# Stochastic Monotonicity in Intergenerational Mobility Tables

Valentino Dardanoni [*], Mario Fiorini and Antonio Forcina

*Università degli Studi di Palermo*

*University of Technology Sydney*

*Università degli Studi di Perugia*

**Abstract**

The aim of this paper is to test for stochastic monotonicity in intergenerational socio-economic mobility tables. In other words we question whether having a parent from a high socio-economic status is never worse than having one with a lower status. Using existing inferential procedures for testing unconditional stochastic monotonicity, we first test a set of 149 intergenerational mobility tables in 35 different countries and find that monotonicity cannot be rejected in hardly any table. We then propose new testing procedures for testing conditional stochastic monotonicity, and investigate how a number of covariates such as education, cognitive and non-cognitive skills can be used to investigate whether monotonicity still holds after conditioning on these variables. Based on the NCDS cohort data from the UK, our results provide evidence that monotonicity holds even conditionally. Moreover, we do not find large differences in our results when comparing social class and wage class mobility.

*Keywords*: Intergenerational mobility, Stochastic monotonicity, Chi-bar-squared, Equality of opportunities, Marginal modelling, Local-global log-odds ratios.

*JEL Classification*: C12, C35, D30.

[*]Corresponding Author: vdardano@unipa.it, Università degli Studi di Palermo; Dipart. di Scienze Economiche, Aziendali e Finanziarie; 90128 - Palermo, Italy; +39 091 6626221

# 1 Introduction

The extent to which individuals inherit their socio-economic status in a society has important implications for policy debates on equal opportunities and social justice. Inspection of typical mobility tables and theoretical reasoning indicate that in most societies there is a general tendency for children from higher status parents to somehow fare "better" in social achievement than children from lower status parents. To substantiate this general idea, first, one needs to agree on the exact meaning of this statement, second, one has to devise a sort of testing procedure for verifying the hypothesis, and finally, one has to apply the testing procedure with real world data.

This paper develops a test of stochastic monotonicity in intergenerational socio-economic mobility tables. We question whether having a parent from a high socio-economic status is never worse than having one with a lower status, and whether this statement holds conditional on discrete and continuous covariates such as education and ability.[1] The first section of this paper will propose a precise definition of the idea that a child is better off by having a parent with a higher social status, based on the theory of monotone Markov chains. We then employ a rich data set on intergenerational social mobility, which has been made available in Ganzeboom, Luijkx, and Treiman (1989), to test the monotonicity assumption using the stochastic dominance testing procedure of Dardanoni and Forcina (1998). This large data set contains 149 mobility tables combining information from 35 different countries and different years. It is a very comprehensive data set for comparative mobility analysis, and it has the distinctive advantage of employing a consistent and well defined definition of social status. Perhaps not surprisingly, we find that for most societies the monotonicity assumption cannot be rejected.

In the second part of the paper, we consider testing for stochastic monotonicity conditional on some appropriate set of covariates $\boldsymbol{z}$. Starting with Becker and Tomes (1979), researchers have proposed economic models of intergenerational mobility. It is by now widely accepted that parental transmission of skills, beliefs, motivation and social connections are all important in explaining the strong dependence between father and son social status. It is therefore natural to test whether the positive dependence between father's and son's status is still present after conditioning for some of these characteristics. On a similar line of thought, some researchers have turned their attention to the concept of equality of opportunity (*EoP*).[2] Dardanoni, Fields, Roemer, and Puerta (2006) for example, following the seminal contribution of Putterman, Roemer, and Silvestre (1998) and Roemer (2000), describe *EoP* by distinguishing between circumstances and effort. Circumstances are aspects of the environment affecting the socio-economic status and for which society does not wish to hold individuals responsible. Effort is the set of actions affecting the status for which individuals are responsible. *EoP* implies that differences in status are ethically acceptable if they are due to differential effort but not if they are due to differential circumstances. This requires independence conditional on those covariates $\boldsymbol{z}$ that we consider effort.

[1] Stochastic monotonicity is a property of a *single* mobility table; Formby, Smith, and Zheng (2004) provide an excellent discussion of the statistical properties of various partial orderings and summary measures used to compare the degree of mobility of *different* tables.

[2] Defining the appropriate concepts of equality of opportunity and testing them empirically is an area which is undergoing much current research, see e.g. Bourguignon, Ferreira, and Menendez (2003), Peragine (2004), Lefranc, Pistolesi, and Trannoy (2006) and Fleurbaey (2008).

The main theoretical challenge is then to devise statistical inference procedures to test for stochastic monotonicity conditional on observed covariates. To this purpose, we extend the Dardanoni and Forcina (1998) test for stochastic monotonicity by allowing for conditioning on covariates. Technically, our approach employs recent advances in marginal modeling (see e.g. Bergsma and Rudas (2002) and Bartolucci, Colombi, and Forcina (2007)), along the lines of Bartolucci, Forcina, and Dardanoni (2001), who consider (unconditional) testing for a notion of positive dependence (*positive quadrant dependence*) which is weaker than the monotonicity assumption analyzed in this paper.

The importance of stochastic monotonicity in economics is highlighted in a recent paper by Lee, Linton, and Whang (2008), who propose a new test of stochastic monotonicity of a given continuous random variable $Y$ with respect to another continuous random variable $X$. They then consider stochastic monotonicity in intergenerational income mobility as a relevant field of application of their procedure by using the Panel Study of Income Dynamics (PSID). Our approach can be considered complementary to theirs, since in order to apply our approach to continuous $X$ and $Y$ some grouping is required; on the other hand, our approach can be applied to discrete ordered data without the need to replace categories with arbitrary scores. In addition, our approach allows, under some parametric restrictions, to condition on continuous and discrete covariates, while the approach taken by Lee, Linton, and Whang (2008) allows conditioning on covariates only when these are discrete and take only few values.[3]

We apply our methodology using the National Child Development Survey (NCDS), a UK data set, which follows a cohort born in 1958 over its lifetime. Information on social class and wages is available both for the cohort members and their parents. The data also provide information on the educational attainment, cognitive and non-cognitive skills of the cohort members. Given the amount of characteristics observable to the researcher, the data are particulary fit to test for conditional dependence. Our results indicate that even though our control variables explain part of the intergenerational mobility mechanism, stochastic monotonicity in this sample holds both unconditionally and conditionally.

# 2 Monotone transition matrices

Let $X$ and $Y$ denote, respectively, father's and son's lifetime socio-economic status, and assume they take $k$ possible values, which correspond to the $k$ socio-economic classes that are ordered from best to worst. Consider then a standard discrete Markov chain of intergenerational social mobility: if the unit of observation is a family of father and child, the chain can be described by the equation $\boldsymbol{p}_y' = \boldsymbol{p}_x' \, \boldsymbol{P}$, where $\boldsymbol{P}$ denotes the $(k \times k)$ transition matrix, with typical element $P_{ij} = \Pr(son\ in\ j \mid father\ in\ i) \geq 0$, and $\boldsymbol{p}_x$, $\boldsymbol{p}_y$ denote respectively the marginal distributions of father's and son's social status. The typical row $i$ of an intergenerational transition matrix indicates the probability distribution faced by a son whose father belongs to social class $i$. As argued above, it is natural to expect that when social states are ordered, sons whose fathers are in a higher social class are somewhat at an advantage with respect to the sons whose fathers are in a lower class. In a stochastic setting, this translates into the assumption that the "lottery" faced by the son of a father in class $i$ is better than the "lottery" faced by the son of a father

[3] See Remark 2.2 in their paper.

in class $i+1$. A natural definition of a "better lottery" in this context is given by the stochastic dominance ordering $\succeq$: given two $(k \times 1)$ probability vectors $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$, we say that $\boldsymbol{q}_1 \succeq \boldsymbol{q}_2$ if $q_{11} + q_{12} + \cdots + q_{1l} \geq q_{21} + q_{22} + \cdots + q_{2l}$ for all $l = 1, 2, \cdots, k-1$.

The stochastic dominance ordering gives a precise definition to the intuitive notion of differential advantage. Let $\boldsymbol{s}$ denote a real valued $(k \times 1)$ vector of "social status scores", where the typical element $s_l$ reflects a quantitative measure of the value of social class $l$, and let $\boldsymbol{p}_i$ (a row vector) denote the $i$th row of $\boldsymbol{P}$. An equivalent characterization of the stochastic dominance ordering is obtained in terms of expected social status: $\boldsymbol{p}_i\boldsymbol{s} \geq \boldsymbol{p}_j\boldsymbol{s}$ (here and hereafter, when an inequality symbol involves vectors and matrices, we mean that the inequality is satisfied elementwise) for any increasing vectors $\boldsymbol{s}$ if and only if $\boldsymbol{p}_i \succeq \boldsymbol{p}_j$.[4] The intuitive notion of background advantage is captured in the discrete Markov chain by the so called 'monotonicity' assumption. The transition matrix of a discrete Markov chain with ordered states is called monotone (see Keilson and Kester (1977), Conlisk (1990), Dardanoni (1993) and Dardanoni (1995) for applications) if each row stochastically dominates the row below it, $\boldsymbol{p}_1 \succeq \boldsymbol{p}_2 \succeq \cdots \succeq \boldsymbol{p}_k$. Notice that under the assumption of a constant transition matrix, this relationship also holds for the grandfather, grand-grandfather and so on, since if $\boldsymbol{P}$ is monotone, so is $\boldsymbol{P}^t$ for $t = 1, 2, \cdots$.

A key monotone transition matrix is the so-called "equal opportunities" transition matrix (see e.g. Prais (1955)), where at time $t$ each son faces an identical probability distribution regardless of his father's background. Given the transition equation, the equal opportunities mobility matrix is equal to $\mathbf{1}\boldsymbol{p}_y'$, so that the stochastic dominance constraint is satisfied as an equality. This particular monotone mobility matrix will play an important role in the hypothesis testing of the monotonicity assumption.

# 3 Testing unconditional monotonicity in 149 mobility tables

There is now an extensive statistical literature (see e.g. Robertson, Wright, and Dykstra (1988) and Silvapulle and Sen (2005)), on estimation and hypothesis testing in problems involving stochastic orderings. In particular, Robertson and Wright (1982) derive testing procedures based on maximum likelihood estimates of two stochastically ordered distributions, and Dykstra, Kochar, and Robertson (1991) obtain the maximum likelihood estimates of several multinomial distributions under uniform stochastic ordering, which is stronger than stochastic dominance. Dardanoni and Forcina (1998) extend these results and propose a nonparametric test for stochastic dominance which can be used to test monotonicity of the Markov chain of intergenerational mobility. In particular, Theorem 2 in their paper gives conservative bounds to the distribution of the likelihood ratio test statistic for testing monotonicity against an unrestricted alternative.

We perform Dardanoni and Forcina's procedure on a sample of cross-classification tables presented in Ganzeboom, Luijkx, and Treiman (1989). This data set, which contains 149 intergenerational class mobility tables from 35 countries, is one of the most comprehensive and well structured data set on intergenerational social mobility to date. Ganzeboom, Luijks and Treiman present the cross-classification of father's occupation by

[4] This is a well known result in the stochastic dominance literature; see e.g. Lehmann (1955).

son's current occupation for representative national samples of men aged 21-64, with the characteristic that the tables conform to a well specified six category scheme. The six social classes, in descending order of socio-economic status, are the following: 1) Large proprietors, higher and lower professionals and managers; 2) Routine non-manual workers; 3) Small proprietors with and without employees; 4) Lower grade technicians, manual supervisors and skilled manual workers; 5) Unskilled and semiskilled manual workers; 6) Self employed farmers and (unskilled) agricultural workers. The use of a common and well structured classification of social classes results in a substantial degree of comparability among the different tables.

Maximum likelihood estimation of each mobility matrix subject to the monotonicity constraint has been carried out by the iterative quadratic programming algorithm described by Dardanoni and Forcina (1998). Convergence to the fifth decimal place of the likelihood function is usually obtained within the first three iterations. The value of the likelihood ratio test statistic is reported in table 11 in the appendix, along with the number of monotonicity constraints actually binding in the sample.

Following Dardanoni and Forcina (1998, Theorem 2), the $\alpha$ critical values of the conservative unconditional chi-bar-squared test can be found by solving the equation

$$\sum_{i=0}^{(k-1)^2} \binom{(k-1)^2}{i} 2^{-(k-1)^2} Pr[\chi_i^2 = c] = \alpha.$$

By numerical integration we obtain a value $c = 22.78$ at the 95% significance level, while it equals 28.61 at the 99% level. Comparing these significance levels with the value of the likelihood ratio test statistic, it emerges that out of 149 intergenerational class mobility tables, monotonicity is rejected at the 99% significance level only for the transition matrices of Hungary 1963, Philippines 1968, Poland 1972 and Spain 1975. In addition, the monotonicity hypothesis is rejected at the 95% level for Hungary 1973 and 1983 and India 1963c. Thus, it appears that monotonicity of the intergenerational transmission mechanism can generally be considered as an assumption supported by the real world.

# 4 Conditional Stochastic Monotonicity

A large number of studies have investigated the degree of intergenerational mobility. Irrespective of country, time period, measure of socio-economic status and measure of association, almost all studies have found that parent's and offspring's adult status are not independent, but exhibit some form of positive dependence. The previous section has confirmed this "fact of life", where positive dependence is precisely formulated in terms of stochastic dominance, and formal statistical inference procedures have been employed.

Starting with Becker and Tomes (1979) researchers have proposed economic models of intergenerational mobility to uncover the mechanism behind the transmission of social status. Becker and Tomes (1986), Solon (1999), Mulligan (1999), Han and Mulligan (2001) and Restuccia and Urrutia (2004) are all attempts in that direction. At a basic level, a simple model that assumes intergenerational transmission of ability and a human capital return to parental investment that is increasing in the child's ability can already generate a high degree of immobility. Adding imperfect capital markets to the model results in even less mobility. These models also show that the degree of mobility can be highly non-linear across the father and child's socio-economic distribution, whether socio-economic status is measured by wage, consumption or education. On a more intuitive ground, Bowles and Gintis (2002), Erikson and Goldthorpe (2002) and Blanden, Gregg, and Macmillan (2007) suggest that more than a simple transmission of ability might be in place. Other factors such as race, geographical location, wealth, risk aversion, discounting of the future, non-cognitive skills, but also height and beauty can be transmitted and generate the correlation in status. Not surprisingly most of these papers have also tried to empirically investigate the black box. However, none of these studies can explain more than 60% of the overall correlation. Bjorklund, Lindahl, and Plug (2006) use unique Swedish data with information on adopted children's biological and adoptive parents to estimate intergenerational mobility associations in earnings and education. They find that both pre- and post-birth factors contribute to intergenerational earnings and education transmission. The distinction between nature and nurture is particularly important if we are asked to judge meritocracy and equality of opportunity.

On a slightly different line of thought Dardanoni, Fields, Roemer, and Puerta (2006) discuss different notions of equality of opportunity based on the distinction between circumstances and effort: "Agreement is widespread that equality of opportunity holds in a society if the chances that individuals have to succeed depend only on their own efforts and not upon extraneous circumstances that may inhibit or expand those chances. What is contentious, however, is what constitutes effort and circumstances". In their paper they describe four channels through which parents affect status in an intergenerational context: social connections, the formation of social beliefs and skills, the transmission of native ability and the instillation of preferences and aspirations. Various notions of *EoP* depend on whether these channels are regarded as circumstances or effort. In other words, if we consider all those channels as circumstances out of an individual's control, than *EoP* implies perfect intergenerational mobility. This is perhaps the strongest definition of *EoP* where parent's and offspring's status must be independent. Less stringent notions of *EoP* allow for some of those channels to be influenced by the offsprings. In turn this requires independence conditional on those covariates $\boldsymbol{z}$ that we consider individual effort. $\boldsymbol{z}$ could include measures of preferences and aspirations, native ability, social beliefs and skills, and social connections.

# 5 Testing conditional monotonicity

Recall that $X$ and $Y$ denote, respectively, father's and son's social class, and let $\boldsymbol{z}$ be a vector of covariates which may affect the joint distribution and denote with $\boldsymbol{\pi}(\boldsymbol{z})$ the vector containing the joint distribution of $X$ and $Y$ conditional on $\boldsymbol{z}$, with $Y$ categories running faster. If $\boldsymbol{z}$ was discrete and a sufficient number of observations were available for each distinct configuration of $\boldsymbol{z}$, Dardanoni and Forcina (1998)'s unconditional test procedure could be performed for each subpopulation. However these conditions are unlikely to be satisfied whenever the number of covariates is reasonably large and/or certain covariates assume a large number of distinct values. Therefore, a meaningful approach is to model the effect of covariates by a suitable link function and a regression model.

The link function that we propose is based on the mapping of the conditional distribution of $X, Y \mid \boldsymbol{z}$ into a set of row and column marginal parameters and $(k-1)^2$ association parameters. Both the row and column marginal parameters are *Global Logits* (see e.g. Agresti (2002)):

$$\rho_i(\boldsymbol{z}) = \log\left[\frac{P(X > i \mid \boldsymbol{z})}{P(X \leq i \mid \boldsymbol{z})}\right],\ i = 1, \ldots, k-1;$$

$$\xi_j(\boldsymbol{z}) = \log\left[\frac{P(Y > j \mid \boldsymbol{z})}{P(Y \leq j \mid \boldsymbol{z})}\right],\ j = 1, \ldots, k-1;$$

while the association parameters are *Local-Global Log-Odds Ratios* (see e.g. Agresti (2002)):

$$\tau_{ij}(\boldsymbol{z}) = \log\left[\frac{P(X = i, Y \leq j \mid \boldsymbol{z})P(X = i+1, Y > j \mid \boldsymbol{z})}{P(X = i, Y > j \mid \boldsymbol{z})P(X = i+1, Y \leq j \mid \boldsymbol{z})}\right],\ i, j = 1, \ldots, k-1.$$

Douglas, Fienberg, Lee, Sampson, and Whitaker (1990) show that $\boldsymbol{\pi}(\boldsymbol{z})$ is monotone if and only if the set of $(k-1)^2$ LG-log odds ratios are nonnegative. Global logits can be seen as the natural generalization of the standard binary logits when the variable is ordered; in fact, global logits can be interpreted as binary logits computed on successive splits of the ordinal response categories into "low" and "high" levels.

Collect now the corresponding parameters into the vectors $\boldsymbol{\rho}(\boldsymbol{z}), \boldsymbol{\xi}(\boldsymbol{z})$ and $\boldsymbol{\tau}(\boldsymbol{z})$ (by letting the $j$ index run faster than $i$), then let

$$\boldsymbol{\lambda}(\boldsymbol{z}) = [\boldsymbol{\rho}(\boldsymbol{z})', \boldsymbol{\xi}(\boldsymbol{z})', \boldsymbol{\tau}(\boldsymbol{z})']'$$

this has dimension $2(k-1) + (k-1)^2 = k^2 - 1$, which equals the number of free parameters in $\boldsymbol{\pi}(\boldsymbol{z})$. The results of Bartolucci, Colombi, and Forcina (2007) on marginal parameterizations may be applied to this context and imply that there exists a matrix of row contrasts $\boldsymbol{C}$ and a matrix $\boldsymbol{M}$ of zeros and ones such that, $\forall \boldsymbol{z}$

$$\boldsymbol{\lambda}(\boldsymbol{z}) = \boldsymbol{C} \log[\boldsymbol{M}\boldsymbol{\pi}(\boldsymbol{z})]$$

and the mapping from $\boldsymbol{\pi}(\boldsymbol{z})$ to $\boldsymbol{\lambda}(\boldsymbol{z})$ is invertible and differentiable for all strictly positive $\boldsymbol{\pi}(\boldsymbol{z})$ (see their Theorem 1). Thus, the set of marginal and association parameters $\boldsymbol{\lambda}(\boldsymbol{z})$ is a one-to-one mapping of $\boldsymbol{\pi}(\boldsymbol{z})$ with no modeling restriction. However, as argued above, when $\boldsymbol{z}$ takes a large number of distinct configurations, to gather information from such sparse data, the different $\boldsymbol{\lambda}(\boldsymbol{z})$ may be constrained to satisfy a linear model

$$\boldsymbol{\pi}(\boldsymbol{z}) = g[\boldsymbol{\lambda}(\boldsymbol{z})] = g[\boldsymbol{Z}\boldsymbol{\psi}] \tag{1}$$

where the design matrix $\boldsymbol{Z}$ and the model parameters $\boldsymbol{\psi}$ are derived by stacking the following system of linear equations

$$\begin{aligned} \rho_i &= \alpha_i^X + \boldsymbol{z}_X'\boldsymbol{\beta}_i^X,\ i = 1, \ldots, k-1 \\ \xi_j &= \alpha_j^Y + \boldsymbol{z}_Y'\boldsymbol{\beta}_j^Y,\ j = 1, \ldots, k-1 \\ \tau_{ij} &= \alpha_{ij}^{XY} + \boldsymbol{z}_{XY}'\boldsymbol{\beta}_{ij}^{XY},\ i, j = 1, \ldots, k-1 \end{aligned}$$

where $\boldsymbol{z}_X$, $\boldsymbol{z}_Y$ and $\boldsymbol{z}_{XY}$ denote respectively the subset of observed covariates $\boldsymbol{z}$ which are supposed to affect the marginal distribution of $X$ and $Y$ and their dependence structure.

## 5.1 Hypotheses of interest

A convenient feature of the parametrization defined above is that the hypothesis of stochastic monotonicity conditionally on relevant covariates can be expressed in the form of linear inequality constraints on the appropriate sub-vector of the $\boldsymbol{\psi}$. In particular, the hypothesis of stochastic monotonicity can be written as

$$\mathcal{H}_1 : \tau_{ij} = \alpha_{ij}^{XY} + \boldsymbol{z}'_{XY}\boldsymbol{\beta}_{ij}^{XY} \geq 0 \ \forall \ \boldsymbol{z}_{XY};\ i,j = 1,\ldots,k-1. \tag{2}$$

If we rewrite the set $\mathcal{H}_1$ compactly in terms of the model parameters as $\{\boldsymbol{\psi} : \boldsymbol{D}\boldsymbol{\psi} \geq \boldsymbol{0}\}$, notice that in typical applications the matrix $\boldsymbol{D}$ may have many more rows than columns and thus many inequalities are likely to be redundant. For example, in the application discussed below,there are more than 4000 inequalities with only 64 variables. There are algorithms, like the Fourier-Motzkin (see e.g. Schrijver (1986)) that could be used to spot and remove redundant inequalities; however, in our experience, redundant constraints do not slow significantly down the estimation algorithm anyway.

On the other hand, the hypothesis of equality of opportunities can be written as

$$\mathcal{H}_0 : \tau_{ij} = \alpha_{ij}^{XY} + \boldsymbol{z}'_{XY}\boldsymbol{\beta}_{ij}^{XY} = 0 \ \forall \ \boldsymbol{z}_{XY};\ i,j = 1,\ldots,k-1; \tag{3}$$

which can be equivalently rewritten in the standard form as

$$\mathcal{H}'_0 : \boldsymbol{\alpha}^{XY} = \boldsymbol{0} \ \&\ \boldsymbol{\beta}^{XY} = \boldsymbol{0}.$$

Finally, by $\mathcal{H}_2$ we will denote the unrestricted model.

## 5.2 Parameter estimates

Suppose now we have independent observations $(X_i, Y_i, z_i)$ for a sample of $n$ units. Let $\boldsymbol{t}(z_i)$ be a vector of size $k^2$ obtained by stacking one above the others the rows of a table having 1 in the cell $X_i, Y_j$ and 0 elsewhere. To simplify notations, in the sequel we write $\boldsymbol{t}(i)$ instead of $\boldsymbol{t}(z_i)$; a similar convention will be adopted for any vector which depends on $z_i$. Under independent sampling, conditionally on $z_i$, $\boldsymbol{t}(i)$ has a multinomial distribution with vector of probabilities $\boldsymbol{\pi}(i)$. An algorithm for maximizing the multinomial log likelihood

$$L = \sum_i \boldsymbol{t}(i)' log[\boldsymbol{\pi}(i)] \tag{4}$$

is described by Colombi and Forcina (2001) and Dardanoni and Forcina (2008), and is based on an extension of an algorithm due to Aitchison and Silvey (1958). Essentially, at each step the algorithm does the following, until convergence:

- compute a quadratic approximation of the log likelihood in terms of the canonical (log-linear) parameters;
- compute a linear approximation of the canonical parameters in terms of $\boldsymbol{\psi}$,
- solve a weighted least square problem.

When inequality constraints are present, the weighted least square problem to be solved at each step requires a quadratic optimization which is itself iterative: there are many algorithms for quadratic optimization under inequality constraints, which are usually very fast and reliable.

## 5.3 Hypotheses testing

In the following let $\boldsymbol{\psi}_2$ denote the unrestricted maximum likelihood estimate (MLE) of $\boldsymbol{\psi}$, $\boldsymbol{\psi}_1$ be the MLE of $\boldsymbol{\psi}$ under the stochastic monotonicity hypothesis $\mathcal{H}_1$ and $\boldsymbol{\psi}_0$ be the MLE of $\boldsymbol{\psi}$ under the equality of opportunity hypothesis $\mathcal{H}_0$. Let $\boldsymbol{F}(\boldsymbol{\psi})$ denote the expected information matrix with respect to $\boldsymbol{\psi}$. From standard asymptotic results it follows that, if as $n$ increases, $\boldsymbol{F}(\boldsymbol{\psi})/n$ is of full rank, $\boldsymbol{\psi}_2$ has an asymptotic normal distribution $N(\boldsymbol{\psi}, \boldsymbol{F}(\boldsymbol{\psi})^{-1})$. Therefore, hypotheses on single elements of $\boldsymbol{\psi}$ may be tested by comparing the estimate with the corresponding standard error. Joint testing may be based on the asymptotic distribution of the $LR$ statistic. Recall the well known result that the $LR$ for testing the unrestricted model against $\mathcal{H}_0$

$$T_{02} = 2(L(\boldsymbol{\psi}_2) - L(\boldsymbol{\psi}_0)) \tag{5}$$

has asymptotic $\chi^2_r$ distribution where $r$ is the sum of the dimensions of $\boldsymbol{\alpha}^{XY}$ and $\boldsymbol{\beta}^{XY}$.

When inequalities are involved, the testing problem may be split into testing the unrestricted model $\mathcal{H}_2$ against $\mathcal{H}_1$ and testing $\mathcal{H}_1$ against $\mathcal{H}_0$. The corresponding $LR$ statistics may be written as

$$\begin{aligned} T_{01} &= 2(L(\boldsymbol{\psi}_1) - L(\boldsymbol{\psi}_0)) && (6)\\ T_{12} &= 2(L(\boldsymbol{\psi}_2) - L(\boldsymbol{\psi}_1)) && (7) \end{aligned}$$

It is also useful to recall the following:

**Definition 1** *Let $\boldsymbol{b} \sim N(\boldsymbol{0}, \boldsymbol{V})$ be a $k$-dimensional normal random vector, and let $\mathcal{C}$ be a polyhedral cone in $R^k$. The squared norm of the projection of $\boldsymbol{b}$ onto $\mathcal{C}$ is a chi-bar-squared random variable $\bar{\chi}^2(\mathcal{C}, \boldsymbol{V})$*

$$\bar{\chi}^2(\mathcal{C}, \boldsymbol{V}) = \boldsymbol{b}'\boldsymbol{V}^{-1}\boldsymbol{b} - min_{\boldsymbol{a}\in\mathcal{C}}(\boldsymbol{b}-\boldsymbol{a})'\boldsymbol{V}^{-1}(\boldsymbol{b}-\boldsymbol{a}) \tag{8}$$

*and has distribution function:*

$$Pr(\bar{\chi}^2(\mathcal{C}, \boldsymbol{V}) \leq x) = \sum_{i=0}^{k} w_i(\mathcal{C}, \boldsymbol{V}) F_{\chi}(x, i) \tag{9}$$

*where $F_{\chi}(x, i)$ denotes the distribution function of a chi-square with $i$ d.f. and $w_i(\mathcal{C}, \boldsymbol{V})$ is the probability that the projection of $\boldsymbol{b}$ onto $\mathcal{C}$ belongs to a face of dimension $i$.*

We recall the following, which can be derived e.g. from Shapiro (1985) or Dardanoni and Forcina (1999):

**Proposition 1** *Under the assumption that the true value $\boldsymbol{\psi}^o$ belongs to the interior of $\mathcal{H}_0$, the asymptotic distributions of $T_{01}$ and $T_{12}$ are:*

$$\begin{aligned} T_{01} &\to \bar{\chi}^2(\mathcal{H}_1, \boldsymbol{F}^{-1}(\boldsymbol{\psi}^o))\\ T_{12} &\to \bar{\chi}^2(\mathcal{H}_1^o, \boldsymbol{F}^{-1}(\boldsymbol{\psi}^o)) \end{aligned} \tag{10}$$

*where $\mathcal{H}_1^o$ is its dual of $\mathcal{H}_1$ in the metric determined by the information matrix at $\boldsymbol{\psi}^o$.*

Asymptotic critical values for these statistics depend on the probability weights $w_i(\mathcal{H}_1, \boldsymbol{F}^{-1}(\boldsymbol{\psi}^o))$. Unfortunately, except in very small problems, no closed form expression is available for the computation of these weights. However, reliable estimates may be obtained by Monte Carlo simulations as described by Dardanoni and Forcina (1998).

It is worth recalling briefly the idea upon which the estimation of the probability weights is based. Let $\hat{\boldsymbol{b}}$ denote the projection of $\boldsymbol{b} \sim N(\boldsymbol{0}, \boldsymbol{V})$ onto $\mathcal{H}_1$, $\boldsymbol{D}(\boldsymbol{b})$ be the subset of rows of $\boldsymbol{D}$ such that $\boldsymbol{D}(\boldsymbol{b})\hat{\boldsymbol{b}} = \boldsymbol{0}$, and $\boldsymbol{Z}(\boldsymbol{b})$ the orthogonal complement of $\boldsymbol{D}(\boldsymbol{b})$. Then $w_g(\mathcal{H}_1, \boldsymbol{F}^{-1}(\boldsymbol{\psi}^o))$ is the probability that $\boldsymbol{Z}(\boldsymbol{b})$ has rank $g$. Clearly, only non redundant rows can appear in $\boldsymbol{D}(\boldsymbol{b})$ and thus the presence of possibly redundant inequalities has no effect on the estimation of weights.

Since $\mathcal{H}_1$ is a composite hypothesis, one should search for the value of $\boldsymbol{\psi} \in \mathcal{H}_1$ which gives the least favorable asymptotic null distribution for $T_{12}$ and, as Wolak (1991) has shown, this value does not necessarily belongs to $\mathcal{H}_0$. Dardanoni and Forcina (1998) discuss some practical solutions to this problem. Finally, notice that the joint distribution of $T_{01}$ and $T_{12}$ can also be derived (see Dardanoni and Forcina (1999) for details), where use of this joint distribution for hypotheses testing is also compared with alternative testing procedures.

# 6 Conditional monotonicity in the NCDS data set

In order to test for conditional monotonicity we use the National Child Development Study (NCDS), an ongoing survey that originally targeted over 17,000 babies born in Britain in the week 3-9 March 1958. Surviving members of this birth cohort have been surveyed on seven further occasions in order to monitor their changing health, education, social and economic circumstances: in 1965 (age 7), 1969 (age 11), 1974 (age 16), 1981 (age 23), 1991 (age 33), 1999 (age 41) and 2004 (age 46). At the age of 7, 11 and 16 mathematics, reading and general skills tests were taken by the cohort member, while at the age of 7 and 11 information on non-cognitive skills was also collected.

From the age of 16 individuals could leave education and enter the labor market. For those who stayed, the surveys from 1981 onwards together with a 1978 school survey provide information on the qualifications attained. Data on wages and social class were gathered at age 23, 33, 41 and 46. To study intergenerational mobility we also need data on parental socio-economic status. The first 4 surveys (1958,1965,1969,1974) contain data about parental background including age, education (1974), wage (1974), social class of father (1965, 1969, 1974) and mother (1974). These data sets therefore bring together information on socio-economic status for two consecutive generations.[5]

## 6.1 Choosing how to measure socio-economic status

To apply our stochastic monotonicity tests we first have to find suitable variables representing socio-economic status $X$ and $Y$. Since true socio-economic status is not observed,

[5]The NCDS data managers have also collected information on the cohort members' children in 1991. However, back then these children were still very young and had not entered the labor market yet. No further information on these and new children of the cohort members was gathered in the 1999 and 2004 surveys.

intergenerational mobility scholars typically employ either wage (income) or social class in their analysis.

Economists often look at wages or income as the most important observed measure of socio-economic status. However both can be very sensitive to measurement error or temporary shocks such as short periods of unemployment, health shocks, or even short business cycles. In the standard linear model, using current wages rather than true lifetime socio-economic status can result in attenuation bias. Researchers try to solve this problem either by using average wage (income), whenever the data provide repeated observations, or by using an instrumental variable approach (see e.g. Zimmerman (1992) for a discussion on the effect measurement error on measured mobility in the linear regression model). Notice that the attenuation bias holds also in our discrete mobility tables setting (Carroll, Ruppert, Stefanski, and Crainiceanu (2006) contains a thorough discussion of measurement error in non linear models); see e.g. Neuhaus (1999) for an analysis in the logit model.

On the other hand, sociologists prefer to use social class as a measure of lifetime socio-economic status (see for instance Erikson and Goldthorpe (2002)). They argue that not only social class is less sensitive to temporary shocks but also that it includes immaterial aspects such as prestige and power. The main limitations of social class originate from its subjectivity, since it is the researcher, using a combination of labor market occupation, education and other factors, that imputes the social class of the individual, sometimes also in an ordered manner, from the more prestigious occupation downwards. The way occupations are coded into social classes can sometimes affect the results. Moreover, the prestige associated with a social class can be time varying, i.e. being in, lets say, the skilled manual category in the 1960's is very much different than being in this category today, and this is a relevant problem in the case of intergenerational mobility where we look at individuals born twenty to forty years apart. Finally, within a class there could clearly be a large degree of heterogeneity; a painter and a carpenter may both be defined as skilled manual workers, but of course the socio-economic status of, say, Picasso is very different from that of an unknown painter. Yet, some of these problems affect wages (income) too. A miner might earn even more than an Academic professor due to the risk associated to his job, yet not many professors would choose to become miners.

Choosing how to measure socio-economic status inevitably depends on the data available. In our data there is not enough information to construct a reliable measure of father's permanent wage since this is observable only at one point in time. To overcome this problem Dearden, Machin, and Reed (1997) regress current wage on non time-varying factors such as education and social class, and then use the predicted variable as a measure of permanent wage. However, while there is no guarantee that this procedure eliminates attenuation bias, it also leads to a mix of wage and social class mobility (because social class is used to predict wages) and it is not really suited to test for conditional mobility (because it directly uses education to predict wages). Moreover, as we show below, there are several individuals for whom wages are not available while we can observe social class. For these reasons, we consider social class a more reliable measure of lifetime socio-economic status than wages.

In this application we use the 1991 data on the cohort member socio-economic status coupled with the 1974 data on father's status. These are also the NCDS surveys used by Dearden, Machin, and Reed (1997) and Blanden, Gregg, and Macmillan (2007) in their studies on wage mobility. We first select all male cohort members for which we observe cognitive and non-cognitive skills at age 7, 11 and 16 (cognitive skills only), educational attainment and father's age. We then select those individual's for whom we observe both social class in 1991 and father's class in 1974.[6]

Table 1 shows summary statistics for the social class measures. Social Class is a status variable grouping occupations into six broad categories, ordered on a skill basis.[7] In the data, sons are more skilled than their parents were. The distribution of son's socio-economic status actually stochastically dominates the father's one.

Table 1: Social Class - Males

| | Son | | Father | |
|---|---|---|---|---|
| Professional | 7.21 | ( 7.21) | 5.97 | ( 5.97) |
| Intermediate | 35.84 | ( 43.05) | 21.63 | ( 27.60) |
| Skilled Non-manual | 12.92 | ( 55.97) | 10.87 | ( 38.47) |
| Skilled Manual | 28.89 | ( 84.86) | 45.01 | ( 83.47) |
| Semiskilled | 12.77 | ( 97.63) | 13.18 | ( 96.65) |
| Unskilled | 2.37 | (100.00) | 3.35 | (100.00) |
| Observations | 1942 | | 1942 | |

Values are in percentages. Numbers in brackets are cumulated percentages.

Table 2 shows summary statistics for parent's and son's weekly net wages, son's highest educational qualification and father's age. The NCDS collected information on parental net wages only in 1974, with separate questions about father's and mother's wages and other sources of income. Note that in the original coding wage was grouped into 12 wage bands and we assign to each observation the median value of the observed band. Finally, in the data there are several individuals for whom father's social class are available but wages are not, while the opposite is quite rare. There are a number of explanations. Some individuals (or their fathers) are self-employed, and their wages are not reliable. For other individuals the wage is not available either because they were unemployed or because they chose not to report it.

## 6.2 Mobility Tables

In this section we present some unconditional transition matrices. We group social classes into three categories roughly corresponding to a high/medium/low skilled partition.

3 Professional+Intermediate

2 Skilled Non-Manual + Skilled Manual

[6]We select males to make our results comparable to previous studies.

[7]Social Class variables are derived according to the Registrar Classification (RG). This classification imputes social class using only information about occupation. This is a quite common and simple grouping methodology, even though some sociologists have proposed alternative ones. The Goldthorpe class schema for instance (see Erikson and Goldthorpe (2002)) aims to capture qualitative differences in employment relations. Unfortunately, the classes distinguished by this schema are not consistently ordered according to some inherent hierarchical principle. Therefore the Goldhtorpe class schema does not suit our statistical model.

Table 2: Summary Statistics

| | Obs | Mean | S.D. |
|---|---|---|---|
| Son's Net Wage | 1341 | 301.07 | 102.79 |
| Father's Net Wage | 1486 | 232.56 | 83.93 |
| No Qualification | 1942 | 0.45 | 0.49 |
| O Levels | 1942 | 0.33 | 0.46 |
| A Levels | 1942 | 0.09 | 0.29 |
| Higher Education | 1942 | 0.13 | 0.33 |
| Father's Age | 1942 | 46.55 | 6.12 |

Father and Son net wages are in January 2001 prices.

1 Semiskilled + Unskilled

There is no compelling reason to use three categories rather than two, four or any other number. Having more categories allows the researcher to have a better understanding of the heterogeneity across the joint distribution but at the same time it can make the identification of the parameters cumbersome. The larger the number of categories the lower the sample size within each cell. This problem only exacerbates when conditioning on other covariates.

Table 3 shows the mobility tables using son's social class in 1991 and father's class in 1974. In the table we also include the chi-square statistic for independence, with the degrees of freedom in brackets. The chi-square statistic is very large leading to reject the hypothesis of independence.

Table 3: Social Class Mobility Table

| Father/Son Class | 1 | 2 | 3 | |
|---|---|---|---|---|
| 1 | 4.53 | 8.08 | 3.91 | 16.53 |
| 2 | 8.81 | 26.06 | 21.01 | 55.87 |
| 3 | 1.80 | 7.67 | 18.13 | 27.60 |
| | 15.14 | 41.81 | 43.05 | 100.00 |

1942 observations. Numbers on table are percentages. Chi-Square (4) = 192.73

Table 4 shows the *Local-Global Log-Odds Ratios* in the unconditional table. The first two odds correspond to the first two rows of the mobility table. If both $\tau_{11}$ and $\tau_{12}$ are positive than the first row (low class) is stochastically dominated by the second one (medium class). The same reasoning applies to $\tau_{21}$ and $\tau_{22}$ (medium class is stochastically dominated by high class). All the odds in the table are positive. Yet the odds [$\tau_{11}$,$\tau_{12}$] are smaller than the [$\tau_{21}$,$\tau_{22}$] indicating stronger stochastic dominance toward the top of the joint distribution.

Table 4: Local Global Odds Ratios

| Row | Log-Odds | |
|---|---|---|
| $R_{12}$ | $\tau_{11}$ | 0.7181 |
| | $\tau_{12}$ | 0.6641 |
| $R_{23}$ | $\tau_{21}$ | 0.9851 |
| | $\tau_{22}$ | 1.1551 |

## 6.3 Control Variables

As we explained in the section 5, our aim is to test for dependence conditional on some characteristics of the parents and offsprings. Since most economic models of intergenerational mobility assume that the transmission of the ability endowment across generations is one of the main reasons behind immobility (see Becker and Tomes (1979) or Grawe and Mulligan (2002) for instance) a starting point is to investigate monotonicity conditional on cognitive skills. However, as Becker and Tomes (1979), Bowles and Gintis (2002), Erikson and Goldthorpe (2002) and Dardanoni, Fields, Roemer, and Puerta (2006) suggest, cognitive ability is just one dimension of the endowment stock. Recently Heckman, Stixrud, and Urzua (2006) and Cunha and Heckman (2007) show that non-cognitive skills can also explain a diverse array of outcomes such as schooling choices, wages, employment and work experience. It is quite likely that non-cognitive skills are also transmitted across generations, if not genetically, because of parental behavior and education. Finally, the human capital models predicts that high-status parents invest more in their children. In turn this implies that these children have more human capital. Therefore we choose to test for monotonicity conditional on the educational attainment, cognitive and non-cognitive skills of the offspring (son). Since the fathers were of different ages at the moment of the survey, we also control for father's age.

Given the education system faced by the 1958 cohort, its educational attainment is measured by 4 ordered categories corresponding to 'No Qualification', 'O Levels', 'A Levels' and 'Higher Education'. In the UK, schooling is compulsory up to the age of 16, when individuals can, at the end of the scholastic year, stay in education or enter the labor market. Those who stay on at age 16 enrol for the O Levels or CSE qualifications, which are taken immediately at the end of the scholastic year. These students are still aged 16 when they obtain the qualification. In the Autumn term of the same year, those who successfully obtained 5 or more O Levels/CSE can enrol for A Levels. These last 2 years, until individuals are aged 18. Passing 2 A-level constitutes the minimum level required for entry in Higher Education. Once the student has completed A Level, he can gain admission to the Universities, Polytechnic or Colleges of Higher Education where a first degree is obtained. The time needed to gain a degree varies by subject but in the majority of cases it takes 3 years.

To control for cognitive skills we use the mathematics and reading test scores. These tests were taken by the cohort members at the age of 7, 11 and 16. We use all these multiple age-skills observations but in order to reduce the control variables space, at each age we replace the original maths and reading scores with the principal component. In all cases the principal component explains no less than 90% of the total variance. In the case of non-cognitive skills things are a bit more complex. Both at age 7 and age 11 there are 12 scores for non-cognitive skills, such as depression, anxiety, hostility etc., as reported by teachers in schools. (No score is available at age 16.) In order to keep our problem computationally tractable, we do a factor analysis of the non-cognitive scores using the iterated principal factor method. Out of 12 scores, only two eigenvalues are larger than 1, with the third being sensibly smaller. Therefore, at each age point, we retain only two factors. In Table 5 we show the rotated loading factors. The first factor captures the skill to relate to other individuals, either adults or other children. The second factor captures emotional problems. There are no large differences between age 7 and 11. The final factors are obtained using the regression method.

Table 5: Loading Factors - Non-cognitive scores

| | Age 7 | | Age 11 | |
|---|---|---|---|---|
| | F1 | F2 | F1 | F2 |
| Unforthcomingness | -0.0690 | 0.7279 | -0.0807 | 0.7161 |
| Withdrawal | 0.0649 | 0.6693 | 0.0908 | 0.6856 |
| Depression | 0.2727 | 0.6757 | 0.3496 | 0.6305 |
| Anxiety for acceptance of adults | 0.4061 | -0.1096 | 0.4257 | -0.0487 |
| Hostility toward adults | 0.6042 | 0.2118 | 0.6455 | 0.1988 |
| Writing off adults and standards | 0.4362 | 0.5162 | 0.4696 | 0.4690 |
| Anxiety for acceptance by kids | 0.6662 | -0.0429 | 0.6862 | -0.0386 |
| Hostility toward children | 0.6736 | 0.0942 | 0.6593 | 0.1311 |
| Restlessness | 0.5467 | 0.1447 | 0.5258 | 0.1190 |
| Inconsequential behavior | 0.7761 | 0.2020 | 0.7771 | 0.1931 |
| Miscellaneous symptoms | 0.2354 | 0.5894 | 0.3329 | 0.5408 |
| Miscellaneous nervous symptoms | 0.2636 | 0.1910 | 0.3057 | 0.1689 |

Number of obs = 14931 (Age 7), 14158 (Age 11). Retained factors = 2.
Number of params = 23.

Finally, since father's age was recorded only in the original 1958 survey, we restrict our sample to those cohort members living with a biological father during their childhood.

## 6.4 Equality of Opportunity

Table 6 presents results when we test for equality of opportunity (independence) as shown in section 5.1 (equation 3) and section 5.3 (equation 5). To begin with we do not condition on any covariates, and test for equality of opportunity by restricting the odds to be zero. This is done by setting the constant terms $\boldsymbol{\alpha}^{XY}$ equal to zero. Since we are ultimately interested in stochastic dominance, that is whether one row "dominates" the other, we first impose the restrictions by looking at pairs of rows. Therefore, we first impose $[\tau_{11}, \tau_{12}] = 0$ ($R_{12}$) and then $[\tau_{21}, \tau_{22}] = 0$ ($R_{23}$) . We then impose the restriction on all rows simultaneously: $\boldsymbol{\tau} = 0$ (AR). The numbers in the table are the $p$-values (likelihood-ratios in brackets) from a standard chi-square distribution. The numbers in bold indicate the number of restrictions. In the unconditional case there are only the constant terms to restrict. All the $p$-values are equal to zero at the fourth decimal digit.

Table 6: P-values Independence

| Row | Unconditional | | Conditional | |
|---|---|---|---|---|
| | P-values | Restrict. | P-values | Restrict. |
| $R_{12}$ | 0.0000 | **2** | 0.0160 | **24** |
| | ( 31.3557) | | ( 41.1642) | |
| $R_{23}$ | 0.0000 | **2** | 0.0007 | **24** |
| | (114.3083) | | ( 52.4605) | |
| AR | 0.0000 | **4** | 0.0000 | **48** |
| | (189.9833) | | (103.1234) | |

Likelihood ratios in brackets.

Next, we test independence conditional on our set of covariates: cognitive skills (measured at age 7, 11 and 16), non-cognitive skills (measured at age 7 and 11) up to the third principal component, educational attainment (at age 23) and father's age. There are 13 variables in total. Table 12 in the Appendix presents the estimated coefficients and standard errors for the unrestricted model.

We model the four odds ratios ($\boldsymbol{\tau}$) and two marginal parameters for the son ($\boldsymbol{\xi}$) as a function of the full set of covariates. On the other hand we model the two marginal parameters for the father ($\boldsymbol{\rho}$) as a function of father's age only, that is we rule out that father's class can be a function of son's education, cognitive and non-cognitive skills. Note also that since all the covariates are centered, the constant terms have a direct interpretation. They measure the *Local-Global Log-Odds Ratios* for the individual with average covariates. For $\boldsymbol{\xi}$, education and cognitive skills coefficients have the expected positive sign, meaning that those sons with better education and cognitive skills are more likely to have a higher status. None of the non-cognitive skills factors is statistically significant on its own. When interpreting the coefficients for the four odds ratios ($\boldsymbol{\tau}$), one has to remember that a larger $\tau$ indicates less mobility. For instance, the Higher Education coefficient is positive and statistically significant for $\tau_{11}$ and $\tau_{12}$. Intuitively, this result suggests that those individuals with Higher Education in the medium class were very unlikely to have a father in low class, and at the same time those individuals with Higher Education in the low class were unlikely to have a father in medium class.

In last two columns of Table 6 we present the conditional $p$-values. To test for independence we impose 12 restrictions given by 1 constant term + 11 coefficients (one for each covariate) for each LG log-odds ratio. Thus the total number of restrictions equal 12 times the number of odds. Once we condition on this rich set of covariates the $p$-values become larger. At a 1% significance level independence is no longer rejected in the first row. This result confirms that our set of covariates do explain part of the mobility mechanism.

## 6.5 Conditional stochastic monotonicity

We now formally test our mobility tables for equality of opportunity against stochastic monotonicity, as explained in section 5.3, equation (6). To test the null hypothesis we need to compute the $\bar{\chi}^2$ distribution function. Being one-directional, this test is more powerful than the previous one where the null hypothesis of equality of opportunity (independence) is tested against the unrestricted model. Table 7 column (*EoP vs SM*) shows the computed $p$-values obtained under the null hypothesis of equality of opportunity against stochastic monotonicity conditional on our rich set of control variables. As expected, since this test is more powerful, all the conditional $p$-values are now smaller than in table 6. We now reject independence even at a 1% significance level.

Since this results suggests that independence is rejected even conditionally on a rich set of controls, we now test whether our data give evidence of conditional stochastic monotonicity. In column (*SM vs Unr*) we display the $p$-values obtained under the null hypothesis of stochastic monotonicity against the unrestricted model as shown in section 5.3, equation (7). The $p$-values are all very high meaning that there is strong evidence of stochastic monotonicity, or in other words very few *Local-Global Log-Odds Ratios* are negative in the sub-tables when conditioning on $\boldsymbol{z}$.

Table 7: P-values Stochastic Monotonicity

| Row | *EoP vs SM* | *SM vs Unr* |
|---|---|---|
| $R_{12}$ | 0.0003 | 0.8450 |
| $R_{23}$ | 0.0000 | 0.8284 |
| AR | 0.0000 | 0.9460 |

Numbers in table are $p$-values.

## 6.6 Wage Mobility

For completeness, in this section we replicate the analysis using the wage mobility table illustrated in table 8. As we discussed above, the sample size is much smaller. As in the case of social class, we group the individuals into three categories based on their wage percentile. Unfortunately, given that the original father's wage variable was coded into 12 bands, it is not possible to exactly partition it in three terciles. The chi-square statistic is smaller than for social class (table 3) though we still reject independence.

Table 8: Wage Class Mobility Table

| Father/Son Class | 1 | 2 | 3 | |
|---|---|---|---|---|
| 1 | 14.04 | 10.78 | 8.79 | 33.61 |
| 2 | 12.41 | 10.51 | 10.42 | 33.33 |
| 3 | 7.61 | 11.32 | 14.13 | 33.06 |
| | 34.06 | 32.61 | 33.33 | 100.00 |

1104 observations. Chi-Square (4) = 37.03.

In table 9 we test for equality of opportunity both unconditional and conditional on our rich set of control variables. In the Appendix, table 13, we show the estimated coefficients and standard errors for the unrestricted model. Anew, the p-values become larger once we condition on $\boldsymbol{z}$. Note however, that for $R_{12}$ we can not reject unconditional independence. Overall the p-values are larger than in the case of social class. From a comparison of tables 12 and 13 we also notice that three out of four *Local-Global Log-Odds Ratios* constant terms (the odds for the average individual) are larger when using social class. This difference between wage and social class might be due to measurement error or temporary shocks affecting wages more than social class. However, we can not rule out that sample selection might also be driving this difference.

Table 9: P-values Independence - Wage Mobility

| Row | Unconditional | | Conditional | |
|---|---|---|---|---|
| | P-values | Restrict. | P-values | Restrict. |
| $R_{12}$ | 0.3267 | **2** | 0.4808 | **24** |
| | ( 2.2372) | | ( 23.6668) | |
| $R_{23}$ | 0.0001 | **2** | 0.0002 | **24** |
| | ( 18.8047) | | ( 55.8950) | |
| AR | 0.0000 | **4** | 0.0044 | **48** |
| | ( 36.0567) | | ( 77.5813) | |

Likelihood ratios in brackets.

In table 10 we show the p-values obtained while testing for stochastic monotonicity. As for social class, when testing *EoP* against stochastic monotonicity, column (*EoP vs*

*SM*), all the conditional $p$-values are now smaller than in table 9, though we still can not reject independence in the first row ($R_{12}$). Column (*SM vs Unr*) provide again evidence of conditional stochastic monotonicity.

Table 10: P-values Stochastic Monotonicity - Wage Mobility

| Row | *EoP vs SM* | *SM vs Unr* |
|---|---|---|
| $R_{12}$ | 0.3342 | 0.5574 |
| $R_{23}$ | 0.0005 | 0.0701 |
| AR | 0.0007 | 0.3064 |

Numbers in table are $p$-values.

# 7 Conclusions

The aim of this paper was to test for stochastic monotonicity in intergenerational socio-economic mobility tables. To do so we apply and extend the methodology discussed in Dardanoni and Forcina (1998) and Bartolucci, Forcina, and Dardanoni (2001). We first test for unconditional stochastic monotonicity using a set of 149 intergenerational mobility tables in 35 different countries, where it emerges that monotonicity cannot be rejected in hardly any table.

We then explain how a number of controls such as education, cognitive and non-cognitive skills can be used to investigate whether monotonicity still holds after conditioning on these factors. In the economic literature, no previous work on intergenerational mobility tables has dealt with continuous controls. Since current research on mobility is focussing on the determinants dependence in socio-economic status between parents and offspring, conditioning on discrete and continuous covariates is increasingly important.

To apply our test of conditional monotonicity we use the NCDS, a UK cohort data with information on the socio-economic status of the cohort members and their parents, individual's educational qualifications, cognitive and non-cognitive skills. Our tests show evidence of stochastic monotonicity both unconditionally and conditionally. While it is not surprising that the unconditional joint distribution exhibits monotonicity, it is interesting to find that such a strong form of dependence subsists even conditional on educational achievement, cognitive and non-cognitive skills. This result reinforces the findings of Solon (1999), Bowles and Gintis (2002), Restuccia and Urrutia (2004), Dardanoni, Fields, Roemer, and Puerta (2006), Blanden, Gregg, and Macmillan (2007) indicating that part of the mechanism linking parent's and offspring's socio-economic status is still a black box. Finally, we observe only minor differences between social and wage class tables; if anything we find that there seems to be more dependence when using social rather than wage class.

# References


Agresti, A. (2002): *Categorical Data Analysis*. John Wiley and Sons, Inc.

Aitchison, J., and S. D. Silvey (1958): "Maximum Likelihood Estimation of Parameters Subject to Restraints," *Annals of Mathematical Statistics*, 29, 813–828.

Bartolucci, F., R. Colombi, and A. Forcina (2007): "An extended class of marginal link functions for modelling contingency tables by equality and inequality constraints," *Statistica Sinica*, 17(2).

Bartolucci, F., A. Forcina, and V. Dardanoni (2001): "Positive Quadrant Dependence and Marginal Modelling in Two-Way Tables With Ordered Margins," *Journal of the American Statistical Association*, 96, 1497–1505.

Becker, G. S., and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, 1153–1189.

——— (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4, S1–S39.

Bergsma, W. P., and T. Rudas (2002): "Marginal models for categorical data," *Annals of Statistics*, 30, 140–159.

Bjorklund, A., M. Lindahl, and E. Plug (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121(3), 999–1028.

Blanden, J., P. Gregg, and L. Macmillan (2007): "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education," *Economic Journal*, 117(519), C43–C60.

Bourguignon, F., F. H. G. Ferreira, and M. Menendez (2003): "Inequality of Outcomes and Inequality of Opportunities in Brasil," DELTA Working Papers 2003-24, DELTA.

Bowles, S., and H. Gintis (2002): "The Inheritance of Inequality," *Journal of Economic Perspectives*, 16(3), 3–30.

Carroll, R. J., D. Ruppert, L. A. Stefanski, and C. M. Crainiceanu (2006): *Measurement Error in Nonlinear Models*. Chapman Hall.

Colombi, R., and A. Forcina (2001): "Marginal regression models for the analysis of positive association," *Biometrika*, 88(4), 1007–1019.

Conlisk, J. (1990): "Monotone mobility matrices," *Journal of Mathematical Sociology*, pp. 173–191.

Cunha, F., and J. J. Heckman (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2).

Dardanoni, V. (1993): "Measuring Social Mobility," *Journal of Economic Theory*, 61(2), 372–394.

——— (1995): "Income Distribution Dynamics: Monotone Markov Chains Make Light Work," *Social Choice and Welfare*, 12(2), 181–192.

Dardanoni, V., G. S. Fields, J. E. Roemer, and M. L. S. Puerta (2006): *Mobility and Inequality: Frontiers of Research from Sociology and Economics*chap. How Demanding Should Equality of Opportunity Be, and How Much Have We Achieved? Stamford University Press.

Dardanoni, V., and A. Forcina (1998): "A Unified Approach to Likelihood Inference on Stochastic Orderings in a Nonparametric Context," *Journal of the American Statistical Association*, 93, 1112–1123.

——— (1999): "Inference for Lorenz Curve Orderings," *Econometrics Journal*, 2, 48–74.

——— (2008): "Multivariate Ordered Regression," Mimeo.

Dearden, L., S. Machin, and H. Reed (1997): "Intergenerational Mobility in Britain," *Economic Journal*, 107(440), 47–66.

Douglas, R., S. E. Fienberg, M. T. Lee, A. R. Sampson, and L. R. Whitaker (1990): *Topics in Statistical Dependence*chap. Positive Dependence Concepts for Ordinal Contingency Tables, pp. 189–202. Hayward, CA: Institute of Mathematical Statistics.

Dykstra, R., S. Kochar, and T. Robertson (1991): "Statistical Inference for Uniform Stochastic Ordering in Several Populations," *The Annals of Statistics*, 19(2), 870–888.

Erikson, R., and J. H. Goldthorpe (2002): "Intergenerational Inequality: A Sociological Perspective," *Journal of Economic Perspectives*, 16(3), 31–44.

Fleurbaey, M. (2008): *Fairness, Responsibility, and Welfare*. Oxford University Press.

Formby, J. P., W. J. Smith, and B. Zheng (2004): "Mobility Measurement, Transition Matrices and Statistical Inference," *Journal of Econometrics*, 120, 181–205.

Ganzeboom, H. B. G., R. Luijkx, and D. J. Treiman (1989): "Intergenerational Class Mobility in Comparative Perspective," *Research in Social Stratification and Mobility*, 8, 3–79.

Grawe, N. D., and C. B. Mulligan (2002): "Economic Interpretations of Intergenerational Correlations," *Journal of Economic Perspectives*, 16(3), 45–58.

Han, S., and C. B. Mulligan (2001): "Human Capital, Heterogeneity and Estimated Degrees of Intergenerational Mobility," *Economic Journal*, 111(470), 207–43.

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24(3).

Keilson, J., and A. Kester (1977): "Monotone matrices and monotone Markov chains," *Stochastic Processes and their Applications*, 5, 231–41.

Lee, S., O. Linton, and Y. Whang (2008): "Testing Stochastic Monotonicity," *Econometrica*, forthcoming.

Lefranc, A., N. Pistolesi, and A. Trannoy (2006): "Inequality of opportunities vs. inequality of outcomes: Are Western societies all alike?," Working Papers 54, ECINEQ, Society for the Study of Economic Inequality.

Lehmann, E. (1955): "Ordered families of distributions," *Annals of Mathematical Statistics*, 26, 399–419.

Mulligan, C. B. (1999): "Galton versus the Human Capital Approach to Inheritance," *Journal of Political Economy*, 107(6), S184–S224.

Neuhaus, J. M. (1999): "Bias and Efficiency Loss Due to Misclassified Responses in Binary Regression," *Biometrika*, 86(4), 843–855.

Peragine, V. (2004): "Ranking Income Distributions According to Equality of Opportunity," *Journal of Economic Inequality*, 2(1), 11–30.

Prais, S. J. (1955): "Measuring social mobility," *Journal of the Royal Statistical Society: Series A*, 118, 56–66.

Putterman, L., J. E. Roemer, and J. Silvestre (1998): "Does Egalitarianism Have a Future?," *Journal of Economic Literature*, 36(2), 861–902.

Restuccia, D., and C. Urrutia (2004): "Intergenerational Persistence of Earnings: The Role of Early and College Education," *American Economic Review*, 94(5), 1354–78.

Robertson, T., F. Wright, and R. Dykstra (1988): *Order Restricted Statistical Inference*. J. Wiley and Sons, New York, N.Y.

Robertson, T., and F. T. Wright (1982): "Bounds on Mixtures of Distributions Arising in Order Restricted Inference," *The Annals of Statistics*, 10(1), 302–306.

Roemer, J. E. (2000): *Equality of Opprotunity*Meritocracy and Economic Inequality. Princeton University Press, editors K. Arrow, Samuel Bowles and S. Durlauf.

Schrijver, A. (1986): *Theory of Linear and Integer Programming*. John Wiley & Sons.

Shapiro, A. (1985): "Asymptotic Distribution of Test Statistics in the Analysis of Moment Structures Under Inequality Constraints," *Biometrika*, 72(1), 133–144.

Silvapulle, M. J., and P. K. Sen (2005): *Constrained Statistical Inference: Inequality, Order, and Shape Restrictions*.

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

Wolak, F. A. (1991): "The Local Nature of Hypotheses Tests Involving Inequality Constraints,," *Econometrica*, 59, 981–995.

Zimmerman, D. J. (1992): "Regression Toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–429.

Table 11: Ganzeboom, Luijks and Treiman tables

| Country/yr | LR | constr. | Country/yr | LR | constr. | Country/yr | LR | constr. |
|---|---|---|---|---|---|---|---|---|
| AUS65 | 16.6230 | 6 | GER80a | 3.3808 | 4 | NOR72l | 5.6674 | 8 |
| AUS67 | 4.5578 | 5 | GER80p | 4.1145 | 9 | NOR72s | 3.0935 | 3 |
| AUS67l | 5.1737 | 8 | GER80z | 4.5121 | 7 | NOR73 | 8.9025 | 4 |
| AUS73 | 11.239 | 4 | GER82a | 12.5920 | 6 | NOR82w | 4.2696 | 8 |
| AUS87 | 3.8706 | 6 | GER84a | 1.6064 | 4 | NZE76 | 2.8615 | 4 |
| AUT69n | 5.2945 | 4 | HKG67 | 1.8590 | 8 | PHI68 | 66.9629 | 1 |
| AUT74p | 5.0644 | 5 | HUN62 | 70.2130 | 2 | PHI73 | 9.2018 | 2 |
| AUT78 | 15.4645 | 4 | HUN73 | 28.5077 | 4 | POL72 | 50.5442 | 4 |
| BEL71e | 13.7803 | 6 | HUN73l | 16.5144 | 4 | POL82 | 3.5511 | 5 |
| BEL75 | 11.0925 | 5 | HUN82 | 1.3950 | 3 | POL87 | 9.7428 | 6 |
| BEL76 | 5.4211 | 4 | HUN83 | 26.3894 | 3 | PUE54 | 0.1228 | 1 |
| BRA73 | 9.4740 | 3 | HUN86 | 5.3734 | 6 | QUE60 | 9.7055 | 8 |
| CAN73 | 4.4114 | 4 | IND62c | 8.3105 | 6 | QUE73 | 4.9998 | 3 |
| CAN82w | 3.9487 | 5 | IND63a | 14.2253 | 8 | QUE77 | 1.4496 | 6 |
| CAN84 | 3.4051 | 3 | IND63c | 26.7984 | 7 | SCO74 | 8.9427 | 6 |
| CSK67 | 21.2481 | 2 | IND71n | 5.2398 | 4 | SCO75 | 16.0839 | 5 |
| DEN71 | 6.4718 | 9 | IRE74 | 8.1673 | 3 | SPA65 | 2.8188 | 2 |
| DEN72l | 7.5135 | 6 | ISR62c | 8.9032 | 7 | SPA67t | 9.1810 | 7 |
| DEN72s | 1.7511 | 4 | ISR74 | 4.6268 | 1 | SPA75 | 68.8972 | 4 |
| ENG51 | 5.0602 | 4 | ITA63 | 6.5001 | 4 | SWE50 | 9.1278 | 5 |
| ENG63 | 1.2945 | 5 | ITA68 | 4.3434 | 6 | SWE60 | 1.9379 | 4 |
| ENG67t | 0.2951 | 6 | ITA72 | 9.5852 | 6 | SWE72l | 5.5451 | 7 |
| ENG69 | 0.9969 | 4 | ITA74 | 6.6980 | 4 | SWE72s | 2.0677 | 3 |
| ENG72 | 10.4063 | 4 | ITA75p | 14.7003 | 7 | SWE73 | 0.2150 | 2 |
| ENG74 | 2.7902 | 5 | JAP55 | 6.6026 | 4 | SWE83w | 4.7170 | 5 |
| ENG74p | 6.0480 | 5 | JAP65 | 1.0588 | 5 | SWI76p | 4.1620 | 4 |
| ENG83 | 3.4236 | 3 | JAP67 | 7.6923 | 7 | TAI70 | 10.0489 | 9 |
| ENG86 | 2.0181 | 5 | JAP69t | 4.3844 | 9 | TAI70l | 7.6641 | 6 |
| FIN67t | 8.6317 | 10 | JAP71n | 2.2285 | 3 | USA47 | 4.5285 | 4 |
| FIN72l | 6.8491 | 8 | JAP75 | 1.0965 | 5 | USA47l | 1.7094 | 5 |
| FIN72s | 3.5080 | 8 | MAL67 | 22.1726 | 7 | USA59c | 3.2706 | 5 |
| FIN75p | 2.1559 | 6 | NET58 | 6.3140 | 6 | USA62o | 2.3330 | 3 |
| FIN80 | 12.2139 | 4 | NET67t | 2.6724 | 7 | USA72g | 5.7139 | 8 |
| FIN82w | 10.7844 | 9 | NET70 | 2.6992 | 4 | USA73g | 8.7437 | 8 |
| FRA58 | 4.6786 | 4 | NET71 | 2.4754 | 5 | USA73o | 7.7145 | 3 |
| FRA64 | 9.3367 | 4 | NET71e | 3.6097 | 5 | USA74g | 3.3402 | 5 |
| FRA67 | 7.8008 | 8 | NET74p | 1.1753 | 4 | USA74p | 8.4365 | 5 |
| FRA70 | 11.7893 | 2 | NET76 | 1.2016 | 4 | USA75g | 2.8330 | 7 |
| FRA71e | 7.0699 | 4 | NET77 | 5.8811 | 4 | USA76g | 3.3676 | 5 |
| GER59 | 3.0701 | 5 | NET77x | 2.1617 | 3 | USA77g | 5.7246 | 4 |
| GER69 | 3.3086 | 8 | NET79p | 12.1392 | 8 | USA78g | 4.8763 | 6 |
| GER69k | 10.5659 | 5 | NET82 | 4.6030 | 7 | USA80g | 5.0530 | 7 |
| GER75p | 4.4880 | 5 | NET82u | 5.1157 | 8 | USA81w | 1.8263 | 5 |
| GER76z | 8.3095 | 6 | NET85 | 1.5353 | 4 | USA82g | 7.8830 | 5 |
| GER77z | 3.6863 | 6 | NIG71n | 9.6726 | 5 | USA83g | 3.0734 | 4 |
| GER78 | 6.0179 | 7 | NIR68 | 2.5272 | 7 | USA84g | 4.5427 | 8 |
| GER78x | 3.5050 | 7 | NIR73 | 4.1850 | 6 | USA85g | 6.5725 | 5 |
| GER78z | 6.2942 | 5 | NOR57 | 11.0213 | 7 | USA86g | 6.6470 | 8 |
| GER79z | 2.2963 | 8 | NOR65 | 1.1214 | 4 | YUG67t | 6.2877 | 8 |
| GER80 | 4.8059 | 5 | NOR67t | 1.8285 | 6 | — | — | — |

## Table 12: Unrestricted Model - Social Class (1974)

| Parameter | Marginal | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\rho_1$ | | $\rho_2$ | | $\xi_1$ | | $\xi_2$ | |
| Constant | 1.6352 | ( 26.4510) | -0.9620 | (-18.9698) | 2.1078 | ( 24.1718) | -0.2674 | ( -5.2418) |
| Father Age | -0.0338 | ( -3.5990) | -0.0003 | ( -0.0396) | -0.0290 | ( -2.8677) | -0.0004 | ( -0.0438) |
| O Level | — | ( — ) | — | ( — ) | 0.5343 | ( 2.9851) | 0.2329 | ( 1.8124) |
| A Level | — | ( — ) | — | ( — ) | 1.6015 | ( 3.0112) | 1.0127 | ( 4.7179) |
| High. Educ. | — | ( — ) | — | ( — ) | 0.7777 | ( 2.2970) | 1.3276 | ( 6.5289) |
| C. Skills 7 | — | ( — ) | — | ( — ) | -0.0798 | ( -1.0457) | -0.0537 | ( -0.8778) |
| C. Skills 11 | — | ( — ) | — | ( — ) | 0.1519 | ( 1.4253) | -0.0109 | ( -0.1395) |
| C. Skills 16 | — | ( — ) | — | ( — ) | 0.4170 | ( 3.9301) | 0.5326 | ( 6.1489) |
| NC. Skills 7 (1st) | — | ( — ) | — | ( — ) | 0.0127 | ( 0.1694) | 0.0089 | ( 0.1438) |
| NC. Skills 7 (2nd) | — | ( — ) | — | ( — ) | -0.0371 | ( -0.4785) | -0.0749 | ( -1.0838) |
| NC. Skills 11 (1st) | — | ( — ) | — | ( — ) | 0.0309 | ( 0.4194) | -0.0038 | ( -0.0592) |
| NC. Skills 11 (2nd) | — | ( — ) | — | ( — ) | -0.1022 | ( -1.4225) | -0.0955 | ( -1.4514) |
| Parameter | Odds | | | | | | | |
| | $\tau_{11}$ | | $\tau_{12}$ | | $\tau_{21}$ | | $\tau_{22}$ | |
| Constant | 0.7785 | ( 3.4852) | 0.5249 | ( 3.0082) | 0.4118 | ( 1.5316) | 0.6935 | ( 5.7366) |
| Father Age | 0.0255 | ( 0.9328) | 0.0517 | ( 2.0670) | -0.0878 | ( -3.1548) | -0.0375 | ( -1.9414) |
| O Level | -0.4090 | ( -0.8632) | -0.0218 | ( -0.0527) | 0.4563 | ( 0.9364) | 0.1094 | ( 0.3608) |
| A Level | 0.9018 | ( 0.6769) | 0.0666 | ( 0.1026) | -0.1924 | ( -0.1311) | -0.1019 | ( -0.1968) |
| High. Educ. | 1.7188 | ( 1.9630) | 1.7835 | ( 2.9363) | 0.8091 | ( 0.6344) | -0.1658 | ( -0.3114) |
| C. Skills 7 | 0.1127 | ( 0.5435) | 0.0896 | ( 0.4560) | -0.2430 | ( -1.1552) | 0.0658 | ( 0.4555) |
| C. Skills 11 | -0.0001 | ( -0.0004) | 0.2169 | ( 0.9342) | -0.2386 | ( -0.8128) | -0.1569 | ( -0.8348) |
| C. Skills 16 | -0.0305 | ( -0.1055) | -0.8564 | ( -2.9761) | -0.0636 | ( -0.2181) | -0.1141 | ( -0.5569) |
| NC. Skills 7 (1st) | 0.1527 | ( 0.7321) | -0.0168 | ( -0.0837) | -0.4423 | ( -2.1842) | -0.1249 | ( -0.8676) |
| NC. Skills 7 (2nd) | -0.2082 | ( -0.8959) | -0.0596 | ( -0.2756) | -0.4707 | ( -2.2672) | -0.2638 | ( -1.6378) |
| NC. Skills 11 (1st) | -0.1114 | ( -0.5616) | 0.2076 | ( 0.9010) | 0.5944 | ( 2.4360) | -0.0147 | ( -0.0987) |
| NC. Skills 11 (2nd) | -0.0560 | ( -0.2812) | -0.1678 | ( -0.8019) | 0.0958 | ( 0.5054) | 0.0340 | ( 0.2208) |

Columns 1 and 2 correspond to the father's marginals.

## Table 13: Unrestricted Model - Wage Class (1974)

| Parameter | Marginal | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\rho_1$ | | $\rho_2$ | | $\xi_1$ | | $\xi_2$ | |
| Constant | 0.7025 | ( 10.8281) | -0.7153 | (-11.0466) | 0.7790 | ( 11.3201) | -0.7829 | (-11.2551) |
| Father Age | -0.0604 | ( -5.3298) | -0.0491 | ( -4.1184) | 0.0103 | ( 0.8803) | -0.0031 | ( -0.2556) |
| O Level | — | ( — ) | — | ( — ) | 0.1923 | ( 1.1480) | 0.0386 | ( 0.2096) |
| A Level | — | ( — ) | — | ( — ) | 0.6051 | ( 2.0390) | 0.5021 | ( 1.9052) |
| High. Educ. | — | ( — ) | — | ( — ) | 0.7753 | ( 2.7621) | 0.7313 | ( 2.9104) |
| C. Skills 7 | — | ( — ) | — | ( — ) | 0.0499 | ( 0.6402) | 0.1655 | ( 1.9331) |
| C. Skills 11 | — | ( — ) | — | ( — ) | 0.0086 | ( 0.0813) | 0.0760 | ( 0.7121) |
| C. Skills 16 | — | ( — ) | — | ( — ) | 0.3645 | ( 3.2359) | 0.2502 | ( 2.0866) |
| NC. Skills 7 (1st) | — | ( — ) | — | ( — ) | -0.0552 | ( -0.6698) | -0.1992 | ( -2.0448) |
| NC. Skills 7 (2nd) | — | ( — ) | — | ( — ) | 0.0268 | ( 0.3110) | 0.1316 | ( 1.4177) |
| NC. Skills 11 (1st) | — | ( — ) | — | ( — ) | 0.0434 | ( 0.5449) | 0.1690 | ( 1.9518) |
| NC. Skills 11 (2nd) | — | ( — ) | — | ( — ) | -0.2199 | ( -2.7225) | -0.3527 | ( -3.5865) |
| Parameter | Odds | | | | | | | |
| | $\tau_{11}$ | | $\tau_{12}$ | | $\tau_{21}$ | | $\tau_{22}$ | |
| Constant | 0.1589 | ( 0.8712) | 0.1128 | ( 0.5894) | 0.5186 | ( 2.6211) | 0.3566 | ( 1.9529) |
| Father Age | -0.0529 | ( -1.7568) | -0.0502 | ( -1.5437) | 0.0973 | ( 2.8043) | 0.0001 | ( 0.0024) |
| O Level | 0.4536 | ( 1.0696) | 0.6547 | ( 1.3205) | 0.3801 | ( 0.8438) | -0.4857 | ( -1.0260) |
| A Level | 2.3471 | ( 2.4014) | 1.7889 | ( 2.5130) | -2.8069 | ( -2.8402) | -2.1742 | ( -3.0755) |
| High. Educ. | 0.6454 | ( 0.9285) | 0.6428 | ( 0.9749) | -0.2310 | ( -0.3092) | -0.3686 | ( -0.5742) |
| C. Skills 7 | 0.4129 | ( 2.0222) | 0.2796 | ( 1.1973) | -0.1209 | ( -0.5801) | -0.3467 | ( -1.5591) |
| C. Skills 11 | -0.5185 | ( -1.8872) | -0.2244 | ( -0.7904) | 0.6906 | ( 2.3944) | 0.1713 | ( 0.6224) |
| C. Skills 16 | 0.1640 | ( 0.5558) | 0.1466 | ( 0.4545) | -0.5773 | ( -1.8904) | -0.2207 | ( -0.7104) |
| NC. Skills 7 (1st) | 0.0837 | ( 0.3861) | 0.1728 | ( 0.6398) | -0.0475 | ( -0.2135) | -0.2570 | ( -1.0288) |
| NC. Skills 7 (2nd) | 0.2386 | ( 1.0826) | 0.0087 | ( 0.0363) | -0.0420 | ( -0.1831) | -0.4537 | ( -1.8804) |
| NC. Skills 11 (1st) | 0.1253 | ( 0.5918) | 0.4016 | ( 1.5518) | 0.4297 | ( 1.7630) | 0.0807 | ( 0.3501) |
| NC. Skills 11 (2nd) | -0.0871 | ( -0.4103) | 0.1781 | ( 0.6943) | 0.1454 | ( 0.6697) | -0.3068 | ( -1.1944) |

Columns 1 and 2 correspond to the father's marginals.