Exploring the sources of earnings transmission in Spain

María Cervini-Plá

Universitat de Girona

20. January 2012

Exploring the sources of earnings transmission in Spain.*

María Cervini-Plá[†]

January 20, 2012

Abstract

This paper explores the mechanisms behind the intergenerational earnings mobility in Spain by means of three exercises: calculating the transition matrix, decomposing the sources of earnings elasticity and estimating quantile earnings regressions. By calculating the transition matrices we find a strong degree of persistence in educational attainment and especially in occupation. By decomposing the sources of earnings elasticity across generations, we find that the correlation between children's and their fathers' occupations is the most important component. Finally, quantile regressions estimates show that the influence of the father's earnings is greater when we move to the lower tail of the offspring's earnings distribution, especially in the case of daughters' earnings.

Keywords: Intergenerational mobility, earnings, transition matrix, quantile regression, two sample two stage least square estimator, Spain.

JEL classification: D31, J31, J62.

*I would like to thank Xavier Ramos, Olga Cantó and Cheti Nicoletti for their useful suggestions. I am also grateful to participants at XXXIII Simposio Análisis Económico in Zaragoza and XVI Encuentro de Economía Pública in Granada for helpful comments on a previous version of this paper. This research was initiated while I was visiting the Institute for Social and Economic Research (University of Essex) with an ECASS Scholarship. Financial support from projects ECO2010-21668-C03-02 (Ministerio de Ciencia y Tecnologa), 2009SGR-307 and XREPP (Direcci General de Recerca) is gratefully acknowledged.

[†]Departament d'Economia. Universitat de Girona. Campus Montilivi, s/n. 17071 Girona, Spain. Phone number +34 972418781 / Fax +34 972418032. E-mail: maria.cervini@udg.edu

1 Introduction

Intergenerational mobility is a measure of the changes in socioeconomic status that occur from the parents' to the children's generation and can also be seen as an indicator of equality of opportunity. A society with more intergenerational mobility implies that the socioeconomic status of children is less determined by the socioeconomic status of their parents and in this sense we can say that there is greater equality of opportunity.

Intergenerational mobility is generally measured in terms of intergenerational elasticity, or a statistical correlation between parents' and children's economic standings. The higher the intergenerational elasticity, the less social mobility a society offers. Economists have primarily concentrated on the relationship between parents and their offspring's permanent income or earnings, while sociologists explore measures of the association between ordered categorical variables, such as social and economic class.[1]

In this paper, we follow the economic approach and focus on intergenerational mobility measured by the intergenerational elasticity of offspring's earnings with respect to their fathers' earnings. The main objective of our paper is to explore the intergenerational elasticity of earnings in Spain. We investigate the underlying mechanisms of this earnings transmission by means of three exercises. First, we calculate the transition matrices; second, we decompose the sources of earnings elasticity and, finally, we estimate the influence of fathers' earnings by quantiles.

Intergenerational mobility in Spain has primarily been studied by sociologists. For example, Carabaña (1999) studied occupational mobility. From an economic point of view, Sanchez-Hugalde (2004) analyses intergenerational income and education mobility using the Family Expenditure Survey (Encuesta de Presupuestos Familiares) for 1980 and 1990; however, she only estimates the elasticity when children and their parents live together.

More recently, Cervini-Plá (2011) provides new evidence on intergenerational earnings and income mobility for Spain. Since there are no Spanish surveys covering long-term information on both children and their fathers' income or earnings, she deals with this selection problem using the two-sample two-stage least squares estimator.

[1]See Solon (1999), Björklund and Jäntti (2000), Bowles and Gintis (2002), Erikson and Goldthorpe (2002) for a review.

She finds that intergenerational mobility in Spain is similar to France, lower than in the Nordic countries and Britain and higher than in Italy and the United States. Furthermore, she uses the Chadwick and Solon (2002) approach to explore intergenerational mobility in the case of daughters overcoming employment selection, and she find similar results by gender.

In order to overcome the co-residence selection problem[2], as in Cervini-Plá (2011), we estimate intergenerational earnings mobility using the two-sample two-stage least squares (TSTSLS) estimator.[3] This method combines information from two separate samples: a sample of adults (sons and daughters) with observations of their earnings and their parents' characteristics, and a sample of potential parents with observations of earnings and the same characteristics. The latter sample is used to estimate an earnings equation for parents using their characteristics as explanatory variables, while the former is used to estimate an intergenerational earnings equation by replacing the missing parental earnings with its best linear prediction.

The second problem we try to correct is the employment selection, wherein we only have earnings for adults who are employed. Since the decision to work or not work is not random, especially in the case of women, estimating intergenerational earnings mobility only for those who are working gives us biased estimators. We deal with this selection problem using a different approach to that of Cervini-Plá (2011); we use the Heckman-type of correction estimation described in Vella (1998) and used in Ermisch, Francesconi, and Siedler (2006). Therefore, another contribution of this paper is the consideration of employment selection in this way.

Correcting for co-residence and empoloyment selection, we find an elasticity of 0.38 for sons between 30 and 40 years and an elasticity of 0.42 for sons between 40 and 50

[2]Following Nicoletti and Francesconi (2006) we refer to the co-residence selection problem by the fact that, in a panel, we have information regarding offspring's and parents' earnings when they live together in at least one wave; however, the probability of observing offspring living with their parents decreases as the children grow older. This selection problem is particularly important in Spain, where we have only short panels, and thus, we do not have information on both children's and their fathers' permanent earnings.

[3]Following the paper written by Angrist and Krueger (1992) on two-sample instrumental variables (TSIV) estimation, numerous empirical researchers have applied a computationally convenient TSTSLS variant to the study of intergenerational mobility, such as Björklund and Jäntti (1997) in Sweden; Fortin and Lefebvre (1998) in Canada; Grawe (2004)) in Ecuador, Nepal, Pakistan, and Peru; Lefranc and Trannoy (2005) in France; Nicoletti and Ermisch (2007) in Britain; and by Mocetti (2007) in Italy.

years. In the case of daughters, we obtain elasticities of 0.36 and 0.49 respectively.

The high values of the diagonals of our transition matrices show a high degree of persistence in education and occupation. By decomposing the sources of earnings correlations, we find that the correlation between children's and fathers' occupation is the most important component. Despite the strong correlation in education between parents and children, the father's occupation is a good indicator of his social position and is better than his education as a predictor of his children's earnings. Finally, when we estimate the elasticity between children's and fathers' earnings by quantiles, we find that the influence of the fathers' earnings is greater when we move to the lower tail of the distribution, especially in the case of daughters.

The rest of the paper is organised as follows. In the next section, we present a very simple theoretical framework that allows us to understand some of the sources of earnings transmission between generations. Section 3 describes how we implement the two-sample two-stage least square estimator. In Section 4 we describe the data source, the selection sample, and the variables used in the empirical analysis. Section 5 reports the results, and finally, Section 6, concludes with some final remarks.

2 Sources of earnings transmission

Why do some children obtain better jobs and higher earnings when they become adults, while others do not? Through what channels are earnings transmitted?

As Nicoletti and Ermisch (2007) point out, an important number of institutions affect intergenerational mobility, such as the education system, the labour market and the family (particularly its investment in children). Furthermore, public policy affects these institutions and hence intergenerational mobility.

Following Checchi (2006) and Lefranc and Trannoy (2005), we present here a simple model that allows us to better understand some of the sources of intergenerational earnings transmission.[4]

Let us suppose an individual belonging to family i and to generation t, whose permanent earnings W_{it} derives from two components: ability endowment A_{it}, and

[4]Here we will summarise the main channels of earnings transmission; however, most of these channels are the same when we analyse the transmission of income.

human capital (i.e., education E_{it}). If we do not consider on-the-job training, education comes before entering employment, and therefore, also comes prior to earnings. If we consider the fact that ability increases labour productivity, we can express these channels as follows:

$$W_{it} = \lambda E_{it} + \pi A_{it} + \mu_{it} \tag{1}$$

Where for simplicity's sake, the relationship between earnings, education and ability is assumed to be linear, and μ_{it} is an independent and identically distributed (i.i.d.) error term that captures the idea of luck in the labour market.

Education is one of the most important channels of intergenerational earnings transmission. Educational attainment can be determined by the cultural influence of the family. We represent the education of the previous generation by E_{it-1} and the coefficient η shows this channel. There is a large body of empirical evidence about how the children of educated parents are more likely to acquire education. As Checchi (2006) points out, this may be partly due to parental imitation, but in most cases, it works through induced educational choices. An educated parent is more aware of the psychological and economic value of education, and therefore, puts more pressure on his or her children to achieve more at school.

Furthermore, in the presence of liquidity constraints, education is also determined by family earnings, reducing the optimal investment in education by poor families. We indicate this channel with γ and we write:

$$E_{it} = \eta E_{it-1} + \gamma W_{it-1} \tag{2}$$

Therefore, education is determined by the education and earnings of the previous generation; however, if we substitute E_{it-1} with the expression with one-period lag successively we can observe that education depends on the earnings of the parents, grandparents, and previous generations.

The second component we have in equation 1 is ability. Considering that ability is genetically (or mechanically) inherited, we can indicate this effect with α, while $t-1$ again represents the previous generation. Ability can influence earnings directly through the type of job obtained because people with greater ability are more productive. Therefore, we have:

$$A_{it} = \alpha A_{it-1} \tag{3}$$

This component can be considered to reflect some aspects of earnings determinants that "money can't buy", and at the same time, are transmitted from one generation to the next. Examples of this are the transmission of intelligence quotient (IQ), social networks, and preferences.

Another source of intergenerational earnings persistence emerges from territorial segregation. One's neighbourhood can influence earnings through education (better quality of schools) or through social capital (good neighbours help obtain better jobs). Neighbourhoods with better schools and better neighbours generally have higher house prices. Therefore, residential choices are ultimately determined by family earnings.

Another channel is networks *per se*. Obtaining a good and well-paid job may depend on friends and social networks rather than on one's curriculum.

We consider the combined effect of family networking and residential choices on offspring's earnings. Since both components are related to family earnings, we indicate this channel with θ and extend equation (1) with an additional term:

$$W_{it} = \lambda E_{it} + \pi A_{it} + \theta W_{it-1} + \mu_{it} \tag{4}$$

Taking into account each of these channels, we can observe that intergenerational persistence is a dynamic system. From an empirical point of view, it is not easy to distinguish between alternative explanations of intergenerational persistence of earnings. It is important to note that in a simple regression of child's earnings on parents' earnings, the coefficient will simultaneously capture all of the effects "that money can buy". Hence, standard estimates of intergenerational earnings regressions will provide an upward-biased estimates of the causal effect of parents' earnings on their children's earnings. Concretely, we will estimate:

$$W_{it} = \beta W_{it-1} + \mu_{it} \tag{5}$$

However, from a policy point of view, the distinction between the different components is relevant in the prediction of the impact of economic policies and in the knowledge of which policy could better improve mobility.

3 Estimation method and sample selection problems

As we explained above, the standard measure of intergenerational mobility is earnings elasticity. More precisely, we consider the following intergenerational mobility equation:

$$W_{it} = \alpha + \beta W_{it-1} + \mu_{it} \tag{6}$$

where W_{it} is the offspring's log earnings, W_{it-1} is the fathers' log earnings (the earnings of the previous generation), α is the intercept term representing the average change in the child's log earnings, and μ is a random error. The coefficient β is the intergenerational elasticity of offspring's earnings with respect to their father's earnings, and is our parameter of interest.

Children's earnings are positively related to their fathers' earnings of parents, therefore, the β coefficient varies between 0 and 1. A high value of β implies a high persistence of wages and therefore a less mobile society. Conversely, a value of β close to zero is representative of a very mobile society where the children's socioeconomic status is not strongly influenced by the position of the parents.

If we had permanent income for successive generations in our sample, we would have no problem with directly estimating equation (6) using the ordinary least squares estimator. Unfortunately, we do not have this information in one data set.

First, most data sets only provide measures of current earnings and fail to provide measures of individual permanent income. Solon (1992) and Zimmerman (1992) show that the use of current earnings as a proxy for permanent earnings leads to downward OLS estimates of β. Different solutions can be implemented to reduce or eliminate this bias. If we work with panel data, we can calculate an average of current earnings over several years as a proxy of permanent income. Another possibility lies in using instrumental variables to estimate β. In this paper, we estimate father's earnings using auxiliary variables. Therefore, the estimated earnings are an average that can be considered a proxy of the father's permanent earnings. In the case of children, we select adult ages that are as close as possible to the age at which earnings are similar to permanent income. In particular, Haider and Solon (2006) suggests the use of offsprings of around 40 years of age.

Second, we only observe earnings for pairs of parents and children when they live together in at least one wave of the panel. On the contrary, we do not have information for sons and daughters who never co-reside with their parents during the panel. This selection problem is particularly important in short panels and could lead to a sub-estimation of the offsprings' earnings, since their living in the parental household is due to the fact they are still students or they do not have enough income to live independently. Therefore, this selection problem causes an overestimation of intergenerational mobility (an underestimation of the elasticity between parents' earnings and offsprings' earnings).

In this paper we deal with this selection problem by linking two samples and using the two-sample two-stage least squares estimator (TS2SLS). The TSTSLS estimator is a computationally easier variant of the two-sample instrumental variable estimator (2SIV) described by Angrist and Krueger (1992), Arellano and Meghir (1992), and Ridder and Moffit (2006).[5] Concretely, in the two-sample context, unlike the single-sample situation, the IV and 2SLS estimators are numerically distinct. Inoue and Solon (2010) derive and compare the asymptotic distributions of the two estimators and find that the commonly used TSTSLS estimator is more asymptotically efficient than the TSIV estimator because it implicitly corrects for differences in the distribution of variables between the two samples. They therefore explain that, although computational simplicity was the original motive that drew applied researchers to use the TSTSLS estimator instead of the TSIV estimator, it turns out that the TSTSLS estimator is also theoretically superior.

Since we do not have information about W_{it-1}, but do have a set of instrumental variables Z of W_{it-1}, we can estimate equation (6) in two steps. We consider two different samples: the first, which we call the main sample, has data on offspring log earnings, W_{it}, and characteristics of their fathers, Z, while the second, which we call

[5]For a detailed description of the properties of this estimator, see Arellano and Meghir (1992), Angrist and Krueger (1992) and Ridder and Moffit (2006).

the supplemental sample, has information on fathers' log earnings, W_{t-1}, and their age, education, and occupational characteristics, Z.

In the first step, we use the supplemental sample to estimate a log earnings equation for fathers using, as explanatory variables, their characteristics, Z, that is:

$$W_{t-1} = Z_{t-1}\delta + v_i \tag{7}$$

In the second step, we estimate the intergenerational mobility equation (6) using the main sample and replacing the unobserved W_{it-1} with its predictor,

$$\widehat{W_{it-1}} = Z_{it-1}\hat{\delta}, \tag{8}$$

where $\hat{\delta}$ represents the coefficients estimated in the first step, and Z represents the variables observed in the main sample. Thus, we estimate equation (6) by using the fathers' imputed earnings.

$$W_{it} = \alpha + \beta(Z_{it-1}\hat{\delta}) + u_i \tag{9}$$

The $\hat{\beta}$ we obtain is the TSTSLS estimate of intergenerational earnings elasticity. The standard errors are properly estimated as Murphy and Topel (1985) and Inoue and Solon (2010) propose. In order to take into account the life-cycle profiles, the estimation of both equations includes additional controls for individual's and father's ages.

The properties of the two-sample estimator depend on the nature of the instrument used. Nicoletti and Ermisch (2007) express how important it is to choose instrumental variables that are strongly correlated with the variable to be instrumented. Therefore, we have to choose the instruments in order for the R^2 of the regression be as high as possible.

Furthermore, consistency requires the error term in the intergenerational mobility equation to be independent of the instrumental variables or that the instrumental variables explain perfectly the father's missing earnings.

As Nicoletti and Ermisch (2007) point out, the TSTSLS estimator of the intergenerational elasticity could be under- or overestimated when the auxiliary variables are endogenous. Moreover, since the instruments we use -paternal educational and occupational characteristics- are likely to be positively related to the sons' earnings even after controlling for the fathers' earnings, the bias is probably positive. Therefore, the potential endogeneity problem is likely to affect most of the empirical papers on intergenerational mobility applying 2SIV and TSTSLS estimators.

Finally, another problem arises when we study intergenerational mobility, is the employment selection wherein we only observe earnings for adults when they are employed; however, the decision to work or not to work is not random, especially in the case of women. Therefore, those who are working are a self-selected sample. Estimating intergenerational earnings mobility exclusively for those who are working yields biased estimators. We deal with this selection problem in the case of daughters by using the Heckman-type of correction estimation described in Vella (1998) and used in Ermisch, Francesconi, and Siedler (2006). In particular, the variables included in the selection equation are dependent children, marital status, age and father's earnings. In all regressions, these are good predictors of participation.

4 Data Sources and Sample Selection Rules

As we explained above, we combine two separate samples to estimate intergenerational earnings mobility, a main sample and a supplemental sample.

In our case, the main sample is the Survey of Living Conditions (Encuesta de Condiciones de Vida (ECV)) for the year 2005, that is, the Spanish component of the European Union Statistics on Income and Living Conditions (EU-SILC).[6]

The ECV has annually interviewed a sample of about 14,000 households that are representative of the Spanish households, and has kept each household in the sample for four years. Personal interviews are conducted at approximately one-year intervals with adult members of all the households.

From the ECV, we have information about adults' earnings and a set of characteristics of their fathers when they were between 12 and 14 years old.

[6]The EU-SILC is an instrument that aims to collect timely and comparable cross-sectional and longitudinal multidimensional microdata on income, poverty, social exclusion, and living conditions. This instrument is anchored in the European Statistical System (ESS).

Our supplemental sample is the Family Expenditure Survey of 1980-1981 (Encuesta de Presupuestos Familiares). This survey was designed with the purpose of estimating consumption and the weights of the different goods used in the consumer price index. In addition, we also have information regarding earnings, occupation, and the education level of the head of the household. Therefore, in this sample we have data on the father's earnings and the same set of their characteristics that are available in the main sample.

Although we have the same characteristics in both samples, we have to recode some variables to have a homogenous classification across surveys.[7]

Our main sample is composed of individuals, either the head of the household or their spouse, born between 1955 and 1975, self-employed or in paid employment, who report positive labour earnings and are full-time workers. Thus, in the year 2005, these adults were between 30 and 50 years old and were 12 or 14 years old between 1969 and 1989. This is why we use the Family Expenditure Survey of 1980-1981 as the supplemental sample to estimate fathers' earnings.

We suppose that when the children were 12 or 14 years old, their fathers were between 37 and 57 years old. Thus, when we estimate the fathers' earnings regression we select males between those ages.

As we have mentioned above, one problem that can bias intergenerational mobility studies is measurement error with regard to earnings. Theoretically, we would like to consider the intergenerational elasticity in long-run permanent earnings, but we can observe earnings only in a single or a few specific years. Thus, the question is, at what age should the current earnings should be observed in order to provide the closest measure of permanent earnings? Haider and Solon (2006) show that it is reasonable to choose sons around the age of 40 and fathers with aged between 31 and 55. Therefore, assuming that these results hold for other countries, we choose similar age intervals in our empirical application.

After the exclusions, we have a total of 4,352 pairs, and in this sample, we have employed fathers and children that reported positive earnings.

The earnings variable we use in all the specifications is the log of current gross

[7]For a detailed description of the frequencies of the different characteristics in the main and supplemental samples see Table A.1 in the Appendix.

Table 1: Descriptive statistics: Sons in the main sample after exclusions.

	sons 30-40	sons 40-50
Observations	1,334	1,322
annual earnings	19,728.35	22,403.7
log of annual earnings	9.72	9.84
Education		
Primary education	13.49%	19.48%
Secondary education (first step)	24.47%	25.00%
Secondary education (second step)	25.42%	24.59%
Vocational qualification	2.64%	1.73%
Higher education (university)	33.97%	29.21%
Occupation		
Higher-grade professionals	5.01%	6.6%
Higher-grade manager	11.65%	10.94%
Low grade professional	12.06%	9.97%
Routine non-manual employees high grade	7.99%	10.80%
Routine non-manual employees low grade	10.98%	9.28%
Skilled agriculture workers	2.37%	3.09%
Skilled manual workers	23.51%	22.70%
Low grade technician	12.33%	13.69%
Unskilled workers	14.09%	12.93%

annual earnings, which is almost directly collected (not imputed), and is not distorted by the national taxation system.

Tables 1 and 2 present the main descriptive statistics of our final sample of sons and daughters, respectively.

5 Results

5.1 Intergenerational earnings elasticity

In this subsection, we present the empirical results for intergenerational mobility estimation correcting for the sample selection problems. As we have explained before, we use a two-sample two-stage estimation, whose first step consists of the estimation of the fathers' earnings regression using the supplemental sample. The results of this regression are presented in the Appendix (Table A.2). These coefficients are then used to impute the fathers' earnings in the main sample, since we have the same characteristics in both samples (main and supplemental). Therefore, in the second step, using the

Table 2: Descriptive statistics: Daughters in the main sample after exclusions

	daughters 30-40	daughters 40-50
Observations	875	821
annual earnings	13,539.65	15,584.45
log of annual earnings	9.2	9.31
Education		
Primary education	10.39%	17.44%
Secondary education (first step)	19.95%	21.54%
Secondary education (second step)	21.78%	23.35%
Vocational qualification	2.35%	1.11%
Higher education (university)	45.52%	36.67%
Occupation		
Higher-grade professionals	1.59%	1.96%
Higher-grade manager	17.44%	19.54%
Low grade professional	11.68%	9.90%
Routine non-manual employees high grade	21.76%	16.89%
Routine non-manual employees low grade	21.08%	19.80%
Skilled agriculture workers	0.91%	0.85%
Skilled manual workers	4.85%	5.38%
Low grade technician	2.35%	1.71%
Unskilled workers	18.35%	23.98%

coefficients from the supplemental sample and the characteristics of the main sample, we estimate earnings for each father in the main sample.

Table 3 reports the second step, the coefficients of the intergenerational regression between annual children's earnings (sons and daughters) and the fathers' imputed earnings correcting for the co-residence selection problem and employment selection in the case of daughters.

The increase in female labour force participation in Spain began at the end of the 70s, but this participation is still lower compared to men. It is intuitive that full-

Table 3: Intergenerational elasticity correcting for the sample selection problems

	sons 30-40	sons 40-50	daughters 30-40	daughters 40-50
father's earnings	**0.380**	**0.427**	**0.369**	**0.498**
	(0.042)	(0.041)	(0.074)	(0.062)
Obs.	1334	1322	875	821
R^2	0.061	0.08	0.072	0.10

Note:Dependant variable is log of annual labor earnings. Fathers' earnings refers to the log of father annual labor earnings. Standard errors are corrected using Murphy and Topel (1985) and Inoue and Solon (2010) procedure.

time women workers are probably more common in some types of household (highly educated households or very poor households). We use the variables of being married, having children, and the father's earnings and age to correct for the employment selection.

In all columns, the father's predicted log earnings has a significant positive effect on child's earnings.

We estimate the elasticity for sons and daughters for two different cohorts, those aged between 30 and 40 and those between 40 and 50 in 2005. For sons (first and second columns), regression coefficients are 0.38 for the first cohort and 0.43 for the second cohort. In the case of daughters (third and fourth columns), the elasticities are 0.37 and 0.50, respectively. The elasticities differences between genders are not statistically significant and we can therefore say that the persistence of earnings is nearly the same by gender.

We observe smaller correlations for the younger cohorts. However we do not have enough information to know whether this is due to a changing trend in that there is greater mobility or whether this is only a matter of age in the sense that when these young sons grow older they become more correlated with their parents.

The figure of intergenerational earnings elasticity *per se* does not give much information. It is always useful to compare our estimates of intergenerational earnings mobility in Spain with the results obtained for other countries. However, we must be careful because the comparability of studies is problematic and very difficult since the estimates are sensitive to different factors such as the income measure used, the adequacy of the database, the different criteria for sample selection and the different estimation methods followed. Therefore, we must choose the studies that are most similar to ours in terms of choice of sample, using a two-sample approach.

In Table 4 we present the elasticities for different studies using the two-sample approach. As we can see, our elasticities are similar to those found in Cervini-Plá (2011) and show that Spain has similar economic mobility to France, higher than Italy and Brazil and lower than the Nordic countries, Canada and UK. The case of the U.S. is interesting because, contrary to popular belief, it has a high persistence of earnings. As Lefranc and Trannoy (2005) point out, one possible explanation for why Europe shows more intergenerational mobility than the United States is the way in which

Table 4: Comparable international elasticities

Studies:	Country	$\hat{\beta}$	Set of instruments
Björklund and Jäntti (1997)	Sweden	0.28	Education, occupation
Björklund and Jäntti (1997)	US	0.52	Education, occupation
Nicoletti and Ermisch (2007)	UK	0.29	Occupational prestige
Mocetti (2007)	Italy	0.50	Education, occupation
Piraino (2007)	Italy	0.51	Education, work status
Lefranc and Trannoy (2005)	France	0.41	Education, social class
Dunn (2007)	Brazil	0.69	Education
Ferreira and Veloso (2006)	Brazil	0.58	Education, occupation
Cervini-Plá (2011)	Spain	0.41	Education, occupation
Fortin and Lefebvre (1998)	Canada	0.21	Occupation

Note:The β is the father-son earnings elasticity. The dependent variable is log annual son's earnings.

higher education is financed.

Once we have estimated elasticities in the next subsections, as explained before, we shall explore intergenerational mobility by means of three exercises that we present in the following three sections.

5.2 Transition matrices

Another way to characterize intergenerational mobility is using mobility matrices. The idea is to characterize the conditional probabilities of transition between ordered groups. Table 5 gives the fraction of sons or daughters in each earnings quantile given the earnings quantile of their fathers. Each cell in Table 5 can be interpreted as the probability of a son or daughter being in quantile i^{th}, conditional on his or her father being in quantile j^{th}.

Table 5: Transition matrices of earnings between fathers and child

		Quantil of child				
		1	2	3	4	5
Quantil of father	1	30.08%	24.40%	19.12%	15.74%	10.66%
	2	23.93%	22.34%	23.54%	15.69%	14.50%
	3	16.98%	19.17%	20.26%	22.64%	20.95%
	4	16.20%	18.29%	21.67%	23.26%	20.58%
	5	13.23%	16.20%	15.66%	22.41%	32.49%

We note strong persistence at the extremes of the distribution. About 30 % of children remain in the bottom quintile of the earnings distribution if their father belongs to that quantile. The value is also high at the other extrem of the distribution, about 32.49% of the children of richer fathers are also in the richest quantile. Although we observe more mobility in the center of the distribution, high diagonal values of the matrix indicate high persistence of earnings throughout the distribution.

Education obviously plays a crucial role in understanding intergenerational mobility. Therefore, in Table 6 we present the transition matrix of education between fathers and children. It is observed that the generation of children is on average more educated than that of their parents. However, the values found in the diagonal are even higher than the earnings' transition matrix, indicating greater educational immobility. For example, around 74% of children with educated fathers are remain in the same group.

Table 6: Transition matrix of education between fathers and children

Father/ Child	1	2	3	4
1	23.73%	28.52%	24.34%	23.42%
2	4.44%	25.23%	34.58%	35.75%
3	3.37%	9.21%	35.28%	52.13%
4	0.57%	4.82%	20.71%	73.90%

Note:1:none or primary school, 2:lower secondary, 3:upper secondary, 4:university degree.

In Spain, as in other southern European countries, children leave the parental home at a very late age. Therefore, most of the school decisions are made when they still live with their parents, a fact that reinforces the influence of parents on children.

Furthermore, cost dimension is also important in understanding schooling decisions. In Spain, access to public tertiary education is not extremely expensive. However, universtity fees only represent a small proportion of annual expenditure in order to attend university, clearly the greatest cost is the opportunity cost of not working and this cost is particularly important in low income families.

Spain is also characterized by strong occupational persistence. Table 7 shows how occupations are linked across generations. Again, the high values of the diagonal of the matrix, even higher than in the education matrix, show that a high proportion of children have the same occupations as their parents. Intergenerational occupational immobility is linked with the existence of entry barriers limiting access to certain professions. Furthermore, in other cases, it is the natural result of educational stratification. Finally, another channel through which persistence in occupations works are family ties, for many jobs are filled through social referral.

Table 7: Transition matrix of occupations between fathers and children

Father/ Child	1	2	3	4
1	33.40%	40.61%	9.32%	16.67%
2	15.39%	50.11%	12.15%	22.35%
3	7.07%	24.08%	26.44%	42.41%
4	6.70%	20.82%	16.30%	56.17%

Note:1:unskilled workers, 2:skilled agriculture and manual workers, 3:low grade professionals, 4:higher grade professionals and managers.

Both the education and occupation matrices show that that major education and occupation immobility lies behind the strong persistence of earnings.

5.3 Decomposing earnings elasticity

The second exercise we perform to explore intergenerational mobility is a decomposition of the sources of earnings elasticity across generations. Using the decomposition developed by Bowles and Gintis (2002) and followed by Lefranc and Trannoy (2005), we can express offsprings' and fathers' earnings as:

$$W_{it} = Educ_i^c\ \delta_{educ}^c + Occup_i^c\ \delta_{occup}^c + \mu_i^c \quad for\ children's\ earnings \tag{10}$$

$$W_{it-1} = Educ_i^f\ \delta_{educ}^f + Occup_i^f\ \delta_{occup}^f + \mu_i^f \quad for\ father's\ earnings \tag{11}$$

where the supra-indices c and f are used to identify children's and fathers' characteristics respectively. The variable *Educ* is the individual's education, while *Occup* is the individual's occupation; these are the variables we have used to estimate fathers earnings in the supplemental sample.[8]

Thus, the elasticity β is simply given by:

[8] In order to provide an easy exposition, the variable age is ignored here; however, it is taken into account in the empirical implementation of the decomposition.

$$\beta = \frac{cov(W_{it}\ ,\ Educ_i^f \delta_{educ}^f + Occup_i^f \delta_{occup}^f)}{V(Educ_i^f \delta_{educ}^f + Occup_i^f \delta_{occup}^f)}$$

Then, we can rewrite β as a decomposition of six terms:

$$\beta = \frac{1}{V(Educ_i^f \delta_{educ}^f + Occup_i^f \delta_{occup}^f)} \times \Big[\delta_{educ}^c cov(Educ_i^c, Educ_i^f)\delta_{educ}^f$$
$$+ \delta_{occup}^c cov(Occup_i^c, Occup_i^f)\delta_{occup}^f + \delta_{educ}^c cov(Educ_i^c, Occup_i^f)\delta_{occup}^f$$
$$+ \delta_{occup}^c cov(Occup_i^c, Educ_i^f)\delta_{educ}^f + cov(\mu_i^c, Educ_i^f)\delta_{educ}^f + cov(\mu_i^c, Occup_i^f)\delta_{occup}^f\Big]$$

Bowles and Gintis (2002) remark that it is important to consider this decomposition as a descriptive device and not as an analysis of causal effects.

The results of applying this decomposition to the estimation of earnings elasticity presented in Table 3 are given in table 8

Table 8: Decomposition of earnings regression coefficient

	sons 30-40	sons 40-50	daughters 30-40	daughters 40-50
$educ_c - educ_f$	0.065	0.084	0.059	0.080
$occup_c - occup_f$	0.143	0.152	0.139	0.173
$educ_c - ocup_f$	0.080	0.082	0.073	0.096
$occup_c - educ_f$	0.055	0.071	0.063	0.093
$res_c - educ_f$	0.002	0.018	0.014	0.018
$res_c - occup_f$	0.035	0.020	0.023	0.038
total	**0.380**	**0.427**	**0.369**	**0.498**

As Lefranc and Trannoy (2005) observe, these results can be interpreted as, assuming that the only channel of intergenerational earnings correlation would work through the correlation of the fathers' and children's education, meaning that the elasticity coefficient for sons between 30 and 40 and their fathers' earnings would be equal 0.065.

Table 8 shows that, for all ages and for both sons and daughters, the correlation between children's and fathers' occupations is the most important component for understanding the intergenerational elasticity between earnings. Furthermore, the correlation between the father's occupation and his offspring's education is also important. If we add the influence of the father's occupation on his child's occupation and education, we explain almost half of the intergenerational elasticity coefficient. However, we can observe a slight contribution of the father's education. This should not be surprising, since the fathers in our sample, who now have adult children, have lower educational levels than their offspring do. Therefore, their occupations are probably better than their education as indicators of their social position for predicting their children's earnings. These results are in line with those obtained by Lefranc and Trannoy (2005) in the decomposition for France and by Österbacka (2001) for Finland. They find that the most important component of the intergenerational correlation in earnings is the correlation between fathers' and children's social positions.

5.4 Quantile regressions

When we regress the children's earnings on their father's earnings we provide a measure of intergenerational mobility at the mean; however, it could be interesting to explore whether this correlation is similar or different at different points of the earnings distribution. If we have homoscedasticity, the coefficient estimated at each percentile will not be statistically different to the coefficient at the mean; however, in the presence of heteroscedasticity, we can obtain different coefficients. After testing the heteroscedasticity with the white test in our sample, we reject the null hypothesis of homoscedasticity.[9] Therefore, as our third exercise it could be interesting to estimate quantile regressions.

Are poor sons and daughters less or more determined by their father's earnings? If low-paid children are more influenced by their father's earnings than children with higher salaries, then the intergenerational elasticity at the mean gives us an incomplete picture of the correlation between fathers' and children's earnings

By estimating quantile regressions, we have a more complete picture of intergenerational transmission of earnings because we have information on the correlation between children's and parents' earnings at different points of the distribution of the children's earnings.[10]

Mean regressions explain how the conditional mean of the children's earnings de-

[9] The results of this exercise are available from the author.

[10] Quantile regression is a statistical technique introduced by Koenker and Bassett (1978) that allows us to estimate conditional functions by quantiles, at different points of the distribution.

Table 9: Intergenerational mobility by quantiles

	Average	10th	25th	50th	75th	90th
sons 30-40	**0.380**	0.428	0.339	0.391	0.356	0.394
	(0.042)	(0.109)	(0.762)	(0.032)	(0.059)	(0.067)
sons 40-50	**0.427**	0.656	0.435	0.468	0.502	0.485
	(0.042)	(0.107)	(0.059)	(0.044)	(0.044)	(0.051)
daughters 30-40	**0.369**	0.813	0.691	0.409	0.346	0.281
	(0.074)	(0.212)	(0.124)	(0.108)	(0.065)	(0.056)
daughters 40-50	**0.498**	0.938	0.864	0.624	0.541	0.410
	(0.062)	(0.177)	(0.064)	(0.067)	(0.081)	(0.069)

Note: Standard error for the estimated coefficients are in parenthesis and are corrected using Murphy and Topel (1985) and Inoue and Solon (2010) procedure. Average refers to mean regression, whereas q-th indicates the q-th percentile regression.

pend on parents' earnings; however, quantile regressions explain how children's earnings depend on parental earnings at each specific quantile of the conditional distribution of the children's earnings, given the fathers' earnings.

In Table 9, we can observe the coefficient of the father's log earnings at different points of the children's earnings distribution. In the first column, we present the mean regression, which shows how important fathers' earnings are on average. In the other columns, quantile regressions evaluate the influence of fathers' earnings at each specific quantile. We consider the 10th, 25th, 50th, 75th and 90th percentiles. We can observe that the influence of fathers' earnings is greater as we move to the poorest quantiles of the distribution. Thus, mobility is lower for the children born in disadvantaged families. This pattern is particularly observed in the case of daughters, where we can observe a monotonic decrease in the elasticity between fathers' and daughters' earnings as we move to the richer percentiles. The results are in-line with those obtained by Nicoletti (2008) for fathers' and daughters' occupations in Britain. For sons, we obtain the highest elasticity at the 10th percentile. Thus, we also observe low mobility for poor sons. However, when we move to richer percentiles the pattern is no longer monotonic, and the coefficients are fairly close to each other and similar to the coefficients in the mean regression.

6 Final remarks

In this paper we examine the intergenerational earnings mobility in Spain with the help of transition matrices, a decomposion of the earnings elasticity and by quantile regressions.

We combine the two samples using the two-sample two-stage least squares estimator in order to correct the co-residence selection problem and we correct the employment selection with the Heckman selection model.

We find an elasticity of 0.38 for sons between 30 and 40 years of age, an elasticity of 0.42 for sons between 40 and 50 years of age. In the case of daughters, we obtain elasticities of 0.36 and 0.49, respectively.

Our transition matrices of education and occupation show a high degree of persistence, especially in the case of occupation. By decomposing the sources of earnings correlations, we find that the correlation between children's and fathers' occupation is the most important component for understanding the intergenerational elasticity between earnings. Furthermore, the correlation between fathers' occupations and their offspring's education is also important. Adding the influence of fathers' occupation on their children's occupation and education, we explain almost half of the intergenerational elasticity coefficient. This should not be surprising since the fathers in our sample, who now have adult children, have a lower educational level than their offspring. So, their occupations are probably better than their education as indicators of their social position for predicting their children's earnings.

Finally, estimating the elasticity between children's and fathers' earnings by quantiles, we find that the influence of the father's earnings is greater when we move to the lower tail of the distribution, especially for daughters' earnings. Thus, mobility is lower for the children born to disadvantaged families.

References

ANGRIST, J. D., AND A. B. KRUEGER (1992): "The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples," *Journal of the American Statistical Association*, 87, 328–336.

ARELLANO, M., AND C. MEGHIR (1992): "Female labour supply and on-the-job search: an empirical model estimated using complementary data set," *The Review of Economic Studies*, 59, 537–559.

BJÖRKLUND, A., AND M. JÄNTTI (1997): "Intergenerational income mobility in Sweden compared to the United State," *American Economic Review*, 87, 1009–1018.

——— (2000): "Intergenerational mobility of socioeconomic status in comparative perspective," *Nordic Journal of Political Economy*, 26(1), 3–32.

BOWLES, S., AND H. GINTIS (2002): "The inheritance of inequality," *Journal of Economic Perspectives*, 16, 3–30.

CARABAÑA, J. (1999): *Dos estudios sobre movilidad intergeneracional.* Fundación Argendaria-Visor (ed.).

CERVINI-PLÁ, M. (2011): "Intergenerational earnings and income mobility in Spain," MPRA Paper 34942, University Library of Munich, Germany.

CHECCHI, D. (2006): *The economics of education: human capital, family background and inequality.* Cambridge University Press, first edn.

DUNN, C. E. (2007): "The Intergenerational Transmission of Lifetime Earnings: Evidence from Brazil," *The B.E. Journal of Economic Analysis & Policy*, 7: Iss. 2 (Contributions), Article 2.

ERIKSON, R., AND J. H. GOLDTHORPE (2002): "Intergenerational inequality: a sociological perspective," *Journal of Economic Perspective*, 16, 31–44.

ERMISCH, J., M. FRANCESCONI, AND T. SIEDLER (2006): "Intergenerational economic mobility and assortative mating," *Economic Journal*, 116, 659–679.

FERREIRA, S. G., AND F. A. VELOSO (2006): "Intergenerational mobility of wages in Brazil," *Brazilian Review of Econometrics*, 6 (2).

FORTIN, N., AND S. LEFEBVRE (1998): "Intergenerational income mobility in Canada," in *Labour Market, Social Institution and the Future of Canada's Children*, ed. by M. Corak. Statistics of Canada, Ottawa.

GRAWE, N. (2004): "Intergenerational mobility for whom? The experience of high- and low-earnings sons in intergenerational perspective," in *Generational Income Mobility in North America and Europe*, ed. by M. Corak. Cambridge University Press, Cambridge.

HAIDER, S., AND G. SOLON (2006): "Life-cycle variation in the association between current and lifetime earnings," *American Economic Review*, 96(4), 1308–1320.

INOUE, A., AND G. SOLON (2010): "Two-sample instrumental variables estimators," *The Review of Economics and Statistics*, 92(3), 557–561.

KOENKER, R. W., AND J. BASSETT, GILBERT (1978): "Regression Quantiles," *Econometrica*, 46(1), 33–50.

Lefranc, A., and A. Trannoy (2005): "Intergenerational earnings mobility in France: Is France more mobile than the U.S.?," *Annales d'Economie et de Statistique*, (78), 03.

Mocetti, S. (2007): "Intergenerational Earnings Mobility in Italy," *The B.E. Journal of Economic Analysis & Policy*, 7: Iss. 2 (Contributions), Article 5.

Murphy, K. M., and R. H. Topel (1985): "Estimation and inference in two-step econometric models," *Journal of Business & Economic Statistics*, 3(4), 370–79.

Nicoletti, C. (2008): "Multiple Sample Selection in the Estimation of Intergenerational Occupational Mobility," ISER working papers 2008-20, Institute for Social and Economic Research.

Nicoletti, C., and J. Ermisch (2007): "Intergenerational earnings mobility: changes across cohorts in Britain," *The B.E. Journal of Economic Analysis and Policy. Contributions*, 7: Iss. 2 (Contributions), Article 9.

Nicoletti, C., and M. Francesconi (2006): "Intergenerational mobility and sample selection in short panels," *Journal of Applied Econometrics*, 21(8), 1265–1293.

Österbacka, E. (2001): "Family background and economic status in Finland," *Scandinavian Journal of Economics*, 103(3), 467–484.

Piraino, P. (2007): "Comparable Estimates of Intergenerational Income Mobility in Italy," *The B.E. Journal of Economic Analysis & Policy)*, 7: Iss. 2 (Contributions), Article 1.

Ridder, G., and R. Moffit (2006): "The econometrics of data combination," in *Handbook of Econometrics*, ed. by Heckman, and Learner, 6. Elsevier Science, North Holland, Amsterdam.

Sanchez-Hugalde, A. (2004): "Movilidad intergeneracional de ingresos y educativa en España (1980-90)," Discussion paper.

Solon, G. (1992): "Intergenerational income mobility in the United States," *American Economic Review*, 82(3), 393–408.

——— (1999): "Intergenerational mobility in the labour market," in *Handbook of Labor Economics*, ed. by O. Ashenfelder, and D. Card, vol. 3, chap. 29, pp. 1761–1800. Amsterdam: Elsevier.

Vella, F. (1998): "Estimating models with sample selection bias: A survey.," *Journal of Human Resources*, 3, 127–169.

Zimmerman, D. (1992): "Regression toward mediocrity in economic stature," *American Economic Review*, 82, 409–429.

Appendix

Table A.1: Distribution of father's education and occupation an coincidences between supplemental and main sample

	supplemental sample	main sample
Observation	5,032	4,352
Education		
No finish primary education	23.82	20.09
Primary education	51.28	57.65
Secondary education (first step)	8.46	6.08
Secondary education (second step)	5.90	5.84
Vocational qualification	2.07	0.49
Higher education (university)	8.47	9.85
Occupation		
Higher grade professionals	9.25	8.04
Higher grade manager	4.28	3.70
Low grade professional	3.43	5.58
Routine non-manual employees high grade	11.04	6.18
Routine non-manual employees low grade	9.85	7.25
Skilled agriculture workers	12.74	12.85
Skilled manual workers	15.88	24.99
Lower-grade technician	13.81	11.82
Unskilled workers	19.71	19.60

Note: All frequencies are weighted using the respective sampling weights.

Table A.2: First step: estimates of father's earnings equation with the supplemental sample

Dependent variable	log father's earnings
age	0.0571 (0.0211)
age square	-0.0006 (0.0002)
Education	
Primary education	0.1873 (0.0148)
Secondary education (first step)	0.3919 (0.0276)
Secondary education (second step)	0.5254 (0.0326)
Vocational qualification	0.5581 (0.0487)
Higher education (university)	0.8455 (0.0281)
Occupation	
Higher grade manager	-0.4381 (0.0404)
Low grade professional	-0.0753 (0.0986)
Routine non-manual employees high grade	-0.0913 (0.0279)
Routine non-manual employees low grade	-0.3158 (0.0320)
Skilled agriculture workers	-0.8155 (0.0306)
Skilled manual workers	-0.1395 (0.0300)
Lower-grade technician	-0.2009 (0.0298)
Unskilled workers	-0.3177 (0.0285)
Constant	11.9961 (0.4918)
Obs	5929
R^2	0.402

Note: standard errors in parentheses. In **Education**: none (reference) and in **Occupation**: Higher-grade professionals (reference).