ISER Working Paper Series

Multiple Sample Selection in the Estimation of Intergenerational Occupational Mobility

Cheti Nicoletti
Institute for Social and Economic Research
University of Essex

No. 2008-20
May 2008

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

www.iser.essex.ac.uk

Non-technical summary

A high level of similarity in the occupational position of parents and children is often thought to be evidence of unequal opportunities between rich and poor resulting in low social mobility. Governments are often encouraged to introduce policies to improve the opportunities of those from disadvantaged backgrounds, to increase social mobility. Therefore, to evaluate whether people's opportunities are unequal it is essential that we have reliable measures of the association between occupational position of parents and children.

In this paper, using the British Household Panel Survey, we estimate the association between occupational prestige of fathers and their daughters, which we call intergenerational association. We measure the occupational prestige by the Hope-Goldthorpe score. This score is strongly related with earnings and provides a measure of the desirability of occupations.

We are especially concerned with the how reliable these measures of intergenerational association are when we are faced with the problem of missing data. Missing data can be caused, for example, by the fact that we do not have information about fathers who have never been interviewed in the survey, and by the fact that that we can not have a measure of occupational prestige for daughters who have never been employed whilst in the survey.

The estimation of the intergenerational association may be biased when we ignore the problem of missing data and focus only on those for whom we have information on both the father and the daughter. To remedy this problem we propose and compare different ways of dealing with this missing data.

The main aim and focus of the paper, therefore, is on the evaluation of different estimation methods to correct for the potential bias caused by missing data. However, we also provide some new results on the intergenerational relationship between fathers and daughters in Britain. Once we have dealt with missing data, we find a correlation of 0.305 between the occupational prestige of fathers and their daughters.

This association is stronger at the bottom of the ladder than at the top, suggesting that a potential increase in the father's occupational prestige is more beneficial for daughters with very low occupational prestige, than for those with very high prestige. This result may suggest that daughters whose fathers have a low occupational prestige have quite similar opportunities to daughters whose fathers occupy high-prestige occupational positions to rise to the top of the occupational prestige scale. On the other hand, there is evidence for a higher inequality in opportunities when looking at daughters whose occupational prestige is especially low.

Multiple Sample Selection in the Estimation of Intergenerational Occupational Mobility

Cheti Nicoletti
ISER, University of Essex

May 2008

Abstract

The estimation of occupational mobility across generations can be biased because of different sample selection issues as, for example, selection into employment. Most empirical papers have either neglected sample selection issues or adopted Heckman-type correction methods. These methods are generally not adequate to estimate intergenerational mobility models. In this paper, we show how to use new methods to estimate linear and quantile intergenerational mobility equations taking account of multiple sample selection.

Keywords: Sample selection; Panel data; Intergenerational mobility; Occupational prestige.

JEL Codes: C23; C24; J24; J62.

Contact: Cheti Nicoletti, Institute for Social and Economic Research, University of Essex, Colchester CO4 3SQ, UK. Email: nicolet@essex.ac.uk.

We would like to thank Marco Francesconi and participants at the LoWER Workshop (Annecy, December 2006) and at the internal seminar series in ISER and in the Department of Economics in the University of Essex. This work was supported by the Economic and Social Research Council through the MiSoC and it is a spin-off from the research project which was financed by the Economic and Social Research Council "Intergenerational mobility in Britain: Does sample selection matter?" Award No. RES-000-23-0185.

1 Introduction

Intergenerational occupational mobility refers to the association between parents' and children's occupational outcomes such as earnings, occupational prestige scores or occupational classes. Measuring the extent of intergenerational mobility is important to verify whether there are inequalities in opportunities between children born in privileged families and children born in disadvantaged families.

Intergenerational occupational mobility studies may be affected by several sample selection issues. In this paper we consider two main selection issues affecting mobility studies based on short panel surveys: the selection into employment and the coresidence selection. The occupational outcome can be observed only for people who are employed, and only for parents and children who are observed to live together in at least one wave of the panel.

Few papers on intergenerational mobility have considered the selection into employment issue (see for example Couch and Lillard 1998, Minicozzi 2003, Ermisch et al 2006, Platt 2005, Fancesconi and Nicoletti 2006, Blanden 2005) and even fewer have considered the coresidence selection one (see Couch and Lillard 1998, Comi 2005, and Francesconi and Nicoletti 2006).

A sample excluding children not in employment probably causes an under-representation of disadvantaged families,[1] while the exclusion of children who were never coresident with their parents during the panel considered could lead to an under-representation of people living parental home very early in their life. More in general the sample selection may depend on specific observed and/or on unobserved variables. Selection on observed variables is presumably a minor problem in models controlling for a large set of explanatory variables. Intergenerational mobility equations usually simply regress children's occupational outcome on their parents' occupational outcome and control only for children's and parents' age. In this context, estimation methods correcting for selection on observables, in particular propensity score weighting estimation (see for example Rosembaum and Rubin 1983, Wooldridge 2002, Hirano et al. 2003, and Wooldridge 2007) can perform better than estimation methods correcting only for selection on unobservables, such as the Heckman-type estimation methods (see for example Vella, 1998).

Selection issues in intergenerational mobility have been either neglected or taken into ac-

[1] This is for example confirmed by the study of O'Neill and Sweetman (1998).

count by using Heckman-type estimation or naive imputation methods, except in Francesconi and Nicoletti (2006) who consider both Heckman type estimation and propensity score methods but only to take account of coresidence selection. In this paper we extend the work of Francesconi and Nicoletti (2006) into two directions.

First, we explicitly model both coresidence and employment selection and propose new estimators to correct for the potential consequent bias. Second, we estimate intergenerational mobility using both mean and quantile regressions. The regression of children's occupational status on their parents' status provides a measure of intergenerational mobility (elasticity) at the mean, while quantile regressions give a more complete picture of intergenerational transmission by providing measures of intergenerational association at different points of the distribution of the children's occupational status.

We focus on the transmission of the occupational prestige from fathers to daughters which has been studied much less extensively than the transmission from fathers to sons. Selection into employment is usually more severe for women, hence the importance to take account of this sample selection issue in the estimation of mobility between fathers' and daughters' occupational outcomes.

The paper is organized as follows. Section 2 describes the coresidence and employment selection problems and summarizes the results found in previous papers which have taken into account those selection issues. Sections 3, 4 and 5 provide theoretical details of estimation methods taking into account sample selection issues when estimating intergenerational elasticity, correlations and persistence at different quantiles. Section 7 gives details on the data and variables used in the empirical analysis, and Section 8 provides the estimation results on the transmission of occupational prestige from fathers to daughters in Britain. Finally, we draw some conclusions in Section 9.

2 Intergenerational occupational mobility and sample selection issues

Intergenerational occupational mobility has attracted a lot of attention in both economic and sociological empirical research. In Britain, for example, Dearden et al (1997), Blanden et al (2004, 2007) and Nicoletti and Ermisch (2007) provide estimates of earnings mobility; Carmichael (2000), Breen and Goldthorpe (2001) and Goldthorpe and Jackson (2007) estimate class mobility; and Ermisch and Francesconi (2004), Ermisch et al (2006) and Francesconi and Nicoletti (2006) look at mobility in occupational prestige.

Economists usually quantify intergenerational occupational (im)mobility by the elasticity of children's occupational status with respect to their parents' one, and measure the occupational status by earnings and occupational prestige scores. On the other hand, sociologists usually consider association measures between occupational classes.[2]

Following the economic approach, we focus in this article on intergenerational occupational (im)mobility measured by the intergenerational elasticity of children's occupational status with respect to their parents one. We restrict our attention on the occupational mobility between fathers and daughters and we measure their occupational status by the Hope-Goldthorpe index, henceforth HG index, which is a continuous measure of occupational prestige.[3] More precisely, we consider the following intergenerational mobility equation:

$$y = \alpha + \beta x + A\gamma + u \tag{1}$$

where y is the daughters log occupational prestige; x is her fathers log occupational prestige; A is a vector of other control variables, specifically the daughters' age and age square; α is the intercept term representing the average change in the children log occupational prestige, β and γ are coefficients; and u is a random error. The coefficient β is the intergenerational elasticity of daughters' occupational prestige with respect to their fathers' one, and it is our parameter of interest.

Notice that β can be alternatively computed by considering the following equation:

$$\tilde{y} = a + \beta\tilde{x} + \epsilon \tag{2}$$

where $\tilde{k}$ is the residual of the regression of k on A, $k = y$ or x, a is a new intercept and ϵ is a new error term. Let ρ be the correlation between $\tilde{y}$ and $\tilde{x}$; then β is related to ρ by the following equation:

$$\beta = \rho\frac{\sigma_{\tilde{x}}}{\sigma_{\tilde{y}}} \tag{3}$$

where $\sigma^2_{\tilde{k}}$ is the variance of $\tilde{k}$, $k = y$ or x. The elasticity coefficient β is therefore related to the correlation between daughters' and their fathers' log occupational prestige net of the control variables A. Moreover, β is exactly equal to ρ when $\sigma_{\tilde{x}} = \sigma_{\tilde{y}}$.

[2] We refer to Solon (1999), Björklund and Jäntti (2000), Bowles and Gintis (2002), Erikson and Goldthorpe (2002) for a review.

[3] The HG index is a continuous measure of occupational prestige computed according to the technique proposed by Goldthorpe and Hope (1974) and related strongly to labour income.

A coefficient β equal to zero indicates a situation where all daughters have equal opportunities. When $\beta = 0$ all daughters have an expected log occupational prestige equal to α plus an additional deterministic component function of their age. When β is instead different from zero, daughters expected occupational prestige depends also on their fathers occupational prestige.

When $\sigma_{\tilde{x}} \neq \sigma_{\tilde{y}}$, β and ρ provide different values for the intergenerational transmission. Changes of β across time and countries may be related to changes in the correlation ρ and/or changes in $\sigma_{\tilde{x}}$ and $\sigma_{\tilde{y}}$, whereas changes in the correlation ρ are invariant to changes in the variance of the marginal distribution of $\tilde{x}$ and $\tilde{y}$. To provide a measure of intergenerational mobility which is more comparable across countries and time, we also estimate the intergenerational correlation.

Intergenerational correlation and elasticity measures are good methods to summarize the association between parents' and children's occupational status by using a single summary statistics. If $\tilde{y}$ and $\tilde{x}$ are distributed as normal, then the correlation ρ summarizes perfectly all we can know about the intergenerational association. More in general, for non-normal distributions, ρ and β captures only the linear relationship between $\tilde{y}$ and $\tilde{x}$.

A more complete picture of intergenerational mobility can be provided by looking at quantile regressions.[4] In this way we can allow the intergenerational association to change at different quantiles of the children's occupational prestige. While the mean regression (2) explains how the conditional mean of y depends on x, the quantile regressions explain how y depends on x at each specific quantile of the conditional distribution of y given x. If, for example, the effect of having an increase in the father's log occupational prestige is more beneficial for daughters with low occupational prestige than for daughters with high occupational prestige, then the intergenerational elasticity does provide only a partial information on how the intergenerational transmission operates. For this reason, we consider both mean and quantile regressions.

The estimation of intergenerational correlation, mean and quantile regressions may be biased when using samples of matched father-daughter extracted from a short panel. These samples are unlikely to be a random sample because affected by at least two different sources of sample selection:

[4] We refer to Eider and Showalter 1999, Grawe 2004, Bratberg et al 2007, and Bratsberg et al 2007, for previous empirical studies considering quantile regressions to allow for a non-linear relationship between children's and parents' socio-economic status.

a. father-daughter coresidence;

b. daughter employment selection.

In the following two subsections we describe these two selection issues and methods which have been adopted in previous papers to correct for the consequent potential biases.

2.1 Employment selection

Intergenerational occupational mobility can be observed only for people who are employed and this may cause a selection bias. Most of intergenerational occupational mobility studies neglect this selection problem and exclude all records of data where parents or children are not in employment and thus implicitly assume exogenous selection into employment. This assumption is not consistent with standard economic results according to which selection into the labour force or into employment is likely to be correlated with potential earnings (see Heckman 1979, and Vella 1998). There are few intergenerational mobility studies taking account of employment selection, especially for chidren, see for example Couch and Lillard (1998), Minicozzi (2003), Francesconi and Nicoletti (2006), Blanden (2005), and Ermisch et al. (2006).

Blanden (2005) and Ermisch et al. (2006) consider the employment selection problem for daughters by using two-step estimation procedures, a Heckman-type correction method. In the first step they estimate the employment selection model and in the second step they estimate the intergenerational mobility equation by adding a correction term which is a specific function of the selection probability - a cubic polynomial expression proportional to the inverse Mill's ratio.[5] Ermisch et al. (2006) estimate intergenerational mobility in Britain and Germany using the two national longitudinal household surveys, the BHPS (British Household Panel Study) and the GSOEP (German Socio Economic Panel). They find that intergenerational mobility results do not change much whether considering or neglecting the employment selection problem. Blanden (2005) uses instead the British Cohort Studies 1958 and 1970 and find a lower intergenerational correlation when taking into account the employment selection issue.

The problem with Heckman-type selection correction methods is that they correct for selection on unobserved variables but do not take account of potential selection on observed

[5] For a more detailed description of the estimation methods see Vella (1998).

variables. In particular, if there are explanatory variables excluded from the intergenerational equation which affect both children's occupational outcome and the selection into employment probability, then Heckman-type correction methods are inconsistent. Considering for example an intergenerational model regressing children' log earnings on their parents' log earnings and an employment selection equation explained by educational level, it is unlikely that the education level be irrelevant in explaining children's log earnings in the intergenerational equation (see, for example, Mincer 1974, Heckman *et al.* 2003).

A different type of approach to take account of the employment selection is by imputing the missing occupational outcome (earnings or occupational prestige scores). Such type of approach has been followed by Couch and Lillard (1998), who study intergenerational earnings mobility in USA using the PSID (Panel Study of Income Dynamics), and Francesconi and Nicoletti (2006), who study instead intergenerational occupational prestige mobility in Britain using the BHPS. Couch and Lillard (1998) impute one dollar income to unemployed people. Francesconi and Nicoletti (2006), similarly, assuming that missing values are for people with low permanent income, replace the missing scores with the minimum occupational prestige score observed for the respondent people. Assuming instead that the missing values refer to people who are randomly spread over the entire distribution of occupational prestige, they replace them with the median observed for the respondent people. Both Couch and Lillard (1998) and Francesconi and Nicoletti (2006) find that more selected samples lead to higher correlations between children' and parents' occupational measures.

Unfortunately, the above imputation methods are very simplified, do not use auxiliary variables and possibly impose too restrictive assumptions.

Minicozzi (2003) relaxes the assumptions imposed by the above methods, but at the cost of failing to produce a point estimate for the intergenerational elasticity coefficient. By using the PSID, as in Couch and Lillard (1998), and adopting the partial identification approach (see Manski 2003), she produces a lower and upper bound for the intergenerational elasticity in earnings. Unfortunately such type of approach produces usually quite large bounds which are not very useful to make inference.

In this paper, we introduce some new methods to correct for selection into employment for children. In particular, we consider weighting and regression adjustment, which correct for selection on observables and do not impose the assumption that all variables used to predict the selection probability be irrelevant in the intergenerational equation (exclusion restriction assumption). Regression adjustment methods relax this exclusion restriction by enlarging the set of explanatory variables considered in the intergenerational equation; whereas the weighting method relaxes the same assumption by using weights which are given by the inverse of the probability of selection. Moreover we combine those two methods, which take account of selection on observables, with methods taking account of selection on unobservables (Heckman-type estimators).

2.2 Coresidence selection

When estimating intergenerational occupational mobility using short household panel, there is an additional sample selection problem which must be taken in consideration. It is possible to observe the occupational status for both children and their parents only for children living with parents in at least one wave of the panel. Following Francesconi and Nicoletti (2006) we call this condition coresidence selection.

Intergenerational mobility studies are usually based on panel surveys which have been running for many years, say more than 20 years, such as the Panel Study of Income Dynamics in the USA or the Socio-economic Panel in Germany. In those surveys the coresidence selection is not a problem because it is easy to observe young children living together with their parents and follow them to adulthood to measure their occupational status, except for possible attrition and employment selection problems. On the contrary, the coresidence selection may be an issue for the estimation of intergenerational mobility using more recent household panel surveys such as the ones in Belgium, Hungary, Luxembourg, Poland and the UK, or panel surveys that have been dismissed too early such as the European Community Household Panel (ECHP).[6] Let us consider for example a 10–year–long panel, then a child 35 years old in the last wave of the panel is 25 years old at the beginning of the panel and by that age she has probably already left her parental home. Considering instead a 25–year–long panel, a child 35 years old in the last wave is 10 years old at the beginning of the panel and at that age she is probably still living with her parents.

The coresidence selection magnitude is sharpened by the fact that we need to observe children occupational status when adult. If the panel is very short then adult children in the last wave may be too old in the first wave to be observed living with their parents. If the subsample of children observed living together with their parents is not random, then

[6]See Comi (2004) and Francesconi and Nicoletti (2006) for two empirical applications using short panels.

the coresidence selection may cause a bias in the intergenerational mobility estimation.

Even with relative long panels, the coresidence selection could become a problem when estimating differences in intergenerational mobility between children born in different periods. Considering as before a 25–year–long panel, if we want to compare intergenerational mobility of children 35 years old in the last wave of the panel with children born 15 years earlier; we observe the former at 10 years old and the latter at 25 years old in the first year of the panel. While the father-child coresidence selection is probably not an issue for the former, it can be an issue for the latter.

Francesconi and Nicoletti (2006) assess the extent of the coresidence selection bias and compare different types of methods to estimate the intergenerational mobility using the British Household Panel Survey (BHPS) and measuring the long-run permanent economic status of sons and their parents by the occupation prestige HG (Hope-Goldthorpe score). In the BHPS it is possible to observe the Hope-Goldthorpe score, HG, for all respondents and their parents regardless of the coresidence condition. This is because all respondents are asked to report the occupation of their parents when they were aged 14. Using the sample of father-son pairs derived from the BHPS Francesconi and Nicoletti (2006) estimate intergenerational elasticities in occupational prestige presumably free from coresidence selection bias. By linking sons to their fathers over the first eleven years of the panel and imposing a standard coresidence condition, they obtain a new selected subsample of father-son pairs, say the restricted sample. Comparing the elasticities estimated using restricted and full samples, they conclude that the coresidence selection causes an overestimation of the intergenerational mobility. They then evaluate two approaches that correct for the coresidence selection bias. The first belongs to the general class of Heckman-type correction methods (see Vella 1998); more precisely they use (1) joint maximum likelihood estimation of the regression equation together with the selection equation, (2) two-step Heckman (1979) estimation, (3) estimation of the intergenerational regression with an additional correction term given by a quadratic polynomial expression proportional to the inverse Mill's ratio, and (4) estimation of the intergenerational equation with dummy variables to control for different level of the selection probability. The second approach is within the class of models based on propensity score weighting estimation (see Rosenbaum and Rubin 1983 and Wooldridge 2002 and 2007). They find that the all estimation methods are unable to correct for the coresidence bias except the weighting method which performs well in most circumstances.

They investigate the reasons for the bad performance of the econometric selection correction methods by testing the validity of the assumptions imposed by different estimators. They find that the exclusion restriction imposed by Heckman-type correction methods is rejected. It is worth noticing that the exclusion restrictions are likely to be rejected, not only for specific empirical example in Francesconi and Nicoletti (2006), but also for all intergenerational mobility studies where the main equation of interest does not consider potential variables affecting both the coresidence selection probability and the children's economic status.

In this paper we again adopt the weighting estimator used in Francesconi and Nicoletti (2006) but we also consider an estimation method combining weights and Heckman-type correction methods in an attempt to control for both selection on observables and unobservables. Regression adjustment estimation methods cannot be used in the case of coresidence selection. This is because they require that the father's occupational status be observable for the whole sample (see Section 3). In the case of coresidence selection father's occupational status is observable only for the subsample of children coresident with their fathers in at least one wave of the panel.

In the empirical application, we first implement the estimation methods separately for the employment and coresidence selection issues, while later we consider the two issues together and we estimate a joint model for the probability of employment and coresidence.

3 Intergenerational elasticity with sample selection

Let us assume to observe a random sample of n father-daughter pairs indexed by i and for whom we observe age and log occupational prestige scores, y for daughters and x for fathers. Then we can estimate the intergenerational occupational mobility equation introduced in Section 2,

$$y_i = \alpha + \beta x_i + A_i \gamma + u_i, \qquad i = 1, \dots, n \tag{4}$$

by simply applying ordinary least squares (OLS) method. Notice that we are not requiring that u_i be independent of x_i. This last assumption is likely to be false because there are omitted variables such as work, cognitive and non-cognitive skills of the daughter, which are related to both daughter's and her father's occupational prestige. This is not a problem as long as we interpret β as a measure of association between x_i and y_i which captures both the direct effect of x and its indirect effect through the omitted variables. Assuming that this total effect is our parameter of interest then the OLS estimation is consistent.

On the contrary, when y_i and/or x_i are missing for some of the father-daughter pairs we have a sample selection problem which can cause a bias of the OLS estimation. Let r_i be a dummy variable, the selection dummy, taking value one if both y_i and x_i are observed and zero otherwise.[7] Let us begin considering the presence of a single selection problem, selection into employment or coresidence selection, and postpone to Section 6 the joint treatment of the two selection problems. In the presence of only coresidence selection, $r_i = 1$ if daughter i is resident with her father at least in one year so that we can observe her father's occupational prestige x_i, and $r_i = 0$ if the daughter is never observed living together with her father and therefore x_i is missing. In the presence of only selection into employment, $r_i = 1$ if daughter i is employed at least in one year and we can observe her occupational prestige y_i and $r_i = 0$ otherwise.

Let us assume that r_i obeys the following latent index model:

$$r_i^* = m(Z_i\theta) + v_i, \tag{5}$$

where r_i^* is a latent variable linked to the selection dummy through the following indicator function $r_i = I(r_i^* > 0)$; Z_i is a vector of explanatory variables (possibly including A_i, but excluding x_i in the case of coresidence selection) that are assumed to be observed for all individuals; v_i is an error term identically and independently distributed with mean zero and unit variance; and θ is a conformable vector of parameters. The estimation of this selection model is possible because both r_i and Z_i are observed for all father-daughter pairs. This allows us to estimate the propensity score,

$$\pi_i = Pr(r_i = 1 \,|\, Z_i) = Pr(v_i > -m(Z_i\theta)),$$

which is used to correct for the potential sample selection bias in some of the methods described below.

Using the subsample of father-daughter pairs with $r_i = 1$ produces a biased estimation of the intergenerational mobility equation (4) when the condition $(y_i \perp\!\!\!\perp r_i \,|\, x_i, A_i)$ is invalid.[8] This last condition fails when u_i depends on v_i (selection on unobservables) and/or when

[7] We assume here that the control variables in A_i are always observed for all father-daughter pairs.

[8] The symbol $\perp\!\!\!\perp$ stands for independence.

u_i depends on Z_i (selection on observables). More in general, we say that data are missing completely at random (MCAR) when r_i does not depend on any of the observable or unobservable variables, whereas we say that data are missing at random (MAR) when r_i depend on observable variables but it is independent on unobservable ones .

In the following of this section we describe various methods to control for a univariate sample selection problem, while in Section 6 we consider how to correct jointly for the two sample selection problems.

Selection on unobservables can be controlled by adopting a Heckman-type of correction estimation (see Vella 1998). Under the assumption that u_i and v_i are independently and identically distributed $N(0, \Sigma)$, with Σ being a two by two full-rank variance-covariance matrix, and (u_i, v_i) independent of Z_i and x_i, it is straightforward to estimate all the parameters in (4) and (5) maximizing the censored likelihood or by the two-step estimation of Heckman (1974, 1979). Alternative two-step estimation methods relaxing the joint normality assumption for the error terms are provided in Newey (1999), Robinson (1988), Powell (1989), Cosslett (1991) and Das et al (2005). All those methods are based on the assumption that

$$E(u_i \,|\, x_i, A_i, Z_i, r_i = 1) = h(\pi_i)$$

where $h(\pi_i)$ is an unknown function of the propensity score π_i.

If $(u_i \perp\!\!\!\perp x_i, A_i, Z_i)$ then

$$u_i \perp\!\!\!\perp x_i, A_i, Z_i \,|\, r_i, \pi_i$$

is satisfied if and only if

$$r_i \perp\!\!\!\perp x_i, A_i, Z_i \,|\, u_i, \pi_i$$

(the proof follows immediately from Proposition 2 in Angrist, 1997). Since any monotonic or more specifically any latent index selection model is such that

$$r_i \perp\!\!\!\perp x_i, A_i, Z_i \,|\, u_i, \pi_i$$

(see Proposition 3 in Angrist, 1997), the condition

$$u_i \perp\!\!\!\perp x_i, A_i, Z_i \,|\, r_i, \pi_i$$

is satisfied whenever $(u_i \perp\!\!\!\perp x_i, A_i, Z_i)$. This implies that, if $(u_i \perp\!\!\!\perp x_i, A_i, Z_i)$ and the selection model (5) is correctly specified, we can rewrite the linear regression model (4) in the presence of selection, $r_i = 1$, as:

$$y_i = \alpha + \beta x_i + A_i\gamma + g(\pi_i) + \epsilon_i, \tag{6}$$

where ϵ_i is independent of all explanatory variables and has zero mean. Newey (1999), Robinson (1988), Powell (1989), Cosslett (1991) and Das et al (2005) propose different way to approximate the unknown function $g(\pi_i)$. More precisely, they propose a two-step control function procedure. In the first step they estimate the selection model, while in the second step they estimate the linear regression (6) controlling for the unknown function $g(\pi_i)$ by using different types of approximations. These methods are called *control function* estimation methods or *Heckman-type* estimation methods (see Vella 1998 for more details).

The control function methods assume that u_i be independent of Z_i. This assumption is restrictive when considering parsimonious intergenerational mobility models as the one in (4). This is because it is likely that some of the explanatory variables included in the selection model and excluded from the intergenerational mobility equation (4) are relevant to explain children's occupational status, and this would contradict the independence between u_i and Z_i. In other words, selection on observables cannot be neglected in intergenerational mobility models. This explains the bad performance of Heckman-type correction methods in intergenerational mobility studies (see Francesconi and Nicoletti, 2006, and empirical result in Section 8).

In the absence of selection on unobservables, selection on observables can be controlled by using either propensity score weighting or regression adjustment methods. Let us suppose that the selection on observables is due to the failure of $(u_i \perp\!\!\!\perp r_i \,|\, x_i, A_i)$ because u_i depends on Z_i. Then we can control for the dependence of u_i on Z_i by extending the intergenerational mobility equation to

$$y_i = \alpha_N + \beta_N x_i + A_i\gamma_N + Z_i\delta + \omega_i, \tag{7}$$

where α_N, β_N and γ_N are new coefficients corresponding to α, β and γ but net of the effect of the additional explanatory variables Z_i.

If the linearity assumption imposed by the new model is satisfied and there is no selection on observables, then $\omega_i \perp\!\!\!\perp r_i \,|\, x_i, A_i, Z_i$ and the OLS estimation of the equation (7) based on the selected sample provides a consistent estimation of β_N. We call this estimation method *regression adjustment estimation.*

It is easy to recover the parameter of interest β from β_N in the following way:

$$\beta = Cov(\tilde{x}, \tilde{y})/Var(\tilde{x}) = \beta_N + Cov(\tilde{x}, \tilde{Z})Var(\tilde{Z})^{-1}\delta, \tag{8}$$

where $\tilde{k}$ is the residual of the regression of k on A, $k = y$, x or Z.[9] Notice that the estimation of β requires that x and Z are observed for the entire sample. This is the case when we consider the problem of employment selection for daughters, while it is not the case when we consider the problem of coresidence selection which implies a missing x for daughters who cannot be matched with their fathers. This is the reason why we will not apply regression adjustment estimation to correct for the coresidence selection problem.

An alternative approach to control for selection on observables is by adopting propensity score methods (see for example Rosembaum and Rubin 1983, Hirano et al. 2003, Wooldridge 2002 and 2007). In this paper we consider the *weighting method* which consists in estimating the intergenerational equation by weighted least squares with weights given by the inverse of the propensity score.

A more robust and new estimation method to take account of selection on observables can be obtained by combining the regression adjustment and the weighting methods. As shown by Robins and Rotnitzky (1995), this *combined regression adjustment/weighting method* has the advantage of being double-consistent, meaning that it is consistent if either the weighting method and/or the regression adjustment method are consistent.

If there is selection on both observables and unobservables, then the combination of a control function method with the weighting estimation should reduce the selection bias. Notice, although, that the *combined weighting/control function* method as well as the control function method require the use of proper instrumental variables restrictions. In other words, the empirical identification of the parameters require that some explanatory variables relevant to explain the selection model be excluded from the intergenerational mobility equation. In our empirical example we will use as instrumental variables house price indexes and religiosity in the coresidence selection model and the presence of children younger than five years in the employment selection model.

[9] The regression adjustment method consists basically in the application of the standard omitted variable bias formula in linear regression models. The regression adjustment method is also equivalent to the method used by Bowles and Gintis (2002) to decompose the intergenerational elasticity or correlation in different additional terms representing the direct effect of the parents' socio economic status and the indirect effects through specific factors (IQ, schooling, race, wealth and personality).

4 Intergenerational correlation with sample selection

The intergenerational correlation measured by the parameter ρ introduced in Section 2 can be derived from the intergenerational elasticity coefficient in the following way:

$$\rho = \beta \frac{\sigma_{\tilde{y}}}{\sigma_{\tilde{x}}} \tag{9}$$

where $\sigma^2_{\tilde{k}}$ is the variance of $\tilde{k}$, and $\tilde{k}$ is the residual of the regression of k on A, $k = y$ or x. Alternatively, ρ can be estimated by computing the correlation

$$\rho = E(\dot{y}\dot{x}),$$

where $\dot{y}$ and $\dot{x}$ are obtained by normalizing the variables $\tilde{y}$ and $\tilde{x}$ subtracting their mean and dividing by their standard deviation.

When $\dot{y}$ and $\dot{x}$ are observed only for a subsample of individuals with the dummy variable $r = 1$ then the estimation of ρ may be biased. But, under the assumption that $(y, x \perp\!\!\!\perp r \,|\, Z)$, we can estimate consistently ρ by using weights which are given by the inverse of the propensity score,

$$\pi = Pr(r = 1 \,|\, Z) = Pr(v > -m(Z\theta)).$$

This is because

$$\begin{aligned} E\left(\dot{y}\dot{x}\frac{r}{\pi}\right) &= E_Z E\left(\dot{y}\dot{x}\frac{r}{\pi} \,|\, X, Z\right) = E_Z\left[E\left(\dot{y}\dot{x}\frac{1}{\pi} \,|\, X, Z, r = 1\right) Pr(r = 1 \,|\, X, Z)\right] \\ &= E_Z\left[E\left(\dot{y}\dot{x} \,|\, X, Z\right) Pr(r = 1 \,|\, X, Z)\frac{1}{\pi}\right] = E_Z\left[E\left(\dot{y}\dot{x} \,|\, X, Z\right)\right] = E(\dot{y}\dot{x}). \end{aligned}$$

5 Intergenerational quantile regression with sample selection

In this section we consider a linear quantile regression, that is we assume that

$$Pr(y_i \leq y_q \,|\, x_i, A_i,) = F_{u_q}(y_q - \alpha - x_i\beta_q - A_i\gamma_q \,|\, x_i, A_i), \tag{10}$$

where $Quant_q(y_i \,|\, x_i, A_i) = (\alpha + x_i\beta_q + A_i\gamma_q) = y_q$ is the conditional quantile of y_i given the explanatory variables (x_i, A_i) and F_{u_q} is the cumulative distribution of the error term u_q

whose q-quantile is equal to zero, $Quant_q(u_{q_i}) = 0$. The above relation can be rewritten as the following linear equation:

$$y_i = \alpha + x_i\beta_q + A_i\gamma_q + u_{q_i}. \tag{11}$$

Given these assumptions we can estimate consistently the q-quantile regression and the parameter of interest β_q by solving the following optimization problem (see for example Koenker and Bassett 1978, Buchinski 1998, and Koenker 2005)

$$\min_{\alpha_q,\beta_q,\gamma_q} \frac{1}{n}\left\{\sum_{i:u_{q_i}\geq 0} q|u_{q_i}| + \sum_{i:u_{q_i}<0}(1-q)|u_{q_i}|\right\} \tag{12}$$

where $u_{q_i} = y_i - \alpha_q - x_i\beta_q - A_i\gamma_q$, or equivalently by solving,

$$\min_{\alpha_q,\beta_q,\gamma_q} \frac{1}{n}\left\{\sum_i [q - I(y_i \leq \alpha_q + x_i\beta_q + A_i\gamma_q)](y_i - \alpha_q - x_i\beta_q - A_i\gamma_q)\right\}, \tag{13}$$

where $I(.)$ is the indicatrice function of the event between parenthesis. The first order condition (F.O.C.) for the above minimization problem is given by:

$$\frac{1}{n}\sum_i [q - I(y_i \leq \alpha_q + x_i\beta_q + A_i\gamma_q)][x_i, A_i] = \mathbf{0}, \tag{14}$$

where $\mathbf{0}$ if a row vector of zeros.

In the presence of missing data the F.O.C. applied to the subsample of units with no missing data is:

$$\frac{1}{n}\sum_i r_i[q - I(y_i \leq \alpha_q + x_i\beta_q + A_i\gamma_q)][x_i, A_i] = \mathbf{0}, \tag{15}$$

where r_i is a dummy taking value one if the are no missing data for unit i and zero otherwise. If r_i depends on observed and/or unobserved variables, then the solution to the minimization problem does not necessarily provide a consistent estimation of the parameters of interest.

When r_i depends only on a set of observed variables Z_i a possible solution is given by weighting the F.O.C. by the inverse of the propensity score

$$\frac{1}{n}\sum_i \frac{r_i}{\pi_i}[q - I(y_i \leq \alpha_q + x_i\beta_q + A_i\gamma_q)][x_i, A_i] = \mathbf{0}, \tag{16}$$

where $\pi_i = Pr(r_i = 1 \,|\, Z_i)$. This type of weighting procedure, which we call *weighting quantile estimation*, belongs to the more general methods of the weighted estimating equation introduced by Robins and Rotnitzky (1995) or the weighted M-estimation described in

Wooldrige (2007), and it provides consistent estimation of the parameters under the assumption of MAR.[10]

Assuming, as in last section, a latent index model for the selection process, that is assuming that r_i obeys the following selection model:

$$r_i^* = m(Z_i\theta) + v_i, \tag{17}$$

then the propensity score π_i is given by $\pi_i = Pr(r_i = 1 \,|\, Z_i, \theta) = Pr(v_i > -m(Z_i\theta))$. Moreover, if we assume a linear mean regression model as in last section,

$$y_i = \alpha + \beta x_i + A_i\gamma + u_i, \tag{18}$$

then $u_{q_i} = \alpha - \alpha_q + x_i(\beta - \beta_q) + A_i(\gamma - \gamma_q) + u_i$.

If u_{q_i} depends on v_i (selection on unobservables), then the conditional quantile is given by

$$Quant_q(y_i \,|\, x_i, A_i, r_i = 1) = \alpha_q + x_i\beta_q + A_i\gamma_q + Quant_q(u_{q_i} \,|\, x_i, A_i, r_i = 1), \tag{19}$$

where $Quant_q(u_{q_i} \,|\, x_i, A_i, r_i = 1)$ is not in general equal to zero. If $(u_i \perp\!\!\!\perp x_i, A_i, Z_i)$, then $(u_i \perp\!\!\!\perp x_i, A_i, Z_i \,|\, \pi_i, r_i = 1)$ and

$$Quant_q(y_i \,|\, x_i, A_i, r_i = 1) = \alpha_q + x_i\beta_q + A_i\gamma_q + h_q(\pi_i), \tag{20}$$

where $h(.)$ is an unknown function of the propensity score. Then we can rewrite the linear equation (11) as:

$$y_i = \alpha_q + x_i\beta_q + A_i\gamma_q + h_q(\pi_i) + \epsilon_{q_i}, \tag{21}$$

where $Quant(\epsilon_{q_i} \,|\, Z_i, r_i = 1) = 0$. A way to control for selection on unobservable is then by extending the control function estimation methods shown in last section as suggested in Buchinski (2001). The estimation proceeds in two steps: the first step consists in the estimation of the selection model, while the second step is a quantile regression estimation where the unknown function $h_q(\pi_i)$ is approximated through, for example, a polynomial or a step function in the propensity score π_i. Other types of approximation are possible and we refer for more details to Buchinski (2001). We call this two-step estimation *control function method* for quantile regression.

[10] Lipsitz et al. (1997) provide an application of this method to quantile regression, while Ibrahim et al. (2005) provide a more recent review of this method and a comparison with other estimation approaches.

If Z_i contains variables relevant to explain y_i, then u_{q_i} depends on Z_i and the control function procedure does not necessarily produce consistent estimation. In that case, the combination of the weighting quantile estimation with the control function procedure can be a solution to control for both selection on observables and unobservables. We call this last type of estimation combined weighting/control fucntion estimation.

6 Multiple selection model

In the last three sections we presented different types of estimation taking account of selection on observables and/or unobservables but we focused only on univariate selection models. In this section we extend these estimation procedures to the case of bivariate selection models. More precisely we extend the weighting and control function estimation procedures for both quantile and mean linear regression models .

Let r_{1_i} and r_{2_i} be two dummy variables taking value one if the first and the second type of selection rules are satisfied and zero otherwise, and let $p_i = p_i(Z_i) = Pr(r_{1_i} = 1, r_{2_i} = 1 \,|\, Z_i)$ be the joint propensity score, i.e. the joint probability of selection conditional on a set of relevant explanatory variables Z_i. Then the extension of weighting procedures are very straightforward and still consists of two steps. In the first step we estimate a bivariate binary model for the double selection. In the second step we consider a weighted least squares estimation in the case of linear regression model (4) and a weighted F.O.C. in the case of quantile regression (11), where the weights are given by the inverse of the propensity score estimated in the first step.

The extension of the control function estimation when we consider the linear regression model (4) in the presence of double selection requires the following assumption:

$$E(u_i \,|\, x_i, A_i, Z_i, r_{1_i} = 1, r_{2_i} = 1) = g(p_i)$$

where $g(.)$ is an unknown function of the joint propensity score. If $u_i \perp\!\!\!\perp x_i, A_i, Z_i$ and the double selection model is a bivariate latent index model, then the condition

$$u_i \perp\!\!\!\perp x_i, A_i, Z_i \,|\, r_{1_i}, r_{2_i}, p_i$$

is satisfied and therefore

$$E(u_i \,|\, x_i, A_i, Z_i, r_{1_i} = 1, r_{2_i} = 1) = g(p_i)$$

(the proof is a simple extension of Proposition 2 in Angrist, 1997, to the case of a multivariate selection process). Under this last condition we can rewrite the mean regression (4) in the presence of double selection as:

$$y_i = \alpha + x_i\beta + A_i\gamma + g(p_i) + \eta_i \tag{22}$$

where η_i is an error term with zero mean. To control for the unknown function $g(p_i)$ we can approximate it by considering a polynomial in p_i or a non-parametric step function. Similar control function methods have been proposed also by Das et al. (2003) and De Luca and Peracchi (2006) but only in the case of linear regressions.

When considering the quantile regression (11) the extension of the control function method to the case of double selection requires that

$$Quant(u_{q_i} \,|\, x_i, A_i, Z_i, r_{1_i} = 1, r_{2_i} = 1) = g_q(p_i),$$

where $g_q(p_i)$ is an unknown function of p_i. Again this condition is satisfied when the linear equation (4) holds, the selection model obeys a bivariate latent index model, and $u_i \perp\!\!\!\perp x_i, A_i, Z_i$. Under these conditions we can rewrite (11) given the double selection as:

$$y_i = \alpha_q + x_i\beta_q + A_i\gamma_q + g_q(p_i) + \xi_{q_i}, \tag{23}$$

where $Quant(\xi_{q_i} \,|\, x_i, A_i, Z_i, r_i = 1) = 0$ and $g_q(p_i)$ can be approximated using a polynomial in p_i or a non-parametric step function.

7 Data

Our estimation will be produced using data from the first thirteen waves of the British Household Panel Survey (BHPS) collected over the period 1991-2003.[11]

Since Autumn 1991 the BHPS has annually interviewed a representative sample of about 5,500 households covering more than 10,000 individuals. All adults and children in the first wave are designated as original sample members. On-going representativeness of the non-immigrant population has been maintained by using a following rule typical of household panel surveys: at the second and subsequent waves, all original sample members are followed (even if they moved house or if their households split up).

[11] See Taylor (2003) for a full description of the dataset. Detailed information on the BHPS can also be obtained at <http://www.iser.essex.ac.uk/bhps/doc>.

Personal interviews are collected, at approximately one-year intervals, for all adult members of all households containing either an original sample member, or an individual born to an original sample member. Individuals are defined as adult (and are therefore interviewed) from their sixteenth birthday onwards. The sample therefore remains broadly representative of the population of Britain as it changes over time. The households from the European Community Household Panel subsample (followed since the seventh wave in 1997), those from the Scotland and Wales booster subsamples (added to the BHPS in the ninth wave) and those from the Northern Ireland booster subsample (which started in wave 11) are excluded from our analysis.

We now turn to describe samples and variables selected from the BHPS.

7.1 Samples

Our main analysis is restricted to 2164 women (daughters) born between 1966 and 1985, who have at least one valid interview over the panel period under study. This represents our Full Sample. The BHPS asks all adult respondents aged 16 or more to provide information about their parents' occupations when they (the respondents) were aged 14, and releases data on an index of occupational prestige introduced by Goldthorpe and Hope (1974). This index called Hope-Goldthorpe (HG) ranges from 5 to 95, with greater values indicating higher occupational prestige, and it is highly correlated with earnings.[12] Between 1991 and 2003, the BHPS data indicate a correlation between gross monthly earnings and the Hope-Goldthorpe (HG) index of 0.70 for men and 0.75 for women. Because the position of individuals in the occupational hierarchy is relatively stable over the life cycle, the HG scale is also likely to be an adequate measure of people's permanent socio-economic status (Nickell, 1982). Ermisch et al. (2006) and Nicoletti and Ermisch (2007) show that using occupational prestige scores of fathers and offspring as a proxy of permanent income produces similar results to those using average earnings data for Germany and Britain. Of course, different correlations between occupational score measures and average earnings may arise in countries other than Germany and Britain.

All the individuals in the Full Sample who could be successfully matched to their father are part of our second sample, which we refer to as Restricted Sample. There are 646 of such

[12]Phelps Brown (1977) reports a strong log-linear relationship between median gross weekly earnings and the HG score, with a rise of 1 unit in the index being associated with an increase of 1.03 percent in earnings. Nickell (1982) finds a correlation between the HG score and the average hourly earnings of 0.85.

father-daughter pairs. Differently from the Full Sample, this imposes stringent coresidence conditions. Individuals born in 1966 were aged 25 in the first year of the panel (1991) and 38 in the last (2003): they could have lived with their parents at any age between those two years. With a median home-leaving age of about 23-24 (Ermisch and Di Salvo, 1997; Ermisch and Francesconi, 2000), coresidence at such ages means that the Restricted Sample overrepresents sons who left home at late ages. At the other extreme, individuals born in 1985 were aged 18 in 2003 (the last year of analysis): although they are likely to be a random sample of young people living with their parents, their HG index is arguably a noisy measure of long-run status.

The comparison of the intergenerational elasticities obtained from the Full Sample and the Restricted Sample will provide us with a measure of the extent of the selection in short panels, under the maintained assumption that Full Sample estimates do not suffer from any selection bias. But both Full and Restricted samples are affected by selection into employment. Out of the 2164 daughters in the Full Sample only 1931 are observed into employment, whereas 601 out of the 646 daughters in the Restricted Sample are employed. We refer to the sub-sample of employed daughters in the Full Sample as Employed Full Sample and to the sub-sample of employed daughters in the Restricted Sample as Employed Restricted Sample. The comparison of the intergenerational elasticities obtained from the Employed Full Sample and the Employed Restricted Sample taking account of the employment selection provides us with a measure of the selection bias in short panel associated with the coresidence selection.
.

Admittedly there are also other sample selection issues besides the coresidence and the employment selection ones. In particular, each of the variables used in the estimation of the intergenerational mobility can be missing. Nevertheless, the size of this item non-response is tiny (well below 10%) for all variables used in our analysis, but for the father's occupational prestige (which is collected retrospectively asking to daughters to report their fathers' occupation). This variables is missing in about 50% of the cases. In an attempt to understand whether this is another selection issue which we should take into account, we consider the Restricted sample where the fathers' occupational prestige is known for almost all fathers and derived from questions asked directly to the fathers. We find that the probability of non-response on father's occupational prestige collected retrospectively depends neither on the value of the father's occupational prestige collected directly through questions asked to the fathers, nor on the values of other observed variables such as the education level of the daughters, their religion and their ethnicity. Given this result we decided to ignore this further selection issue in our analysis.

7.2 Variables

Theoretically, we would like to measure intergenerational occupational mobility by considering long-run permanent occupational prestige, but we observe current occupational prestige at a specific age. For this reason we measure the occupational status of daughters by their average Hope Goldthorpe (labelled HG^d) over all waves after excluding the cases with missing status information either because the daughter does not work or because his information is genuinely missing. Several studies have argued that averaging status over time reduces the impact of the transitory component of the status variable (thus reducing the potential of errors-in-variables bias) and yields a more accurate measure of permanent status (see, among others, Solon 1992, Zimmerman 1992, Dearden et al. 1997, and Mazumder, 2005).

Table 1 shows that the mean values of HG^d is about 43 for daughters in the Full Sample and 42 for daughters in the Restricted Sample.

One of the major difficulties in estimating intergenerational elasticities abides in the fact that father's status is measured with error. The key problem is the lack of direct measures of permanent status. In the case of the Full Sample, in particular, the BHPS provides us with only one single-year measure of fathers' occupational prestige (when daughters were aged 14).[13] Although the HG index is an arguably good proxy for long run status, a single-year measure may still be tainted by transitory fluctuations in fathers' careers. In addition, the BHPS elicits this information by asking respondents to report their parents' occupation when they were aged 14. The retrospective questioning of children to obtain data on parents may of course generate recall errors.[14] Both types of errors (due to measurement and recall) may be such that the variance of observed status is greater than the variance of permanent status, leading the OLS estimate of β in (1) to be biased downward.

Table 1 reports mean and standard deviation for the daughters' and the fathers' occupa-

[13] In the case of the Restricted Sample, instead, multiple measures of fathers' HG scores are available in principle. For that sample, however, there are still measurement problems in that fathers and daughters are observed at different points in their life cycle (see below). Clearly, the Restricted Sample is expected to suffer from the selection bias discussed in 2.

[14] Again, our Restricted Sample does not have to face this problem as fathers and daughters report independently their own occupational information.

tional prestige scores, HG^s and HG^f. These statistics do not change significantly between the Full and the Restricted sample.

Table 1 lists also the summary statistics of the other variables used in the analysis. As in several other studies, the intergenerational model includes daughter's age and its square (Solon, 1992; Zimmerman, 1992; Dearden et al., 1997; Couch and Dunn, 1997; Corak and Heisz, 1999). The mean age of daughters in the Full Sample is 23.5, almost 2 years greater than in the Restricted Sample. We do not have information for fathers' age for the entire Full Sample, but in the Restricted Sample their mean age is 46.2. A number of other variables are used to model the probability of observing fathers and daughters living in the same household at least once during the sample period and the probability of daughters to be employed. The mean and the standard deviations of these variables are also reported in Table 1.

The set of explanatory variables used in both selection models are: age and dummies for ethnicity, region and education level. More precisely we consider five ethnic groups (White, Black, Indian, Pakistani/Bangladeshi, Other), eight regional dichotomous variables for the standard regions of Great Britain,[15] and five levels of education (first degree or above, A level or equivalent, GSCE/O level or equivalent, other qualification, and no qualification). Table 1 shows that distributions of the above variables dummies are similar across samples.

Furthermore, we consider the presence of children in the employment selection model, and dummies for religious attendance and house prices in the coresidence selection model. More precisely we consider the number of children between zero and two years old and between three and fours years old, which are found to be important in explaining womens labour decisions and, by consequence, their probability to be employed.

Religious attendance is a factor that is believed to have a deep effect on the likelihood of young people's leaving parental home (Cherlin 1992). Although religious views on family formation are varied, strong religious beliefs are one cultural source of ideas that encourages the maintenance of traditional values (Wilcox and Wolfinger 2007). For each of the three religious denominations considered here (Catholic, Protestant, and other religions), "attendance" is defined as attending religious services at least once a month.[16] In both samples,

[15]The regions are Greater London (which is our base category), South East, South West, East Anglia and East and West Midlands, North West (including Yorkshire and Humberside), Rest of the North, Wales, and Scotland.

[16]The "Protestant" group includes: Church of England (Anglican), Church of Scotland, Free Presbyterian, Episcopalian, Methodist, Baptist, Congregationalist, and other Christian denominations. The "other reli-

about 20 percent of the young women are religiously active, more than 40 percent have no religious affiliation and the rest have an affiliation but do not attend services regularly.

Many studies of household formation have underlined the importance of the price of housing (e.g., Haurin et al., 1994; Ermisch, 1999). House prices indeed can affect the likelihood of observing fathers and daughters living together, and so they may determine the selection into the Restricted Sample. The price of housing is an ambiguous concept when there are different housing tenures, non-neutral tax treatment of them, and probable imperfections in financial markets (Ermisch and Di Salvo, 1997; Ermisch, 1999). In Britain in 2002, nearly 90 percent of households are either owner-occupiers (68 percent) or "social tenants" (22 percent). The latter primarily includes households who rent their dwelling from local authorities. Social housing is not allocated by price, but by administrative procedures, which give priority to families with children and the elderly. While only a small proportion of all households rent from private landlords, it is a relatively important sector for young people leaving their parental home, being the destination of 45 percent of all departures and 33 percent of departures among those who are not full-time students. Owner-occupation is the destination for 56 percent of non-student departures. Information on rents in the private market is not available in the BHPS, and there are barriers to entry into social rental housing for young people. We use two measures for the price of housing. The first house price index is given by average annual house prices from 1991-2003 provided by Halifax Housing Research and aggregated at level of Local Authority District.[17] The second house price index is given instead by the average "mix-adjusted" house price relative to the retail price index in any given year for the region in which a person resided in that year.[18] It adjusts for changes in the mix of the size and type of house (e.g., detached, semi-detached, flat, etc.), but does not adjust for quality change.

Both measures are likely to capture a large proportion of the variation in a measure of the annual "user cost of housing" for owner-occupiers, because mortgage and income tax rates are set nationally and relative house prices show much larger variation over time than these. They also could be viewed as an indicator of housing market conditions, in both rental and owner-occupied markets. For individuals in the Full Sample who could not be matched

gions" include: Muslim, Hindu, Jewish, Sikh, and other non-Christian denominations. The omitted category includes those with no religious affiliation as well as those who have a religious affiliation but attend religious services only infrequently. Distinguishing between such two groups does not change our results.

[17]This price index is the same used by Rabe (2006).

[18]It is the same measure used by Ermisch (1999).

with their fathers, the price of housing refers to the price observed in the first wave they were in the panel. For those who coreside with their fathers (and therefore, all daughters in the Restricted Sample), this variable is measured at the last wave they were observed living together. In Table 1, the averages for the two (log) house price indexes in the Full and Restricted Samples are similar. These averages mask large differences across Local Authority District (for the first measure) or regions (for the second measure) the and over time.

8 Empirical Results

In this section we look at how the occupational prestige differs across daughters whose fathers occupy different positions in the occupational prestige scale. Mean regression results inform us on how much this difference is at the mean, whereas quantile regressions results evaluate that difference at each specified quantile. We consider the 10th, 25th, 50th, 75th and 90th percentiles. The difference at the mean and at the 5 different percentiles are summarized by the coefficients of the fathers' log occupational prestige, that is β, β_{10}, β_{25}, β_{50}, β_{75}, β_{90}, which we report in the Tables 2–4. Moreover we report the correlation between daughters' and fathers' log occupational prestige.

In Tables 2 and 3 we consider the coresidence selection and the employment selection issues separately, while in Table 4 the two issues are considered jointly.

In each of the three tables we report the estimates obtained ignoring the selection problem (i.e. the estimates of the simple correlation (4), the OLS estimates of the linear regression (4) and the solution to the optimization problem (12) for the five quantile regressions) followed by the results of different estimation methods which take into consideration the sample selection: weighting, control function, and combined weighting/control function methods. Moreover, when considering only the employment selection issue, we apply the regression adjustment and the combined weighiting/regression adjustment too. Those two methods cannot be applied when the coresidence selection is present because they require the observation of the fathers' occupational prestige for all father-daughter pairs.

The estimation methods correcting for the potential selection biases consist of two steps: (1) the estimation of the sample selection model, (2) the estimation of the intergenerational transmission taking account of the selection issue using the propensity score estimates from the first step. When correcting for employment and coresidence selection separately, we model the selection process as a univariate probit model; while when jointly modeling the two selection issues, we consider a bivariate probit model. The choice to adopt a parametric instead of a semiparametric specification (as in Robinson 1988, Powell 1989, Cosslett 1991, Das et al. 2003, and De Luca and Peracchi 2006) for the selection models can be criticized, but it is supported by results of the normality test for the error in the selection models.[19] Moreover, the identification in semiparametric latent index models would require the presence of at least one continuous explanatory variable, which we do not have in the case of employment selection (see Li and Racine 2007 for more details of identification conditions).

The tables 5 and 6 report the estimation results of the univariate probit models and of the bivariate probit model for the coresidence and employments selection. There seems to be a slight correlation between the selection and the coresidence process, which is not statistically significant. The most significant variables in the coresidence selection are the region of residence, the house price index and the dummy for no education qualification. People who were resident at the beginning of the panel or before leaving parental home in the Greater London or in the South East are more likely to be observed coresident with their parents than people leaving in the remaining regions. The first house price index, which is at level of local authority district, has a negative impact on the coresidence probability; while the second one, which is measured at level of region, has a positive effect. This could be interpreted as a higher likelihood to leave parental home early for daughters leaving in regions where the house price index is low (houses are more affordable), and especially so if they are leaving in local authority districts with a house price index relatively higher than in the rest of the region. Finally daughters with no qualification have a significant lower probability to be observed living together with their father. The main determinants of employment selection are ethnicity (black, Indian and Pakistani/Bangladeshi daughters are less likely to be in employment than white daughters); the presence of children and the dummy for no qualification, which have a negative effect; and the dummy for being married or cohabiting, which has a positive effect.

The estimated coefficients of the probit models are used to predict the propensity score. In the weighting estimation we use the inverse of predicted propensity score as weight. In the control function estimation we estimate the intergenerational mean or quantile regres-

[19]We verify the normality assumption imposed by the probit selection model by modifying the score test proposed in Machin and Stewart (1990) for ordered probit models, which in turn modifies the score test for a grouped dependent variable introduced by Chesher and Irish (1987). When the selection model includes Z, we find that normality is never rejected at standard levels of significance.

sion with additional variables (correction terms) given by a polynomial of order three in the estimated propensity score. In the weighting, the control function and the combined weighting/control function estimation methods we adopt bootstrapping techniques to compute the standard errors, which otherwise would be underestimated. The regression adjustment approach does not require the estimation of the propensity score, therefore does not require adjustment of the standard errors.

Table 2 compares the intergenerational mobility when using the full sample (where only selection into employment is present) and the restricted sample (where there are both coresidence and employment selection issues). The difference between the full sample and the restricted sample estimates reflects the bias caused by the coresidence selection and this bias is negative for the intergenerational correlation as well as for the intergenerational transmission estimated using mean and quantile regressions. The weighting estimation seems to correct the bias in the right direction; while the control function does not seem to perform well, and especially so for the average regressions and the 10th, 25th and 50th percentile regressions.

Since in the case of coresidence selection we are able to recover the missing information on the father's occupational prestige (through a retrospective question asked to each daughter about her father's occupation when she was 14), we can test whether the assumptions imposed by the different types of estimators used are rejected.

In Table 7 we report the assumptions under which each specific estimator would be consistent even in the presence of coresidence selection. Furthermore, we report a test of each assumption and its p-value in the last two columns of Table 7. The relationship between daughters' and their fathers' log occupational prestige changes significantly between the Full sample (unaffected by coresidence) and the Restricted sample (contaminated by coresidence selection). The equality of the coefficients in the intergenerational mobility between the two samples is strongly rejected (see the test result for the assumption A1 in Table 7). Assumptions A2 and A4 can be tested by using the Full sample and verifying whether x and (y, x) are significantly different form zero in the probit model for the coresidence selection with explanatory variables given by (Z,A,X,y). Both assumptions are not rejected at the standard level of significance. Not surprisingly, the explanatory variables used to explain the coresidence probability (and more in particular the education dummies) are significantly different from zero when included in the intergenerational model, and in consequence the condition A3 is rejected at both 1% and 5% significance levels (see the test reported in Table 7) .

In Table 3, we consider the Full sample, which is not biased by coresidence selection, and we report the results of different estimation methods to correct for the employment selection. More precisely we consider the weighting, the control function, and the combined weighting/control function. Moreover, for the mean regression we consider also the regression adjustment and the combined weighting/regression adjustment methods.[20] The differences between the estimates for these correction methods and the estimates computed ignoring the employment selection problem are very small. The reason for the low difference is probably a consequence of the low percentage of women who are never in employment in our sample, about 10%. When considering the control function estimation, the estimates of the intergenerational immobility are slightly lower than when ignoring selection into employment in line with what Blanden (2005) finds, but the difference is not significant in line with what Ermisch et al (2006) find. Anyway, we would suggest not to adopt the control function method when considering very parsimonious intergenerational mobility models. This method imposes the very implausible assumption that daughter's occupational prestige is independent of all explanatory variables used in the employment selection given the father's occupational prestige.

In Table 4, we finally report the estimates computed using the restricted sample and correcting for both employment and coresidence selection. In this case the first step of the weighting and of the control function methods is the joint estimation of employment and coresidence selection by using a bivariate probit model. While the control function methods produce estimates very close to the one computed ignoring both selection issues, the weighting and the combined weighting/control function methods suggest that the ignoring selection causes a quite large underestimation of the intergenerational transmission coefficients.

Our preferred estimation is the weighting method applied to the Full sample. This preference is justified by two reasons: (1) the Full sample is not biased by coresidence selection, (2) the weighing method does not assume that the error term in the intergenerational equation be independent of x and/or Z. A preference for the combined weighting/control function method could be also justified on the ground that it takes account of both selection on observ-

[20]Notice that the regression adjustment or the combined weighting/regression adjustment methods cannot be applied to the quantile regressions because the omitted variable formula, which we use to recover the β_q coefficients, is not directly applicable to quantile regressions.

ables and unobservables. Nevertheless, this last method imposes an independence condition between u and x given Z. Given that there will be always unobservables, such as variables describing genetic and cultural endowment, which are probably correlated with both u and x given Z, we prefer the weighting method which imposes an independence condition between y and r given Z and x.

The results of this estimation (see Table 3) suggest that intergenerational correlation and elasticities at the mean and at the 25th, 50th, 75th percentiles are quite close and varies between 0.294 and 0.327. The intergenerational persistence seems instead different at the 10th percentile (0.357) and at the 90th percentile (0.222), and it is weaker at the top than at the bottom of the distribution. Furthermore, by allowing the slope coefficient β to change across different levels of the fathers' occupational prestige for each of 5 quantile regressions, we find that β does not change significantly. Summarizing, it seems that a log linear relationship between occupational prestige of fathers and daughters be adequate to represent the intergenerational relationship except maybe at the extremes of the distribution. This result is in line with Bratsberg et al (2007) who find that linear and non-linear regressions provide similar results for the transmission of earnings from fathers to sons when using the UK National Child Development Study. But, while Bratsberg et al (2007) find that intergenerational association is stronger for sons with exceptionally high prestige than for sons with exceptionally low prestige, we find the opposite for daughters.

9 Conclusion

Intergenerational mobility equations are usually very parsimonious models which do not consider all the potential variables relevant to explain both sample selection probability and children's socio-economic status. This implies that instrumental exclusion restrictions imposed by Heckman-type correction methods are unrealistic. This is confirmed by our analysis of the coresidence selection issue, for which we can compare a sample with and a sample without coresidence selection. The best methods to correct for the coresidence selection are the weighting estimation and the combined weighting/control function estimation.

In our empirical analysis we find also that the coresidence selection issue is large both in terms of size and bias, while the employment selection issue seems negligible. Our preferred estimates seem to suggest that intergenerational persistence is stronger at the bottom of the daughters' occupational prestige distribution than at the top. Except for the extremes of the distribution, a log linear relationship between occupational prestige of fathers and daughters seem to provide an adequate description of the intergenerational transmission.

References

Angrist, J.D. (1997) Conditional independence in sample selection models, *Economics Letters*, 54, 103–112.

Atkinson, A.F., Maynard, A.K., Trinder, C.G. (1983) *Parents and Children: Incomes in Two Generations*, London: Heinemann.

Becker, G.S., Tomes, N. (1979) An equilibrium theory of the distribution of income and intergenerational mobility, *Journal of Political Economy*, 87, 1153–89.

Becker, G.S., Tomes, N. (1986) Human capital and the rise and fall of families, *Journal of Labor Economics*, 4, S1–S39.

Björklund, A., Jäntti, M. (2000) Intergenerational mobility of socioeconomic status in comparative prospective, *Nordic Journal of Political Economy*, 26, 1, 3-32.

Blanden, J. (2005) Intergenerational mobility and assortative mating in the UK, Chapter 6 of Essays in *Intergenerational Mobility and its Variation over Time, Place and Family Structure*, PhD Thesis. University of London.

Blanden, J., Goodman, A., Gregg, P., Machin, S. (2004) Changes in intergenerational mobility in Britain, in M. Corak (ed.) *Generational Income Mobility in North America and Europe*, 122–46. Cambridge: Cambridge University Press.

Blanden, J., Gregg, P., Macmillan, L. (2007) Accounting for intergenerational income persistence: noncognitive skills, ability and education, *The Economic Journal*, 197, 43-60.

Bowles, S., Gintis, H. (2002) The inheritance of inequality, *Journal of Economic Perspectives*, 16, 3–30.

Bratberg, E., Nilsen, Ø.A., Vaage K.(2007), Trends in intergenerational mobility across offspring's earnings distribution in Norway, *Industrial Relations*, 46,1, 112–129.

Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T., Österbacka, E. (2007), Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons, *Economic Journal*, 117, 519, C72 C92.

Breen, R., Goldthorpe, J.H. (2001) Class mobility and merit: the experience of two British birth cohorts, *European Sociological Review* 17, 81–101.

Buchinski, M. (1998) Recent advances in quantile regression models: A practical guide for empirical research, *The Journal of Human Resources*, 33, 1, 88–126.

Buchinski, M. (2001) Quantile regression with sample selection: Estimation women return to education in the U.S., *Empirical Economics*, 26, 86–113.

Carmichael, F. (2000), Intergenerational mobility and occupational status in Britain, *Applied Economic Letters*, 7, 391-396.

Cherlin, A.J. (1992) *Marriage, Divorce and Remarriage*, Cambridge, MA: Harvard University Press.

Chesher, A., Irish, M. (1987) Residual analysis in the grouped and censored normal linear model, *Journal of Econometrics*, 34, 33–61.

Comi, S. (2003), Intergenerational mobility in Europe: evidence from ECHP, Dipartimento di Economia Politica e Aziendale, Università degli Studi di Milano, Working paper 2003-3.

Corak, M., Heisz, A. (1999) The intergenerational earnings and income mobility of Canadian men: evidence from longitudinal income tax data, *Journal of Human Resources*, 34, 504–33.

Cosslett, S. (1991) Semiparametric estimation of a regression model with sample selectivity, In W.A. Barnett, J. Powell, and G. Tauchem (eds.) *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, 175–97. Cambridge: Cambridge University Press.

Couch, K.A., Dunn, T.A. (1997) Intergenerational correlations in labor market status: a comparison of the United States and Germany, *Journal of Human Resources*, 32, 210–32.

Couch, K.A., Lillard, D.R. (1998) Sample selection rules and the intergenerational correlation of earnings, *Labour Economics*, 5, 313–29.

Das, M., Newey, W.K., Vella, F. (2003) Nonoparametric estimation of sample selection models, *Review of Economic Studies*, 70, 3358.

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

De Luca G., Peracchi F. (2007) A sample selection model for unit and item nonresponse in cross-sectional surveys, CEIS Tor Vergata, Research Paper Series, Working Paper 99.

Eide, E.R., Showalter, M.H. (1999) Factors affecting the transmission of earnings across generations: A quantile regression approach, *Journal of Human Resources*, 34, 2, 253-67.

Erikson, R., Goldthorpe, J.H. (2002) Intergenerational inequality: a sociological perspective, *Journal of Economic Perspectives*, 16, 31–44.

Ermisch, J. (1999) Prices, parents, and young people's household formation, *Journal of Urban Economics*, 45, 47–71.

Ermisch, J., Di Salvo, P. (1997) The economic determinants of young people's household formation, *Economica*, 64, 627–44.

Ermisch, J., Francesconi, M. (2000) Patterns of household and family formation, In R. Berthoud and J. Gershuny (eds.) *Seven Years in the Lives of British Families*, 21–44. Bristol: Policy Press.

Ermisch, J., Francesconi, M. (2004) Intergenerational mobility in Britain: new evidence from the BHPS, In M. Corak (ed.) *Generational Income Mobility in North America and Europe*, 147-89. Cambridge: Cambridge University Press.

Ermisch, J., Francesconi, M., Siedler, T. (2006) Intergenerational economic mobility and assortative mating, *Economic Journal*, 116, 659-679.

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

Goldthorpe, J.H., Hope, K. (1974), *The Social Grading of Occupations*. Oxford: Clarendon Press.

Goldthorpe, J.H., Jackson, M. (2007), Intergenerational class mobility in contemporary Britain: political concerns and empirical findings, *British Journal of Sociology*, 58, 4, 525–546.

Grawe, N.D. (2004), Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints, *Journal of Human Resources*, 39, 3, 813-27.

Haurin, D.R., Hendershott, P.H., Kim, D. (1994) Housing decisions of American youth, *Journal of Urban Economics*, 35, 28–44.

Heckman, J. (1974) Shadow prices, market wages, and labor supply, *Econometrica*, 42, 679–693.

Heckman, J. (1979) Sample selection as a specification error, *Econometrica*, 47, 153–61.

Heckman, J., Lochner, L., Todd, P.E. (2003), Fifty years of Mincer earnings regressions, NBER working paper 9372.

Hirano, K., Imbens, G., Ridder, G. (2003) Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71, 1307-1338.

Ibrahim, J.G., Chen, M.-H., Lipsitz, S.R., Herring, A.H. (2005) Missing-data methods for generalized linear models: A comparative review, *Journal of the American Statistical Association*, 100, 469, 332–346.

Koenker, R., Bassett, G. (1978) Regression quantiles, *Economtrica*, 46, 1, 33–50.

Koenker, R. (2005), *Quantile Regression*, Econometric Society Monograph Series, Cambridge University Press.

Li, Q., Racine, J.S. (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, Princeton.

Machin, S.J., Stewart, M.B. (1990) Union and financial performance of british private sector establishment, *Journal of Applied Econometrics*, 5, 327–50.

Lipsitz, S.R., Fitzmaurice, G.M., Molenberghs, G., Zhao, L.P. (1997), Quantile regression methods for longitudinal data with drop-outs, *Applied Statistics*, 46, 463-476.

Machin, S.J., Stewart, M.B. (1990) Union and Financial Performance of British Private Sector Establishment, *Journal of Applied Econometrics*, 5, 327–50.

Mazumder, B. (2005) Fortunate sons: new estimates of intergenerational mobility in the United States using social security earnings data, *Review of Economics and Statistics*, 87, 235–55.

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

Minicozzi, A.L. (2003) Estimation of sons' intergenerational earnings mobility in the presence of censoring, *Journal of Applied Econometrics*, 18, 291–314.

Newey, W.K. (1999) Two step series estimator of Heckman selection model, Department of Economics, Massachussets Institute of Technology, mimeo.

Nickell, S. (1982), The determinants of the occupational success in Britain", *Review of Economic Studies*, 49, 45–53.

Nicoletti, C., Ermisch, J. (2007), Intergenerational earnings mobility: changes across cohorts in Britain", *The B.E. Journal of Economic Analysis & Policy, Contributions*, 7, 1-36.

O'Neill, D., Sweetman O. (1998), Intergenerational mobility in Britain: evidence from unemployment patterns, *Oxford Bulletin of Economics and Statistics*, 60, 4, 431–449.

Phelps Brown, H. (1977) *The Inequality of Pay.* New York: Oxford University Press.

Platt, L. (2005), Mobility and missing data: What difference does non-response make to observed patterns of intergenerational class mobility by ethnic group?, ISER working paper 2005-10, Colchester: University of Essex.

Robins, J., Rotnitzky, A. (1995) Semiparametric efficiency in multivariate regression models with missing data, *Journal of the American Statistical Association*, 90, 122–29.

Rabe, B. (2006) Dual-Earner migration in Britain: earnings gains, employment, and self-selection, Working Paper of Institute for Social and Economic Research, paper 2006-01, Colchester: University of Essex.

Powell, J.L. (1989), Semiparametric estimation of censored selection models, Department of Economics, University of Wisconsin-Madison.

Robins, J. M., Rotnitzky, A. (1995), Semiparametric Efficiency in Multivariate Regression Models With Missing Data, *Journal of the American Statistical Association*, 90, 122129.

Robinson, P.(1988), Root-N-consistent semiparametric regression, *Econometrica* , 56, 931–954.

Rosenbaum, P.R., Rubin, D.B. (1983), The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70, 41–55.

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

Solon, G. (1999) Intergenerational mobility in the labor market, In O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Volume 3, Chapter 29, 1761–1800. Amsterdam: Elsevier.

Taylor, M.F. (ed.) with Brice, J., Buck, N., Prentice-Lane, E. (2002) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Vella, F. (1998) Estimating models with sample selection bias: A survey, *Journal of Human Resources*, 3, 127–69.

Wilcox, W.B., Wolfinger N. (2007) Then comes marriage? Religion, race, and marriage in urban America, *Social Science Research*, 36, 569–589.

Wooldridge, J.M. (2002) Inverse probability weighted M-Estimators for sample selection, attrition, and stratification, *Portuguese Economic Journal*, 1, 117–39.

Wooldridge, J.M. (2007) Inverse probability weighted M-Estimation for General missing data problems, *Journal of Econometrics*, 141, 2, 1281–1301.

Zimmerman, D.J. (1992) Regression toward mediocrity in economic stature, *American Economic Review*, 82, 409–29.

Table 1: Summary statistics by sample

Variables	Full Sample Mean (S.D.)	N	Restricted Sample Mean (S.D.)	N
Hope-Golthorpe scores (daughters)				
HG^d	43.960 (13.358)	1931	42.751 (11.018)	601
Hope-Golthorpe score (fathers)				
HG_1^f	48.308 (15.041)	1180	48.231 (14.821)	307
Other characteristics				
Father's age			46.508 (7.323)	595
1st House price index (log)	11.090 (0.280)	2164	11.160 (0.334)	646
2nd House price index (log)	11.091 (0.235)	2164	11.027 (0.262)	646
Daughter's characteristics				
Age	24.374 (4.207)	2164	22.384 (3.630)	646
Ethnic origin:				
White (base)	0.947	2029	0.959	611
Black	0.017	2029	0.013	611
Indian	0.011	2029	0.011	611
Pakistani/Bangladeshi	0.011	2029	0.011	611
Other	0.013	2029	0.005	611
Religious attendance:				
Protestant	0.111	2162	0.115	645
Catholic	0.058	2162	0.064	645
Other denomination	0.017	2162	0.020	645
Region of daughters' residence:				
Greater London (base)	0.127	2164	0.125	646
Rest of South East	0.177	2164	0.176	646
South West	0.079	2164	0.084	646
Anglia and Midlands	0.222	2164	0.228	646
North West	0.093	2164	0.108	646
Rest of North	0.154	2164	0.146	646
Wales	0.051	2164	0.050	646
Scotland	0.159	2164	0.084	646
No. children 0-2	0.159 (0.387)	2164	0.159 (0.379)	646
No. children 3-4	0.139 (0.366)	2164	0.128 (0.348)	646
Education dummies:				
First degree or above	0.227	2164	0.217	646
A level or equivalent	0.406	2164	0.402	646
GSCE/O level or equivalent	0.205	2164	0.240	646
Other qualification	0.094	2164	0.096	646
No qualification	0.059	2164	0.029	646

Note : N indicates the number of daughters in the Full and Restricted Samples for whom the specific variable is observed.

Table 2: Intergenerational mobility in the full and restricted samples ignoring employment selection and correcting for coresidence selection

	Correlation	Average	10th	25th	50th	75th	90th
				Full sample			
Ignoring selection	0.303	0.285	0.447	0.310	0.312	0.246	0.084
	(0.000)	(0.028)	(0.050)	(0.048)	(0.057)	(0.037)	(0.032)
				Restricted sample			
Ignoring selection	0.218	0.156	0.146	0.205	0.141	0.130	0.003
	(0.000)	(0.045)	(0.131)	(0.069)	(0.069)	(0.070)	(0.065)
Weighting	0.300	0.236	0.374	0.413	0.285	0.133	0.109
	(0.000)	(0.097)	(0.151)	(0.129)	(0.137)	(0.106)	(0.087)
Control function		0.157	0.171	0.214	0.148	0.132	0.062
		(0.050)	(0.107)	(0.064)	(0.066)	(0.067)	(0.056)
Weighting/Control f.		0.243	0.491	0.480	0.260	0.150	0.109
		(0.099)	(0.156)	(0.146)	(0.156)	(0.134)	(0.096)

Note : Standard errors for the estimated coefficients are in parenthesis. Average refers to mean regression, whereas q-th indicates the q-th percentile regression.

Table 3: Intergenerational mobility in the presence of only employment selection. Full sample

	Correlation	Average	10th	25th	50th	75th	90th
Ignoring selection	0.303	0.285	0.447	0.310	0.312	0.246	0.084
	(0.000)	(0.028)	(0.050)	(0.048)	(0.057)	(0.037)	(0.032)
				Correcting for employment selection			
Weighting	0.305	0.301	0.357	0.327	0.333	0.294	0.222
	(0.000)	(0.032)	(0.058)	(0.047)	(0.050)	(0.039)	(0.041)
Control function		0.224	0.301	0.229	0.233	0.204	0.080
		(0.029)	(0.062)	(0.043)	(0.046)	(0.033)	(0.028)
Weighting/Control f.		0.230	0.337	0.258	0.254	0.268	0.185
		(0.030)	(0.064)	(0.065)	(0.054)	(0.046)	(0.056)
Reg. adjustment		0.288					
		(0.028)					
Weighting/Reg. Adj.		0.294					
		(0.030)					

Note : Standard errors for the estimated coefficients are in parenthesis. Average refers to mean regression, whereas q-th indicates the q-th percentile regression.

Table 4: Intergenerational mobility in the presence of both employment and coresidence selection. Restricted sample

	Correlation	Average	10th	25th	50th	75th	90th
Ignoring selection	0.218	0.156	0.146	0.205	0.141	0.130	0.003
	(0.000)	(0.045)	(0.111)	(0.069)	(0.069)	(0.070)	(0.065)
Correcting for both sample selection							
Weighting	0.288	0.231	0.396	0.404	0.273	0.129	0.104
	(0.000)	(0.094)	(0.136)	(0.120)	(0.133)	(0.110)	(0.086)
Control function		0.158	0.140	0.213	0.164	0.129	0.006
		(0.048)	(0.119)	(0.084)	(0.071)	(0.065)	(0.081)
Weighting/Control f.		0.242	0.517	0.464	0.283	0.144	0.113
		(0.095)	(0.187)	(0.143)	(0.150)	(0.093)	(0.084)

Note : Standard errors for the estimated coefficients are in parenthesis. Average refers to mean regression, whereas q-th indicates the q-th percentile regression.

Table 5: Univariate probit models for coresidence and employment selection

	Coresidence		Employment	
Variable	Coeff	S.E.	Coeff	S.E.
No qualification	-0.684	(0.391)	-1.274	(0.234)
Other qualification	0.209	(0.180)	-0.473	(0.220)
GSCE/O level or equivalent	0.142	(0.131)	-0.312	(0.173)
First degree or above	0.072	(0.129)	0.004	(0.206)
Age before leaving home	0.040	(0.114)		
Age2	0.004	(0.003)		
log house price index 1	-4.074	(0.431)		
log house price index 2	1.503	(0.269)		
Rest of South East	-0.458	(0.201)	0.260	(0.265)
South East	-0.590	(0.248)	0.275	(0.320)
Anglia and Midlands	-1.285	(0.254)	0.159	(0.241)
North West	-1.324	(0.274)	0.378	(0.314)
Rest of North	-1.656	(0.270)	-0.202	(0.248)
Wales	-1.449	(0.310)	0.558	(0.395)
Scotland	-1.586	(0.278)	-0.130	(0.280)
Black	-0.294	(0.522)	-5.953	(2.150)
Indian	0.182	(0.629)	-1.537	(0.461)
Pakistani/Bangladeshi	0.423	(0.704)	-1.252	(0.521)
Other ethnicity	0.239	(0.508)	-0.419	(0.513)
Catholique	-0.186	(0.214)		
Protestant	0.149	(0.149)		
Other denominations	0.079	(0.369)		
Father's Log HG			0.044	(0.203)
Average age			0.562	(0.155)
Average age square			-0.010	(0.003)
Married			0.673	(0.172)
No. of children aged 0-2			-0.685	(0.149)
No. of children aged 3-4			-0.504	(0.161)
No. of children aged 5-11			-0.253	(0.101)
No. of children aged 12-15			-0.392	(0.143)
No. of children aged 16-18			-0.250	(0.443)
Constant	26.389	(4.533)	-5.953	(2.150)
No. Obs.	1090.000		1090.000	
Log-likelihood	-421.168		-228.072	
Pseudo R^2	0.333		0.265	

Table 6: Bivariate probit model for coresidence and employment selection

	Coresidence		Employment	
Variable	Coeff	S.E.	Coeff	S.E.
No qualification	-0.771	(0.327)	-1.269	(0.232)
Other qualification	0.229	(0.171)	-0.469	(0.219)
GSCE/O level or equivalent	0.125	(0.126)	-0.307	(0.172)
First degree or above	0.056	(0.127)	0.015	(0.205)
Age before leaving home	0.043	(0.108)		
Age2	0.004	(0.003)		
log house price index 1	-4.139	(0.422)		
log house price index 2	1.454	(0.260)		
Rest of South East	-0.438	(0.195)	0.278	(0.264)
South East	-0.617	(0.243)	0.267	(0.315)
Anglia and Midlands	-1.359	(0.250)	0.170	(0.238)
North West	-1.320	(0.267)	0.375	(0.310)
Rest of North	-1.657	(0.262)	-0.203	(0.244)
Wales	-1.510	(0.306)	0.575	(0.394)
Scotland	-1.638	(0.272)	-0.122	(0.278)
Black	-0.275	(0.513)	-6.229	(2.005)
Indian	0.363	(0.518)	-0.250	(0.573)
Pakistani/Bangladeshi	0.246	(0.654)	-1.543	(0.455)
Other ethnicity	0.056	(0.483)	-1.214	(0.521)
Catholique	-0.209	(0.212)	-0.400	0.512
Protestant	0.115	(0.146)		
Other denominations	-0.009	(0.358)		
Father's Log HG				
Average age			0.583	(0.156)
Average age square			-0.010	(0.003)
Married			0.674	(0.172)
No. of children aged 0-2			-0.675	(0.149)
No. of children aged 3-4			-0.513	(0.161)
No. of children aged 5-11			-0.255	(0.099)
No. of children aged 12-15			-0.395	(0.141)
No. of children aged 16-18			-0.277	(0.436)
Constant	27.696	(4.447)	-6.229	(2.005)
ρ	0.186			
Test significance ρ (p-value)	1.995	(0.158)		
No. Obs.	1090.000			
Log-likelihood	-675.315			
Wald test joint signifcance (p-value)	404.090	(0.000)		

Table 7: Assumptions imposed by different estimators

Label	Estimator	Assumption	Wald Test	p-value
A1	Ignoring selection	$(y \perp\!\!\!\perp r \mid x, A)$	3.448	0.008
A2	Control function	$(r \perp\!\!\!\perp x \mid y, A, Z)$	0.001	0.970
A3	Control function	$(y \perp\!\!\!\perp Z \mid x, A)$	12.599	0.000
A4	Weighting	$(r \perp\!\!\!\perp y, x \mid A, Z)$	0.464	0.733