayan pozitif bir ücret veya kazanç belirten baba-oğul ve baba-kız çiftlerini içerir. En Küçük Kareler ve Araç Değişken tahmin stratejilerinin gerçek popülasyon hareketliliği korelasyonunu aşağıdan ve yukarıdan bağladığı gösterilmiştir; bu nedenle, kuşaklar arası hareketliliğin tutarlı bir aralık tahmini ile sonuçlanır. Popülasyondaki gerçek hareketlilik korelasyonlarının erkek çocuklar için (0,10, 0,51) ve kız çocuklar için (0,17, 1,00) aralığında olduğu tahmin edilmektedir. Bulunan büyük hareketlilik korelasyonları, iki nesil arasında yer alanzararlı gelir bağını kırmak için devlet müdahalesini gerektirebilir.

[*] World Bank, erduman@alumni.sabanciuniv.edu, ORCID: 0000-0002-6585-6665

1.INTRODUCTION

Mayer (2010) suggests that children from rich families have higher education, live a more prosperous life and be more successful at later stages of their lives compared to children from poor families. The problem is whether this relation is due to rich families which have more resources to invest in their children actually investing more compared to poor families. If so, government may redistribute income among poor families or invest directly in human capital of children from poor families to fight against income inequality and poverty in the next generation (Shea, 2000). On the other hand, it is wondered whether increasing the income of poor families result in higher investment levels in their children's human capital. Poor families may be unwilling to invest in their children even if they have adequate resources due to their parental characteristics associated with their poverty—culture, neighborhood characteristics and probably parents' low valuation of education. In this case, income redistribution policy would not be a remedy. So, it is necessary to understand whether the intergenerational income mobility relation is causal and what drives persistence in intergenerational wealth.

This study presents an in-depth survey of the recent methodological advancements in consistent estimation of intergenerational income mobility. This study also contributes to the scarce literature on intergenerational income mobility in Turkey by trying to consistently estimate intergenerational mobility coefficients. The econometric techniques applied in this paper allows to present a range for mobility coefficients that includes the true population parameter. The results suggest a high persistence in income across generations.

The next section overviews the literature and discusses the issues of and proposed solutions to consistent estimation of intergenerational income mobility coefficients. The rest of the paper follows as: the dataset used to estimate intergenerational income mobility coefficients in Turkey is described and some summary statistics about the sample are presented in section 3. Section 4 briefly summarizes the estimation strategy. Results are presented in section 5 and lastly section 6 concludes.

2. LITERATURE REVIEW AND ESTIMATION ISSUES

Dearden et al. (1997) and Solon (1992) try to estimate the intergenerational income correlations. They want to see to what extent a child's position in the income distribution can be explained by his parents' positions in their own income distribution. They do not try to account for unobserved ability which may bias the estimates: High able parents may earn higher income compared to low able parents and high able parents are more likely to have high able children who may earn high income. Even though they do not take care of ability bias, they try to draw a boundary around the true parental income effect by some methods. Dearden et al (1997) estimates the equation:

$$y_i^{child} = a + by_i^{parent} + \varepsilon_i \tag{1}$$

where y_i^{child} is the permanent income of the child, y_i^{parent} is the permanent income of the parent and ε_i is an error term. The data used is National Child Development Survey in Britain for individuals born in a specific week in March 1958. There is detailed information on children's educational attainment and earnings; however, paternal income is observed only for 1974. Dearden et al. (1997) argues that having only a cross sectional observation for paternal income will result in error in measuring permanent income and this will cause errors in variables problem and biases all coefficients. Trying to estimate the intergenerational income correlations without taking care of the measurement error will result in downward biased inconsistent estimates (Solon, 1992; Zimmerman, 1992). Dearden et al. (1997) suggests that the observed

income can be decomposed into three components: an observable time varying component (age, region, etc.), permanent income, and a transitory shock. By regressing observed income on observable time varying determinants and taking the residual, Dearden et al (1997) capture the permanent income with a transitory shock. Still, permanent income is measured with error but at least the measurement error is diminished by the extent of the variation captured by the observable time varying components. To deal with the errors in variables problem, Dearden et al. (1997) uses father's total years of schooling and social class as instruments for paternal income. However, Solon (1992) argues that father's educational attainment is a relevant variable in child's income equation, thus, using father's schooling as an instrument for father's income will yield upward biased estimates.

Dearden et al. (1997) estimates the paternal income coefficient to be 0.240 for boys and 0.315 for girls with OLS. When the residual obtained by subtracting the impact of time varying observable components from observed paternal income is used in place of the permanent income, the estimate decreases to 0.216 for boys. When instrumental variables estimation technique is used to correct for the measurement error, the estimates range from 0.56 to 0.59 for boys and 0.63 to 0.70 for girls. As a consequence, with a range of 0.24 to 0.56 for boys and a range of 0.315 to 0.63 for girls, the impact of paternal income should not be despised.

Solon (1992) is another influential study that tries to estimate the intergenerational income correlations. He estimates the equation:

$$y_{1i} = \rho y_{0i} + \varepsilon_i \qquad (2)$$

where y_{1i} is the permanent income of the child, y_{0i} is the permanent income of the parent and ε_i is an error term. ρ represents the true population correlation between y_{1i} and y_{0i}. Solon (1992) argues that there are two main problems in estimating intergenerational income correlations. First, being unable to observe the long run income and replacing it with the cross-sectional observation results in measurement error which biases the estimates. Second, having a homogenous sample of parents or children also biases the estimates. Homogenous sample of parents (children) is described as a sample of parents (children) who have lowervariance in their long run income compared to the population of parents (children).

Starting with the measurement error problem Solon (1992) assumes that the population variances of permanent income are the same for the generations of parents and children and it is denoted by σ_y^2. Solon (1992) takes the sample of fathers-sons to estimate the intergenerational income correlations. He begins with the assumption that both son's and father's long run income is measured with error. That is, son's long run income is proxied with his short run status in period t: $y_{1it} = y_{1i} + v_{1it}$ (3) and father's long run income is proxied with his short run status in period s: $y_{0is} = y_{0i} + v_{0is}$ (4). v_{1it} and v_{0is} stand for transitory movement around the permanent income and random measurement error for son's and father's income, respectively. σ_{v1}^2 and σ_{v0}^2 represent the population variances in the random measurement errors for son's and father's income, respectively. Assume that random measurement errors are not correlated with permanent income and with each other. That is, v_{1it} and v_{0is} are not correlated with y_{1i} and with y_{0i}, in addition, are not correlated with each other. If we do not observe the permanent income of son and father and use instead the short run proxies (3) and (4) in equation (2) we estimate $\hat{\rho}$ with bias. In this case, the probability limit of $\hat{\rho}$ is:

$$plim\hat{\rho} = \frac{Cov(y_{1it}, y_{0is})}{Var(y_{0is})} = \frac{Cov(\rho y_{0is} + \varepsilon_i - \rho v_{0is} + v_{1it}, y_{0is})}{Var(y_{0is})} \qquad (5)$$

$$= \rho - \frac{\rho Cov(v_{0is}, y_{0is})}{Var(y_{0is})} = \rho - \frac{\rho Cov(v_{0is}, y_{0i} + v_{0is})}{Var(y_{0is})}$$

$$= \rho - \frac{\rho \sigma_{v0}^2}{\sigma_y^2 + \sigma_{v0}^2} = \frac{\rho \sigma_y^2}{\sigma_y^2 + \sigma_{v0}^2} < \rho$$

As can be seen from (5), the actual problem is the measurement error in father's income. If father's permanent income is observed, the measurement error in son's income would not cause inconsistency. The bias depends strongly on the size of the variance of the measurement error in father's income. If the variance of the measurement error in father's income is large, then the bias will be large. If the measurement error in father's income is densely distributed, then the bias will be small. It is suggested that for studies using only one-year data like Behrman and Taubman (1985), errors in variables bias could depress the estimates of ρ by more than 30% (Michael Baker, 1990; Solon et al. 1991; MacDonald and Robinson, 1985). Dearden et al. (1997) suggests that if longitudinal data is available on the income of parents and children, averaging the income for a few years decreases the errors in variables bias considerably but does not eliminate it at all.

Another problem is the homogenous samples of fathers or sons. Solon (1992) begins with an assumption that we can observe the long run income for fathers and sons. If the sample of fathers has a lower long run income variance compared to the population $(s_y^2 < \sigma_y^2)$, then we can consistently estimate ρ (Atkinson et al, 1983; Menchik, 1979). The probability limit of $\hat{\rho}$ is:

$$plim\hat{\rho} = \frac{Cov(y_{1i}, y_{0i})}{Var(y_{0i})} = \frac{Cov(\rho y_{0i} + \varepsilon_i, y_{0i})}{Var(y_{0i})} = \rho \frac{Var(y_{0i})}{Var(y_{0i})} = \rho \frac{s_y^2}{s_y^2} = \rho$$

So, if we observe permanent income of the father, then there will be no problem in estimating the intergenerational income correlation consistently even though the sample of fathers is homogenous. However, if the sample of sons is homogenous, then the estimate of ρ will be inconsistent (and probably downward biased) (Goldberger, 1981; Chung and Goldberger, 1984; Greene, 1990 Ch. 21). If we do not observe the permanent income of the father, then homogenous fathers sample will aggravate the errors in variables bias (Solon, 1992). Now assume that for fathers $s_y^2 < \sigma_y^2$ holds and fathers' permanent income is measured with error so $y_{0is} = y_{0i} + v_{0is}$. The probability limit of $\hat{\rho}$ is:

$$plim\hat{\rho} = \frac{Cov(y_{1i}, y_{0is})}{Var(y_{0is})} = \frac{Cov(\rho y_{0is} + \varepsilon_i - \rho v_{0is}, y_{0is})}{Var(y_{0is})} = \rho - \frac{\rho Cov(v_{0is}, y_{0i} + v_{0is})}{Var(y_{0is})}$$

$$= \rho - \frac{\rho \sigma_{v0}^2}{s_y^2 + \sigma_{v0}^2} = \frac{\rho s_y^2}{s_y^2 + \sigma_{v0}^2} < \rho$$

Since $0 < s_y^2 < \sigma_y^2$ for the sample of fathers we conclude that $s_y^2 \sigma_y^2 + s_y^2 \sigma_{v0}^2 < s_y^2 \sigma_y^2 + \sigma_y^2 \sigma_{v0}^2$, from here it follows that $\frac{s_y^2}{s_y^2 + \sigma_{v0}^2} < \frac{\sigma_y^2}{\sigma_y^2 + \sigma_{v0}^2}$ so the bias is larger when errors in variables problem is accompanied with homogenous fathers sample.

Solon (1992) argues that since individuals have different earnings in different stages of their lives and probably low earnings when they are young, it is necessary to control for the age effect in equations (3) and (4); otherwise, the intergenerational income estimate will capture the age effect. Young fathers are more likely to earn less compared to old fathers since they have less experience in the labor force and young fathers are more likely to have younger children

who earn less compared to children of older fathers. So, some part of the relation between fathers' and sons' income is due to the variation in age. Solon (1992) adds age and age squared to equation (3) and (4). He adds age squared since age variable also captures the effect of experience and it is widely acknowledged that income and experience has concave relationship. Solon (1992) revises equations (3) and (4):

$$y_{1it} = y_{1i} + \alpha_1 + \beta_1 A_{1it} + \gamma_1 A_{1it}^2 + v_{1it} \tag{6}$$

$$y_{0is} = y_{0i} + \alpha_0 + \beta_0 A_{0is} + \gamma_0 A_{0is}^2 + v_{0is} \tag{7}$$

where A's represent the age and A^2's represent the age squared. Putting back equations (6) and (7) into equation (2) gives:

$$y_{1it} = (\alpha_1 - \rho\alpha_0) + \rho y_{0is} + \beta_1 A_{1it} + \gamma_1 A_{1it}^2 - \rho\beta_0 A_{0is} - \rho\gamma_0 A_{0is}^2 + \varepsilon_i + v_{1it} - \rho v_{0is} \tag{8}$$

Solon (1992) in (8) regresses son's observed income on father's observed income and controls for the age and age-squared of both son and father. The errors in variables bias arises due to the correlation between father's observed income y_{0is} and father's transitory income shock v_{0is}. Solon (1992) suggests two ways to solve for errors in variables bias. First, averaging father's income over a few years (T years) reduces the noise variance by a factor of T. This will reduce the bias in $\hat{\rho}$; however, does not eliminate it at all. Second way is to use father's education as an instrumental variable for the single year observation of father's income. It is arguable whether father's education is excludable from child's earning equation. Solon (1992) argues that under some plausible assumptions using father's education as an instrument results in upward biased estimates for the intergenerational income correlations. Thus, OLS yields downward biased estimates for the intergenerational income correlations due to the errors in variables bias and IV estimation yields upward biased estimates. Solon (1992) argues that the true intergenerational income correlationis between the OLS and IV results. Solon (1992) finds in USA an estimate of 0.386 for intergenerational income correlation with OLS and 0.526 with IV. Thus, he concludes that the intergenerational income correlation is around 0.4. It is important to remember that he tries to estimate the intergenerational income correlations because if he tries to estimate causal intergenerational income effects, then omitting father's education in child's earning equation where father's education is relevant result in (most likely) upward biased estimate for the causal impact unless father's education is negatively correlated with child's income.Omitted variables bias and errors in measurement bias may work in opposite directions and make it difficult to argue that there is a downward bias in the estimated causal intergenerational income effect.

The instrumental variables estimation method implemented in Solon (1992) is explained next to understand under what conditions the IV estimates of intergenerational income correlations are upward biased.

Solon (1992) suggests that in contrary to equation (2) father's education is relevant to son's income. So, the son's earning equation becomes: $y_{1i} = \beta_1 y_{0i} + \beta_2 E_i + u_i$ where E_i stands for the father's total years of schooling. If this equation represents the true relation between son's income and father's income, then estimating equation (2) results in committing omitted variable bias since father's education is not included in (2). The relation between β_1, β_2 ve $\rho$is:

$$\rho = \beta_1 + \frac{\beta_2 Cov(E_i, y_{0i})}{\sigma_y^2}$$

We can rewrite it as

$$\rho = \beta_1 + \beta_2 \sigma_E \left\{ \frac{Cov(E_i, y_{0i})}{\sigma_y \sigma_E} \right\} \frac{1}{\sigma_y} = \beta_1 + \frac{\beta_2 \lambda \sigma_E}{\sigma_y}$$

where λ is the correlation between father's permanent income and father's education.

Solon (1992) states that measurement error in father's income results in downward biased intergenerational income correlation estimates. To correct for this bias using father's education as an instrument for father's observed income and assuming that E_i is uncorrelated with v_{0is} and v_{1it} results in:

$$plim \hat{\rho}_{IV} = \frac{Cov(E_i, y_{1it})}{Cov(E_i, y_{0is})}$$

$$= \frac{Cov(E_i, \beta_1 y_{0is} + \beta_2 E_i + u_i + v_{1it} - \beta_1 v_{0is})}{Cov(E_i, y_{0i} + v_{0is})}$$

$$= Cov(E_i, \beta_1 y_{0is} + \beta_2 E_i + u_i + v_{1it} - \beta_1 v_{0is}) / (\lambda \sigma_E \sigma_y)$$

$$= \beta_1 + \beta_2 \sigma_E^2 / (\lambda \sigma_E \sigma_y)$$

$$= \beta_1 + \frac{\beta_2 \lambda \sigma_E}{\sigma_y} + \beta_2 \left[\left(\frac{\sigma_E}{\lambda \sigma_y} \right) - \left(\frac{\lambda \sigma_E}{\sigma_y} \right) \right]$$

$$= \rho + \beta_2 \sigma_E (1 - \lambda^2) / \lambda \sigma_y$$

So, IV estimate is consistent only if $\beta_2 = 0$ or $\lambda = 1$. That is, if father's education is excludable from child's earning equation or father's education and his permanent income is perfectly correlated. Sewell and Hauser (1975) and Corcoran et al. (1992) find that once parental income is averaged over several years and is controlled for, the estimated effect of parental education on son's earnings is indistinguishable from zero. So, these arguments may be used as a justification for using father's education as an IV for his income and in this case, there will be no bias in the intergenerational income estimate. If $\beta_2 > 0$ then there is an upward bias in the estimate of intergenerational income mobility and this is the plausible expectation for β_2 since children who have high educated parents are expected to earn more compared to children who have low educated parents. In addition, E_i may also be measured with error which does not affect the estimates (Solon, 1992).

3. DATA AND DESCRIPTIVE STATISTICS

This paper uses data from cross-sectional household budget surveys, "Hanehalkı Bütçe Anketi" conducted by Turkey's national statistical agency (Türkiye İstatistik Kurumu). Nineconsecutive cross-sections from household budget surveys are pooled covering the years from 2003 to 2011 in order to increase variation. Each survey is representative at urban, rural and national levels. The surveys contain information on demographic characteristics including the last finished schooling level, current and previous employment status, wages earned in last 12 months, earnings both in cash and in-kind from last 12 months, expenditures and household asset ownership. The pooled cross-sectional data set contains information on 98,568 households.

Our purpose in this study is to investigate the income mobility in the population. In order to achieve our goal, parents and their children should be identified in the sample. Since the data used is not a longitudinal one, children who left a household and form their own households cannot be matched with their parents. However, the data set allows us to match parents and children if they live in the same household. The sample constructed by choosing the households

where parents and their children live together may be a highly selected sample and may not be representative.

Descriptive statistics will be given for two different samples. Table 1 gives the descriptive statistics for children who worked at least 1 hour and earned wages in the last 12 months and who have fathers worked at least 1 hour and earned non-zero wages in the last 12 months. Table 2 gives the descriptive statistics for children who had non-zero earnings in the last 12 months (total of wage and non-wage earnings) and who have fathers with non-zero earnings in the last 12 months. Selecting fathers and sons/daughters with non-zero wages or non-zero earnings is aimed to capture earnings measures as close as possible to "permanent" measures. This type of a selection eliminates the transitory movements around the permanent earnings measures due to students moving from school to labor force or older workers moving from labor force to retirement.

Each of the samples constitute of sons and daughters who are between ages 25 and 34. In each table, the descriptive statistics are given for the oldest son or oldest daughter present in the household and their matched fathers. Choosing the oldest son or oldest daughter is to preserve independence across observations and to reduce potential life-cycle bias as individuals in their early ages may have less earnings or wages due to having less experience (Zimmerman, 1992). In each table, annual wages and annual earnings are reported in December, 2011 Turkish Lira. In each table, household wealth quintiles are created using household's asset ownership information as basis for the household wealth. Principal component analysis method was used to transform the information in asset ownership into household wealth index which is then utilized in creating the household wealth quintiles. Using household asset ownership information to measure families' wealth is due to Acosta (2006) who argues that asset ownership information is less prone to measurement error, and therefore a more reliable measure of family income. Some of the regressions in this study makes use of household wealth quintile as a proxy for father income, and a measure for family income which may affect children's educational attainment and income.

Table 1 presents the descriptive statistics for the oldest son or oldest daughter who has non-zero wages in the last 12 months and for their matched fathers with non-zero wages in the last 12 months. Descriptive statistics for the mother are not presented because the sample size significantly decreases when it is restricted to families with both parents and their son/daughter having non-zero wages and in the intergenerational mobility studies the focus is on the fathers since in most populations the father is in the labor force whereas the mother stays at home. The sample consists of 2,372 father-son pairs who have non-zero wages. However, the sample size decreases to 202 father-mother-son pairs who all have non-zero wages. Similar decreasing trend applies for the sample sizes of father-daughter pairs with non-zero wages and of father-mother-daughter pairs who all have non-zero wages. In both samples, the mean age of oldest sons and oldest daughters as well the mean age of fathers seem to be similar. The mean annual wage for sons is slightly less than the mean annual wage for daughters. Fathers in daughters' sample earn more wages on average than fathers in sons' sample. It is interesting to observe the change in educational attainments across fathers' and sons' generations, similarly across fathers' and daughters' generations. In fathers-sons sample, fathers with lower secondary education or less than lower secondary education constitute 77% of all fathers. 8% of the fathers have high school education and only 5% have education level over high school. 7% of fathers do not know how to read and write. Illiteracy rate significantly decreases to 0.004% in sons' generation. The share of sons with lower secondary education or less than lower secondary education is less than the corresponding figure for their fathers. The high school share increases from 8% to 31% moving from fathers' generation to sons' generation. Lastly, the share of high-level education tripled in sons' generation. Similar patterns arise for fathers-daughters sample with an important difference; fathers' in daughters' sample have lower share of illiteracy and low-level education

and havehigher shares of middle level and high-level education compared to fathers in sons' sample which is in line with the difference in the average annual wages for fathers' in sons' sample and for fathers' in daughters' sample. In fathers-sons sample, 33% of all fathers work as a farmer or livestock worker which most likely implies that these families are located in rural areas. Corresponding figure for fathers-daughters sample is around 20%. Since the occupational opportunities in rural areas are not that much in number compared to urban areas and on average occupations in rural areas may pay less wage compared to occupations in urban areas, it may be important to control for location of residence to account for heterogeneities in labor market opportunities. As expected, a low share of daughters is married in fathers-daughters sample. Households are evenly distributed with respect to household wealth in fathers-sons sample whereas in fathers-daughters sample there is a higher share of households in top quintile.

The way the families chosen to take place in the sample may create a highly selected sample and as Solon (1992) argues may result in homogenous fathers and children samples. As mentioned before, homogenous fathers sample is not a problem in consistently estimating intergenerational income mobility correlations if permanent income of fathers can be observed. If fathers' permanent income is measured with error and the sample variance of fathers' long run income is lower than population variance of fathers' long run income, then homogenous fathers sample worsens the errors in variables bias and increases the downward bias in the income mobility estimate. Long run income measures are observed in our data set neither for sons/daughters nor for fathers thus, to estimate it for fathers, fathers' wages are regressed on their own ages and the residual is taken to be a proxy for the permanent income (Behrman and Rosenzweig, 2002; Solon, 1992). As a result, the sample standard deviation of fathers' annual wages is estimated to be 17,648 for fathers-sons sample and the sample standard deviation of fathers' annual wages is estimated to be 24,393 for fathers who have at least one child at home, including those fathers who have sons between ages 25 and 34. While selecting the fathers sample with at least one child at home, we make the assumption that at least a share of the fathers who have a child at home may have children aged between 25 and 34 who had left the home to form his own household. The difference in variances between fathers' samples make us suspicious that making use of the fathers-sons sample will result in deepened downward bias in income mobility estimates. The expectation is to have a highly selected sons' sample since in sons' sample we have the sons who live with their parents and the reason for sharing the same household may be the inadequate resources for the sons to live on their own. Thus, it is expected to observe a densely distributed wage distribution for sons in fathers-sons sample. The sample standard deviation of sons' annual wages is estimated to be 10,504 for fathers-sons sample whereas the sample standard deviation of sons' annual wages is estimated to be 14,891 for sons who may live in a household with a father present or not or who may form their own households. In both samples the sons are restricted to be between ages 25 and 34. The assumption made to make use of sons who form their own households is that a portion of sons who form their own households may have fathers who work but they are not observed in the data. As expected, the sample variance of sons' wages in fathers-sons sample is less than the sample variance of sons' wages for sons who may live in a household with a father present or not or who may form their own households. As suggested by Solon (1992), having a homogenous sons' sample with a lower variance in long run income compared to the population may downward bias the income mobility estimates.

Table 1: Descriptive Statistics[a]

Variables	Sons Sample		Daughters Sample	
	Dads	**Sons**	**Dads**	**Daughters**
Age	55.3	28.2	55.16	28.16
Annual wage	12,977	10,150	14,295	10,334
Log annual wage	9.47	9.22	9.56	9.24
Education[b]:				
No qualification	0.074	0.004	0.024	0.004
Low level	0.776	0.529	0.667	0.283
Middle level	0.089	0.31	0.166	0.286
High level	0.059	0.155	0.141	0.424
Occupation[c]:				
Top executive and managerial	0.146	0.084	0.192	0.053
Professional	0.016	0.067	0.05	0.188
Assistant professional	0.03	0.077	0.061	0.16
Clerical	0.029	0.059	0.042	0.231
Service and sales	0.069	0.143	0.087	0.119
Farmer and livestock workers	0.331	0.058	0.194	0.027
Craftsmen and foremen	0.125	0.228	0.136	0.082
Operatives	0.104	0.139	0.118	0.053
Unskilled labor	0.145	0.139	0.116	0.082
Married	0.981	0.486	0.978	0.074
No. Of siblings	-	2.32	-	2.17
Proportion of households by				
quintile of household wealth[d]:				
1st quintile(bottom)	0.158	-	0.063	-
2nd quintile	0.233	-	0.158	-
3rd quintile	0.215	-	0.145	-
4th quintile	0.201	-	0.273	-
5th quintile(top)	0.192	-	0.358	-

Notes: (a) The descriptive statistics are for sons and daughters who have earned wages in the last year and who have fathers with non-zero wages earned in the last year. (b) No qualification represents individuals who are illiterate. Low level represents individuals who are junior high school graduates or have less than junior high school level education. Middle level represents individuals who have high school diploma. High level represents individuals who have 2 year or 4 year or master's or PhD diploma. (c) Professions are categorized according to ISCO 88. (d) Principal component analysis is used to create household wealth index which is based on household's asset ownings and quintiles are created from household wealth index. Sons' sample consists of 2,372 father-son pairs. Girls' sample consists of 617 father-daughter pairs. The descriptive statistics for sons and girls are for the oldest son and oldest daughter in the household. Annual wages are in December 2011 Turkish Liras.

Table 2 presents the descriptive statistics for the oldest son or oldest daughter who has non-zero earnings in the last 12 months and for their matched fathers with non-zero earnings in the last 12 months. Earnings are described as the sum of labor earnings (wages) and non-labor earnings such as rents earned from real estate property. Fathers on average are 57 years old for both fathers-sons sample and fathers-daughters sample and this corresponds to an older fathers sample compared to the ones in Table 1. As in Table 1, fathers' annual earnings are on average higher than sons or daughters annual earnings, in addition sons and daughters on average earn similar. Educational attainment and household wealth patterns resemble their counterparts in Table 1. For the samples in Table 2, the homogenous sample of fathers and sons problem arises in the same pattern as in Table 1.

Table 2: Descriptive Statistics[e]

Variables	Sons Sample		Daughters Sample	
	Dads	**Sons**	**Dads**	**Daughters**
Age	57.31	28.53	57.35	28.61
Annual earnings	14,396	10,691	16,694	10,227
log annual earnings	9.57	9.27	9.72	9.23
Education[f]:				
No qualification	0.083	0.013	0.039	0.028
Low level	0.774	0.482	0.69	0.276
Middle level	0.09	0.338	0.158	0.312
High level	0.052	0.165	0.112	0.382
Occupation[g]:				
Top executive and managerial	0.14	0.086	0.181	0.049
Professional	0.015	0.068	0.044	0.194
Assistant professional	0.031	0.077	0.057	0.145
Clerical	0.028	0.069	0.042	0.248
Service and sales	0.07	0.157	0.081	0.129
Farmer and livestock workers	0.352	0.067	0.212	0.027
Craftsmen and foremen	0.117	0.206	0.131	0.065
Operatives	0.105	0.138	0.127	0.053
Unskilled labor	0.137	0.129	0.12	0.086
Married	0.97	0.455	0.969	0.086
No. Of siblings	-	2.17	-	2.08
Proportion of households by				
quintile of household wealth[h]:				
1st quintile(bottom)	0.14	-	0.07	-
2nd quintile	0.236	-	0.152	-
3rd quintile	0.198	-	0.144	-
4th quintile	0.213	-	0.27	-
5th quintile(top)	0.211	-	0.362	-

Notes: (e) The descriptive statistics are for sons and daughters who have non-zero earnings in the last year and who have fathers with non-zero earnings in the last year. Sons' sample consists of 5,492 father-son pairs. Girls' sample consists of 1,695 father-daughter pairs. See Table 1 notes (b), (c), and (d) for (f), (g) and (h), respectively.

4. METHODOLOGY

Based on the discussion in section 2, I decide to adopt the empirical framework of Solon (1992) and estimate the following equation using OLS and IV:

$$y_{1it} = \alpha + \gamma_t + \rho y_{0it} + \beta_1 A_{1it} + \beta_0 A_{0it} + \varepsilon_{it} \tag{9}$$

the indices 1 and 0 refer to the child and father, respectively. The equation differs from (2) by including the age of the father and of the child which accounts for the life-cycle effects on income[1]. The dependent variable is the log of earnings (wages) of the child and the independent variable is similarly the log of earnings (wages) of the father. The income figures are in December 2011 Turkish Liras to account for the confounding impact of price inflation over the study period 2003-2011. The instrument is the father's highest achieved schooling. Due to measurement error in fathers' permanent income and homogeneity of the fathers sample, OLS yields downward biased intergenerational mobility coefficient estimates. The instrument is most likely not exogenous to equation (9); therefore, leads to inconsistent estimation of mobility coefficients. Nevertheless, owing to the positive association in heritable endowments through transmission of genetic ability across generations, IV estimation strategy is shown to yield upward biased intergenerational mobility coefficient estimates (Solon, 1992). Consequently, the two estimation strategies—OLS and IV—bound the true population mobility coefficients from the bottom and the top.

5. RESULTS

Table 3 presents the intergenerational income mobility estimates for fathers-sons sample. The theoretical framework presented in Solon (1992) was adapted to run the regressions. Dependent variables are log of son's wages and log of son's earnings. Independent variables are log of father's wages, log of father's earnings, household wealth quintile, age of the son and age of the father. I test with three different measures of father's income to check the sensitivity of intergenerational mobility estimates to a change in the definition of father's income. First 3 columns present the OLS results which are subject to errors in variables problem due to having a cross-sectional observation for permanent income measures. Also, as argued in Table 1 and Table 2 the way the samples created may end up with homogenous samples for both sons and fathers which will worsen the downward bias in the mobility estimates (Solon, 1992). Thus, the first 3 columns form a lower bound estimate for the intergenerational income correlations. Columns 4-6 tries to solve the errors in variables and homogenous fathers sample problems by making use of instrumental variable estimation method. Log of father's wage, log of father's earnings and quintile of household wealth are instrumented by father's last finished schooling indicator which is an ordinal variable with 11 distinct values. After controlling for age effects, the OLS results imply that the population is quite mobile with a correlation of 0.11 for log of wages and a correlation of 0.10 for log of earnings. In other words, the place of the son in his wage/earnings distribution is poorly correlated with the place of his father in his own wage/earnings distribution. Furthermore, the similarity of mobility estimates in first two columns suggest that the association in intergenerational income is not due to how income is

---

[1] Year fixed effects combined with the age of the child helps account for unobserved cohort differences affecting the mobility estimates. A national education reform in 1997 increased compulsory schooling from 5 to 8 years. Cohorts born in 1987 or after are exposed to the education reform. Therefore, in our sample of father-child pairs young cohorts are exposed to the education reform and older cohorts are not. The education reform might have created a systematic difference in earnings of young and older child cohorts that is confounding the intergenerational mobility estimates.

defined. Column 3 presents an insightful interpretation to the correlation between father's and incomes. Quintile of household wealth is used as a proxy for the permanent income measure of the father. Looking at the results, it is obvious that a son living in a household placed in the top quintile has a 95% higher wage compared to a son living in a household placed in the bottom quintile. The slope increases monotonically as one climbs to the next step in the wealth distribution which advocates for persistence in intergenerational income across generations. The instrumental variables estimation results imply that the population is not that mobile at all with a correlation of 0.51 for log of wages and a correlation of 0.43 for log of earnings. In addition, the estimated household wealth quintile impacts are higher with IV compared to OLS estimates and only the coefficients for the third and top quintiles are statistically significant at nominal levels. Having the lower and upper bounds for the estimates of intergenerational income correlations we may argue that the true population correlations are in intervals (0.12, 0.51) and (0.10, 0.43) for wages and earnings, respectively. The intervals for the consistent parameter estimates are not narrow; therefore, it does not help us pinpoint the true income mobility correlation around a fixed value. However, we can compare our results with international comparators. Solon (1992) argues that the true intergenerational correlation in log earnings is around 0.4. In our study, the upper bound is around 0.4. Therefore, the population in our study seems to be more mobile with respect to log of earnings compared to the US population examined in Solon (1992).

Table 4 presents the intergenerational income mobility estimates for fathers-daughters sample. After controlling for age effects, the OLS results imply that the population is mobile but at a lower rate compared to the fathers and sons population, with a correlation of 0.16 for log of wages and a correlation of 0.25 for log of earnings. The OLS estimates of household wealth quintile impacts are intensified compared to the estimates for fathers-sons sample. The instrumental variables estimation results imply that the fathers-daughters population can be considered immobile with a correlation of 0.79 for log of wages and a correlation of 1.03 for log of earnings. In addition, the estimated household wealth quintile impacts are higher with IV compared to OLS estimates as in fathers-sons sample and none of the coefficients are statistically significant at nominal levels. This finding suggests that on average wage of a daughter is uncorrelated with her father's income which contradicts with the implications of large IV estimates in columns 4 and 5. The upper bound estimate for the intergenerational correlation in log earnings is not insightful since the mobility correlation is constructed to be between 0 and 1 with 1 implying a perfectly immobile population. What can be deduced from this result is that the true intergenerational correlation in log earnings is above 0.25 for fathers-daughters population. The true intergenerational correlation in log wages is between 0.16 and 0.79. Since the intervals are too wide, we cannot fix the true intergenerational correlations around a point. However, fathers-daughters population seems to be more immobile compared to fathers-sons population.

Table 3: Intergenerational Income Mobility Estimates (Fathers-Sons)

VARIABLES	DEPENDENT VARIABLES					
	OLS results			IV estimation results		
	(1)	(2)	(3)	(4)	(5)	(6)
	log of son's wage	log of son's earnings	log of son's wage	log of son's wage	log of son's earnings	log of son's wage
Quintile of households by household wealth:						
2nd quintile			0.335***			1.223
			(0.0635)			(1.017)
3rd quintile			0.508***			1.254**
			(0.0646)			(0.513)
4th quintile			0.697***			0.915
			(0.0659)			(0.723)
5th quintile(top)			0.957***			1.646***
			(0.0667)			(0.389)
father's age	-0.0586	0.00889	0.00696	0.0514	0.102***	0.0327
	(0.0510)	(0.0297)	(0.0489)	(0.0591)	(0.0326)	(0.0658)
son's age	0.0327	0.0187	0.0249	0.0685	0.0216	0.0340
	(0.0496)	(0.0306)	(0.0475)	(0.0550)	(0.0320)	(0.0559)
log of father's wage	0.119***			0.513***		
	(0.0176)			(0.0690)		
log of father's earnings		0.100***			0.434***	
		(0.0149)			(0.0396)	
Observations	2,343	5,481	2,371	2,343	5,481	2,371
R-squared	0.021	0.008	0.094			

Notes: Year fixed effects are suppressed. Standard errors are in parentheses. *** significant at 1%, ** significant at 5% and * significant at 10%.

Table 4: Intergenerational Income Mobility Estimates (Fathers-Daughters)

	DEPENDENT VARIABLES					
	OLS results			**IV estimation results**		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	log of daughter's wage	log of daughter's earnings	log of daughter's wage	log of daughter's wage	log of daughter's earnings	log of daughter's wage
Quintile of households by household wealth:						
2nd quintile			1.208***			1.851
			(0.211)			(3.014)
3rd quintile			1.344***			1.426
			(0.213)			(1.706)
4th quintile			1.782***			1.713
			(0.198)			(1.567)
5th quintile(top)			2.009***			3.099
			(0.194)			(1.945)
father's age	0.139	-0.0205	0.176	0.289*	0.117	0.265**
	(0.128)	(0.0669)	(0.116)	(0.152)	(0.0771)	(0.120)
daughter's age	0.217*	0.0901	0.229**	0.295**	0.121	0.240
	(0.121)	(0.0678)	(0.111)	(0.141)	(0.0766)	(0.171)
log of father's wage	0.165***			0.796***		
	(0.0424)			(0.144)		
log of father's earnings		0.249***			1.032***	
		(0.0361)			(0.0991)	
Observations	613	1,695	617	613	1,695	617
R-squared	0.031	0.028	0.183			

Notes: Year fixed effects are suppressed. Standard errors are in parentheses. *** significant at 1%, ** significant at 5% and * significant at 10%.

6. CONCLUSION

This study presents a thorough review of the literature on intergenerational income mobility with a focus on the issues regarding consistent estimation of the extent of persistence in income across generations and on the proposed solutions. In the light of the discussions this paper analyzes the intergenerational income association in Turkey using OLS and IV estimation strategies to put a boundary from—below and above—on the true population mobility correlation. The sample is extracted from Household Budget Surveys and includes father-son and father-daughter pairs with either non-zero wages or non-zero earnings living in the same household in any of the years between 2003 and 2011. The sample selection criteria probably worsen the downward bias in mobility estimates using OLS due to the homogeneity of fathers (i.e., the distribution of fathers' income is more densely distributed around the mean compared to the income distribution of an unrestricted sample of fathers) which are estimated around 0.10 for fathers-sons and around 0.20 for fathers-daughters. IV estimates are shown to be upward biased in the presence of genetically transmitted ability; hence, provide an upper bound to the mobility estimates (Solon, 1992). Resultingly, the true mobility correlations in the population are estimated to be in the range of (0.10, 0.51) for sons and (0.17, 1.00) for daughters. The intervals for true population correlations are narrower for sons and provide evidence that is on par with income mobility in US for sons that were born in 1950s. The results for mobility correlations of daughters are inconclusive as the range of estimates includes almost all possibilities (i.e., mobility correlations theoretically are in [0,1] closed interval). However, the results may point to higher persistence in intergenerational income for daughters than for sons.

There may be large variation in mobility estimates across regions of Turkey as prior studies have shown regional disparities in economic and social development. Further work may address the presence and causes of regional disparities in intergenerational income mobility as well the drivers of national persistence in intergenerational income. Persistency in national intergenerational income found in this study proves to be an obstacle in overall economic development which may call for the government to generate policies that break the harmful intergenerational income link either through providing equal opportunity in accessing education for low-income families' children in their early childhood or through redistributing income across poor families that help them invest more in their children's human capital.

Research and Publication Ethics Statement

This study has been prepared in accordance with the rules of scientific research and publication ethics.

Contribution Rates of Authors

The author's contribution to the article is 100%.

Conflicts of Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

REFERENCES

Acosta, P. (2006). "Labor supply, school attendance, and remittances from international migration: The case of El Salvador." World Bank Policy Research Working Paper No. 3903.

Atkinson, A. B., Maynard, A. K., and Trinder, C. G. (1983). *Parents and children: Incomes in two generations*.London: Heinemann Educational Books.

Baker, M. (1990). *Essays on unemployment spell duration and life cycle earnings*. Ph.D. dissertation, University of Michigan, Michigan, US.

Behrman, J. R., and Taubman, P. (1985). Intergenerational earnings mobility in the United States: Some estimates and a test of Becker's intergenerational endowments model. *The Review of Economics and Statistics, 67(1),* 144-151.

Behrman, J. R., and Rosenweig, M. R. (2002). Does increasing women's schooling raise the schooling of the next generation? *The American Economic Review*, *92(1),* 323-334.

Chung, C. and Goldberger, A. S. (1984). Proportional projections in limited dependent variable models. *Econometrica,52,* 531-534.

Corcoran, M., Gordon, R., Laren, D. and Solon, G. (1992). The association between men's economic status and their family and community origins. *Journal of Human Resources, 27(4),* 575-601.

Dearden, L., Machin, S. and Reed, H. (1997). Intergenerational mobility in Britain. *Economic Journal*, *107(440),* 47-66.

Goldberger, A. S. (1981). Linear regression after selection. *Journal of Econometrics, 15,* 357-366.

Greene, W. H. (1990). *Econometric analysis*. New York: Macmillan.

MacDonald, G. M. and Robinson, C. (1985). Cautonary tails about arbitrary deletion of observations; or, throwing the variance out with the bathwater. *Journal of Labor Economics*, *3,* 124-152.

Mayer, S. E. (2010). Revisiting an old question: How much does parental income affect child outcomes? *Focus,27(2),* 21-26.

Menchik, P. L. (1979). Inter-generational transmission of inequality: An empirical study of wealth mobility. *Economica,46(184),* 349-362.

Sewell, W. H. and Hauser, R. M. (1975). *Education, occupation, and earnings: Achievement in the early career*. New York: Academic Press.

Shea, J. (2000). Does parents' money matters? *Journal of Public Economics*, *77,* 155-184.

Solon, G. (1992). Intergenerational mobility in the United States. *The American Economic Review, 82(3),* 393-408.

Solon, G., Corcoran, M., Gordon, R. and Laren, D. (1991). A longitudinal analysis of sibling correlations in economic status. *The Journal of Human Resources*, *26(3),* 509–534.

Zimmerman, D. J. (1992). "Regression toward mediocrity in economic stature." *The American Economic Review*, *82(3),* 409-429.