# High school types, academic performance and early labor market outcomes

Lorenzo Cappellari
Università del Piemonte Orientale, CHILD and IZA

March 2004

Correspondence:
Lorenzo Cappellari
Dipartimento di Scienze Economiche e Metodi Quantitativi
Università del Piemonte Orientale
Via Perrone 18, 28100, Novara, Italy
Email: lorenzo.cappellari@eco.unipmn.it

# High school types, academic performance and early labor market outcomes

Abstract
Using micro data on the 1995 cohort of Italian high school graduates, this paper studies the relationship between the type of high school attended (general *versus* technical; private *versus* public) and indicators of subsequent performance. Simultaneity issues that potentially bias this type of exercise are tackled by instrumental variables. Results indicate that the type of high school attended greatly depends upon the family of origin and prior school performance. General high schools are found to increase the probability of transition to university and to improve performance once at university. On the other hand, private high schools appear to be associated with lower academic performance. Technical schools improve the quality of the school-to-work transition, both in terms of participation and employment probabilities.

Keywords: high school types; academic and economic performance; endogeneity
JEL codes: I21, J24, C35

## 1. Introduction

Reforms at any level of the Italian educational system are the subject of an ongoing debate, recently stimulated by policy interventions on secondary high schools. In March 2003 the Italian parliament approved a law of reform of the high school system (law 53/2003) which introduces a new vocational track that includes class work in the first two years and a mixture of class work and on-the-job training in the last couple of years, plus an additional fifth year for those wishing to enroll into university. The vocational track will parallel the ones of general and technical/professional high schools, based on five years of class work and with free access to university at the end. The new law warrants the possibility of transition between tracks. Proponents of the new system emphasize the role that labor demand will play, via training schemes, in shaping the type of skills that vocational track students would acquire, therefore facilitating their school-to-work transitions. Opponents stress the risks of social segmentation that such system could ingenerate through tracks separation, doubt the feasibility of transitions between tracks, and question the usefulness of learning job-specific skills during school years in a world of fast skill obsolescence.

A second issue debated –at any educational level– is school provision. Since 2000, school vouchers started being introduced by some regional governments, with the aim of helping households in choosing their preferred type of school, whether private or public, and in September 2003 vouchers programmes have been extended nationwide. Given the difference in costs entailed by the two types of provision, vouchers lower the price differential between private schools and public ones, thereby shifting some demand for schooling from the latter to the former. Both freedom of choice and increases in private and public school quality –the latter brought about by greater competition in the education market– can be found among the principles inspiring vouchers programmes. Opponents to these policies question the ability of private schools in providing quality education and argue for employing public funds for improving the effectiveness of public schools.[1]

The policy debate prompts several questions. To what extent the choice of a high school track affects later individual outcomes in terms of access to higher education, academic performance, and the school-to-work transition? How does the performance of private school students compare with that of public schools ones? Do differences in performance across high school types reflect a causal effect or are they due to endogenous sorting of students (due to unobservable ability or unobservable family characteristics, for example) across school types? This paper aims at providing answers to the above questions. Using survey data on a cohort of Italian high school graduates, models of the impact of high school types (general *versus* technical; private *versus* public) on the transition to university, academic performance, labor market participation, employment probabilities and wages are estimated. In order to unravel the causal link between school types and subsequent outcomes, great attention is devoted to endogenous selection issues, which are tackled by employing instrumental variables within systems of simultaneous limited dependent variables equations.

The literature on the evaluation of school type effects on measures of performance has grown rapidly over the past decade. For example, the impact of catholic schools on academic performance has received considerable attention in the United States. Evans and Schwab (1995) study the effect of catholic schools on high school completion and college enrolment probabilities; they highlight the endogeneity issues that can arise from self-selection of students into catholic schools and use instrumental variables to identify the effect of catholic school attendance on measures of academic success, concluding that catholic schools raise subsequent educational outcomes. Neal (1997) shows that as long as

[1] See Neal (2002) for a review of the school vouchers literature.

households chose school types on the basis of expected outcomes, and these depend upon unobserved factors that also influence school choice, estimating the impact of school types on outcomes is subject to endogeneity issues. He uses area-level measures of catholic schools availability as instruments for school choices, showing that the benefits of catholic schools are confined to urban minorities, suggesting that they could act as a remedy in the presence of low quality public schools. An instrumental variable procedure is employed by Figlio and Stone (1999) to assess the effect of religious and non-religious private schools on educational outcomes, finding that, in general, only the former increase individual outputs relative to public schools. The methodological approach developed by Altonji et al. (2000) is instead based on the use observable information as a way of reducing endogeneity bias; they find that catholic schools are effective in favoring high school completion, while the effect on transition to college is less evident. Recently, researchers' interest on the effect of school types have also started spreading onto other spheres of human life: an example is Figlio and Ludwig (2000), who look at the effect of catholic school attendance on youths' crime, drug use and sexual activity, finding that catholic schools are effective in reducing all three.

While catholic/private *versus* public schools effects have been extensively researched, the general *versus* technical or vocational high school divide is less explored, perhaps not surprisingly, given that US high schools are characterized by greater uniformity under this respect compared to other countries, Italy included. Evidence for Germany indicates that family income plays a limited role in determining the probability of attending general high schools, other factors such as parental education being more relevant (see, Buchel et al., 2001, and Jenkins and Schluter, 2002).[2] For France, Margolis and Simonnet (2002) show

[2] A strong association between secondary school track choice and parental education in Germany has been found also by Dustmann (2001), but without controlling for parental income. For Britain, Ermisch and

that technical high school graduates outperform general high schools ones in the school-to-work transition, thanks to the more effective labour market networks they can access.

Given their current political relevance, it is not surprising that the interest surrounding economics of education topics in Italy is increasing at a fast rate. The main issue emerging from the Italian literature is a strong intergenerational persistence in educational achievement. Checchi et al. (1999) compare intergenerational mobility in incomes and schooling attainment between Italy and the US, finding larger persistence in the former case, and argue that the Italian schooling system might have failed to provide low-income families with the incentives to invest in the human capital of their off-springs. Besides family of origin, also high school types appear to be strongly associated to educational outcomes. Bertola and Checchi (2002) study a sample of university students from the University of Milan and find that those coming from general high schools score better than otherwise comparable students on a range of performance indicators. They also consider the differences in academic performance between public and private school students, finding that public schools are associated to better performances, followed by religious private schools and lay private schools. The importance of high school types for academic performance is confirmed by Boero et al. (2001) who study a sample of Italian college graduates and show that the final graduation mark drops significantly if one compares general and technical high school graduates. The issue of public/private school choices is analyzed in Checchi and Jappelli (2002) using subjective assessment of public school quality: they do not find evidence of any quality differential in favor of private school as determinant of the school choice. A theoretical perspective on the optimal school design is provided by Brunello and Giannini (2000), showing that the desirability of educational stratification can not be unambiguously judged from an efficiency point of view.

---

Francesconi (2000) report that while parental education raises educational outcomes for both female and male children, parental income has an impact only for females.

This paper extends the bivariate probit type of models employed by Evans and Schwab (1995) and Neal (1997), and estimates the impact of school types on academic and economic outcomes. There are three novelties relative to the models of previous studies. First, allowance is made for the presence of two —rather than one— endogenous and possibly correlated school types, adapting the framework applied to US catholic schools to the two schooling dimensions relevant to the Italian debate.[3] Second, incidental truncation is taken into account; such an issue arises when an outcome of interest is observable depending upon the realization of another outcome (e.g. going to college is observable conditionally on finishing high school) and overlooking it might induce endogenous selectivity issues. Third, interactions between academic and economic spheres are explicitly modeled. If academic and economic participation are interrelated and high school types affect both, then evaluating the impact of high schools in one sphere should take indirect – i.e. through the other sphere—effects into account, an aspect that has not received attention so far. In particular, in this paper not only academic (economic) performance is conditioned upon economic (academic) behavior, but also unobserved correlation is allowed across the two, thus tackling the endogeneity issues pointed out by the literature on the effects of working while at school on school performance (Stinebrickner and Stinebrickner, 2003).

Estimates show that school types greatly depend upon the family of origin and prior school performance. General high schools are found to increase the probability of transition to university and to reduce labor market participation, while no direct influence can be detected on academic performance. On the other hand, private high schools and labor market participation reduce academic performance. Technical schools improve the quality

[3] Light and Strayer (2000) model the impact of college types on college completion probabilities allowing for four unordered and mutually exclusive realizations of a single treatment, while in the present paper attention is placed on two binary non-mutually exclusive treatments.

of the school-to-work transition, both in terms of participation and employment probabilities, whereas college enrolment and general schools reduce employability.

## 2. Data and descriptive patterns of high school types and subsequent performances

The data used in this paper originate from the "1998 Survey on the school and work experiences of 1995 high school graduates", a cross-sectional sample of 18,843 high school leavers interviewed by the National Statistical Office (ISTAT) three years after graduation. The sample represents approximately 4 percent of the population of Italian high school graduates of 1995 and contains a wide range of information on the high school curriculum and on post high school experiences, either in the tertiary education system – including university – and the labor market. In addition, information on personal characteristics and family background is available.[4]

The Italian high school system as of 1995 may be broadly described by three types of schools: general ('licei'), technical ('istituti tecnici e professionali') and teaching schools ('istituti magistrali'). The first two types are based on a five-years curriculum, at the end of which students can freely chose to enroll in university.[5] General high schools are academic-oriented; technical ones, on the other hand, prepare students for white collar or skilled blue collar careers. The third type is specifically aimed at training primary school teachers and is structured into a four years curriculum, plus an additional year required to students who wish to go to university. In the present paper this latter group (which represented approximately 12 percent of the sample originally available) has been excluded from the estimation sample, due to the difference in the structure of the curriculum compared to the

[4] At the time of interview, compulsory education ended after 8 years, and high school enrolment was not compulsory. Law 53/2003 establishes that : "…everyone has the right to acquire education or training for at least 12 years, or, in any event, until the attainment of a degree before the age of 18, either within the educational system or the education /training one…" (article 2, own translation).

[5] Before 1969 graduates from technical or professional high schools were required to pass an additional exam before they could enrol into university.

other two types. Thence, the choice of high school types analyzed here focuses on the general *versus* technical divide.

As for the other dimension of the Italian debate on educational policies, i.e. the public *versus* private provision of schools, the survey enables identification of the type of school attended by asking respondents if the school they enrolled in after junior high school was private or public, and if they made any transition from private to public school (and *vice versa*) during high school. Differently from the work of Bertola and Checchi (2002), the data do not contain information on the confessional nature of private schools.

The distribution of the type of school attended in the estimation sample is reported in the first panel of Table 1. The estimation sample excludes those who are employed and started their job while at high school (3.6 percent of the original sample), since their post-graduation choices might not be comparable with those of the rest of the sample. In addition, the estimation sample for the analysis of academic performance excludes students who enrolled into short university diplomas (2.2 percent of the original sample) because it was not possible to compute reliable indicators of academic performance due to small cells (see below for the definition of performance used), and those who did not report the information necessary for computing academic performance (0.05 percent). Graduates from general high schools represent 31 percent of the sample, whereas 12 percent of it graduated from a private high school. The two dimension of high school choices are strongly correlated. The probability of having a private school diploma is 9 percent among graduates from technical schools and 19 percent among students from the licei. The other conditional frequencies presented in the table show that the incidence of students from general high schools is some 17 percent larger among private school graduates than in the overall sample. One explanation of the positive association between the two variables can be found in the larger supply of private education in the general high schools system: official statistics for 1989 indicate that while 14 percent of technical schools were private, the proportion was 30 percent for general high schools. A second reason has to do with family resources: both private and general schools entail larger costs than public or technical ones, either in term of fees and expected opportunity costs.

The middle panel of Table 1 considers the links between school types and subsequent academic performance. Two are the dimensions of academic performance taken into account, college enrolment and exams passing speed. Each exam can be taken several times during the year, and there are no restrictions on the number of times an exam can be taken before being passed. Such a system creates great dispersion in the number of exams passed per year, and exam passing speed depend to a large extent on individual ability and commitment. To exploit such variation, for each of eight broad disciplinary areas in which university courses can be grouped (sciences; medicine; engineering; architecture; economics, business administration and statistics; political sciences; law; and the humanities), the median number of exams passed per year has been computed; the speed indicator is a dummy for being above the median in the subject-specific distribution of the numbers of exams passed per year. Clearly, this is an imperfect performance indicator, since it does not take the marks obtained into account.[6] However, no information on marks is reported in the survey. Moreover, one of the main concerns inspiring the recent reform of university degrees has been the long duration of university studies in the old system, so that studying the impact of high school types on the speed in passing exams will yield insights on an issue at the core of Italian educational policies.

Table 1 shows that 40 percent of high school graduates attend university three years after graduation. The proportion halves among technical school graduates and doubles

[6] The performance indicator used by Bertola and Checchi (2002) considers both speed and the marks obtained in each exam. Exams can be passed with marks that range from 18 to 30 out of 30, and students can chose whether to accept them or not, refusal implying that the exam has to be taken again. The average exams mark matters in determining the final graduation mark.

among general school graduates, revealing a strong link between type of degree held and transition to university. Transitions to university are more frequent among private school graduates than they are among public schools ones, although the magnitude of the link is less evident than in the previous case. Due to discontinuities in the distribution of exam passed per year, the sample proportion of cases above the median (the 'fast tracks') is 45 percent. Graduating from a general high school changes this proportion only slightly, by 4 percent; on the other hand, the proportion of fast individuals among technical school graduates is 38 percent. Evidence on the private/public school divide shows little variation in speed.

The second dimension of students performance that is modeled in this paper are early labour market outcomes, in particular participation, employment (conditional on participation) and low pay (conditional on employment). The ISTAT questionnaire contains a wide array of information on both job search strategies and the characteristics of the current job. As for academic performance, the sample for the labor market model excludes graduates from teaching schools and those who started their current jobs while at high school. In addition, after testing for endogenous selection and failing to reject the null hypothesis of exogeneity, those who work as self-employed (4.6 percent of the original sample) have been excluded from the estimation sample. Also, excluded are employees with missing wages (an additional 2.6 percent of the original sample). Therefore, the estimation sample for the analysis of economic performance differs from the one used for analyzing academic performance, and consists of 14,420 observations.[7]

The lower panel of Table 1 illustrates the variation of labor market outcomes in the sample. Participation is defined has having an occupation of any kind or being on job search. The employment status considered refers to regular employment; unemployed

[7] The distribution of high school types in this sample is very similar to the one shown in Table 1, Panel (A).

individuals having done some work in the week prior to interview or those on occasional or seasonal jobs are not counted as employed. Partly, such a choice is driven by the data, since no information on job attributes is collected for seasonal or occasional employment; in addition, it allows to focus on stable employment, which is probably more relevant from a policy perspective. Low pay is defined as the bottom quartile of the sample distribution of net hourly pay, i.e. the low pay threshold is defined relatively to the group of high school graduates. 68 percent of the sample participates to the labor market. The proportion rises to 81 percent among technical school graduates and falls to 40 percent for general school ones, clearly indicating a difference in the propensity for economic activity between the two educational tracks. A slight difference in job search propensities can also be observed between public and private schools graduates. Of those who search, 42 percent are observed in employment. This rather low figure depends upon the definition of employment adopted: if one included also seasonal employment, occasional employment and the unemployed reporting hours of works in the week prior to interview, the employment rate would be 68 percent. The employment rate varies depending upon the type of high school attended, passing from the 49 percent of technical schools graduates to the 16 percent of licei students; some variability is evident also between private and public schools. Finally, some associations between school choices and the probability of earning low wages can be detected, although they are less evident compared to the case of employment.

## 3. High school types and academic performance

This section develops a framework for estimating the impact of school types on academic performance while allowing for the presence of spurious correlations. Endogeneity can arise from selection into school types, as long as unobserved ability and unobserved parental background might influence both choices and outcomes; as shown in Neal (1997) such a circumstance can be relevant if school choices are driven by expected returns on education. Moreover, one of the outcomes of interest –exams passing speed– can only be observed conditionally on university attendance, i.e. on a sample that is potentially subject to endogenous selection. Finally, high school types may well influence labor market participation which, in turn, can affect academic performance, forming an indirect –endogenous– channel through which school types impact on academic outcomes.

3.1 The model of academic performance

Let observations in the estimation sample be indexed by $i$, $i$=1…$n$, and let the latent net benefits to graduate from general schools ($g^*_i$) and private schools ($p^*_i$) be linear functions of individual characteristics:

$$g^*_i=\beta_g' x_{gi}+\varepsilon_{gi} \quad (1)$$

$$p^*_i=\beta_p' x_{pi}+\varepsilon_{pi} \quad (2)$$

where the $x$s are vectors of observable attributes such as gender, parental background and indicators of academic performance at the junior high stage, the $\beta$s associated parameter vectors to be estimated and the $\varepsilon$s error terms assumed to be distributed as standard normal variates. When net benefits are positive, individuals are observed to graduate from general or private schools; let $g_i$=I($g^*_i$>0) and $p_i$=I($p^*_i$>0) be dummy variables indicating the two events, the indicator function I( ) taking value one whenever its argument is true and zero otherwise.

Next, assume that, conditional on high school types, individuals chose whether to enrol into college and whether to participate to the labor market depending upon unobserved net benefits $u^*_i$ and $r^*_i$. In order to capture the patterns observed in the data, the two choices are treated as not mutually exclusive:[8]

$$u^*_i = \beta_u' z_{ui} + \delta_g g_i + \delta_p p_i + \varepsilon_{ui} \quad (3)$$

$$r^*_i = \beta_r' z_{ri} + \eta_g g_i + \eta_p p_i + \varepsilon_{ri} \quad (4)$$

where the error terms $\varepsilon$s are assumed to be distributed as standard normal variates, the $z$s are vectors of characteristics with associated coefficient vectors $\beta$s, and the coefficients $\delta$s and $\eta$s index the effects of school types on academic and economic participation; let $u_i$=I($u^*_i$>0) and $r_i$=I($r^*_i$>0) be dummy variables indicating the occurrence of academic and economic participation. Note that the two decisions are allowed to occur simultaneously but are not conditioned one upon the other; dependency between the two equations is allowed in a reduced-form fashion through correlation of the unobservables, see below.

Finally, for those attending university academic performance can be observed in terms of exams passing speed. Let $a_i$, the average number of exams passed per year by individual $i$, depend upon personal attributes $m_s$, labor market participation and high school types according to the following relationship:

$$s^*_i \equiv h(a_i) = \beta_s' m_{si} + \gamma_r r_i + \gamma_g g_i + \gamma_p p_i + \varepsilon_{si} \text{ ; observed if } u_i=1 \quad (5)$$

where $h(\ )$ is a suitable monotonic unspecified transformation such as the error tem of (5) is distributed as standard normal, while the relationship is truncated when $u_i$=0. An individual is classified to be 'fast' whenever she lies in the upper half of the subject-specific

[8] 16 percent of the estimation sample is active on both the academic and economic side, whereas complete inactivity occurs for 4 percent of the sample; remaining proportions are 25 percent (only academic participation) and 55 percent (only economic one).

distribution of $a_i$; let $s_i$=I($a_i$>$\mu$) signal that event, where $\mu$ is the median of the subject-specific distribution of exam passed per year.

The vector of error terms $\varepsilon^A_i$=($\varepsilon_{gi}$, $\varepsilon_{pi}$, $\varepsilon_{ui}$, $\varepsilon_{ri}$, $\varepsilon_{si}$) is assumed to follow the five-variate normal distribution: $\varepsilon^A_i$~$N_5$(0,$\Omega$), where the correlation matrix $\Omega$ has unit diagonal elements and extra-diagonal elements equal to $\rho_{jk}$, $j,k=g,p,u,r,s$. Overall, the model specified is a five-variate probit with endogenous truncation of one equation.[9] Estimation of cross-equation correlation coefficients allows to control for unobserved heterogeneity, therefore eliminating the issues of endogeneity discussed above.

Improving parameter identification requires valid "instruments", i.e. variables that can be excluded from the outcome equations but significantly affect high school types, selection into university or labor market participation. The instrument used for general high school is a dummy for having at least a grandparent with a high school degree, i.e. it is assumed that second order intergenerational transmission of educational tastes and constraints loses relevance in shaping students' choices as they age from 14 to 19, after its impact at age 14 has been controlled for. The private school equation includes as instruments the same variable, plus dummies for the mother being retired or housewife when the individual was 14; since private schooling can be seen, in many cases, as a purchase of child care, especially for the afternoon hours when there is no activity in public high schools, the two dummies, that indicate the availability of parental time, are intended to proxy the need of purchasing those services. Finally, instrumenting selection into university or labor market participation requires variables that affect post-high school choices but have no residual impact on exams speed; the number of siblings and parental occupations –i.e. proxies of resources available to individuals' choices—are the ones used

[9] Computation of multi-variate normal distributions is performed by simulation, in particular by applying the so-called GHK simulator.

here.[10] The validity of the identification strategy laid out above can be tested parametrically using functional form as identifying restriction.

Besides the exclusion restrictions discussed above and the endogenous variables, the conditioning set of the model is as follows: equations (1) and (2) include gender, birth cohort and regional dummies, number of siblings, parental education and occupation, junior high school performance indicators and age of high school enrolment; equations (3) and (4) add secondary high school performance indicators; equation (5) adds the dummies for mother housewife or retired when the child was 14 and grandparents education.

3.2 Results

Table 2 presents some results from the estimation of the academic performance model, while the full set of estimates is provided in Table A1. Panel B of the table —lower left corner— shows results from tests of significance of the instruments, which indicate that the data support the choice of instruments made: high school types are significantly affected by grandparents' education and by the mother activity status when the individual was 14, but have no residual impact on the transition to university or labor market participation; similarly, the presence and number of siblings and parental occupation affect the probability of attending university at the date of interview, but have no impact on exam passing speed.

The right column of panel (B) provides estimates of the errors correlation structure. Even net of observable attributes, the two high school types are positively correlated, as noted when commenting Table 1; differently from there, however, in this section controls are made for family resources, so it is likely that the positive coefficient reflects a larger supply of private education among general schools than among technical ones. A second

[10] Household incomes are not available in the data; however, the survey contains rather detailed information on parental education and occupation which partly compensates for the unavailability of incomes.

relevant source of unobserved heterogeneity arises considering the link between college enrolment and labor market participation, which is negative and precisely estimated, showing that they tend to be substitutes in the choice set of high school graduates. Moreover, private school and exams speed propensities are positively correlated, an outcome that might either signal larger ability compared to public schools, or a lower selectivity in accepting exams marks (see footnote 6); considering that, as discussed below, school performances indicators prior to high school are lower among private school students compared to public school ones, the second might be the relevant explanation. On the other hand, there is no endogenous selection into college relating to estimation of the exams speed equation, the coefficient linking the two equations being small and not statistically significant. Finally, unobservables of the labor market participation and exam speed equations are positively correlated: as pointed out by, e.g., Stinebrickner and Stinebrickner (2003), such an outcome might reflect unobserved motivation. The table reports results from tests of the ignorability of high school types equations, i.e. tests of the null hypothesis that all correlation coefficients involving a given school type equation are simultaneously equal two zero; as shown in the table, endogeneity of either process should not be ignored when estimating the model. A test of significance of the overall correlation structure also points towards need of a simultaneous equation framework.

Panel (A) of the table presents in column (1) the estimated 'marginal effects' (evaluated at sample means of explanatory variables) of school types on academic performance and labor market participation. General high schools increase the probability of college enrolment and reduce that of labor market participation. Given that factors such as ability (proxied by previous school performance) parental background and unobserved heterogeneity have been controlled for in the model, these effects can be given a causal interpretation, for example based on the skill content of general school curricula. On the other hand, the private/public school divide appears to be irrelevant in shaping behaviors in the academic and economic spheres. Looking at the exams speed equation estimates, while no significant association can be detected concerning general high schools, negative and sizeable effects characterize private schools and labor market participation. In the latter case, the negative effect is explained by the fact that economic activity drags resources (in terms of time and effort) from exams preparation, reducing the probability of success. Explanations are less evident in the case of private schools, while the result suggests that they are less effective, compared to public ones, in favoring one dimension of college performance that received considerable attention in the recent reform of university degrees.

Columns (2) and (3) of panel (A) offer, for comparative purposes, effects estimated using restricted versions of the model. Estimates in column (2) are obtained omitting controls for unobserved heterogeneity, i.e. by separately estimating each equation. Generally, these models overestimate the effects of interest while, as could be expected, standard errors are lower compared to the model with unobserved heterogeneity. The overestimation of effects is consistent with the presence of positive selection effects; for example, general high school students tend to be of a type more inclined to college participation compared to technical schools ones, see the sign of the correlation coefficient reported in panel (B), so that once that source of heterogeneity has been controlled for, the estimated impact diminishes. The loss of estimates precision induced by the controls for unobserved heterogeneity is, in some cases, crucial, since some of the effects dwindle to non significance. It is also evident that ignoring endogeneity leads to underestimate the negative impact of economic participation on academic performance. Column (3) of panel (A) yields insights into the consequences of ignoring economic behavior when studying academic performance, i.e. presents estimates of the model deprived of the participation equation, and with participation excluded from the conditioning set of the speed equation.

The main difference with column (1) emerges considering the effect of general schools on exams speed, which is now positive and precisely estimated. Given that general schools reduce labor market participation which in turn reduces speed, omitting to control for participation leads to conclude that general schools have a positive effect on speed. Put another way, school types can impact on academic outcomes either directly or indirectly, i.e. through their impact on economic behavior, and the effects presented in column (3) are a convolution of the two. A second consequence of omitting the participation equation from the model has to do with selectivity effects. The estimated correlation coefficient between unobservables of enrolment and speed equations is 0.384 (0.101) for the model of column (3), indicating positive selection of 'fast' individuals into college. Given the evidence from column (1), such a result can instead be ascribed to the fact that college students tend not to participate to the labour market and participation worsens college performance.

Table A1 provides the full set of estimates of the academic performance model. Looking at the determinants of high school types, we can see that the presence of a favorable family background affects both choices significantly. In particular, while general high school graduates tend to come from families where one or both parents, are in managerial or professional occupations, private schools are chosen more frequently when the parents are self-employed or entrepreneurs. Also, teachers' children are likely to graduate from general and public high schools. Parental education positively and strongly influences both probabilities. The indicator of academic performance at junior high school is strongly and positively associated to the probability of graduating from a general high school. Much smaller effects, and in the opposite direction, can be observed for private high schools. Overall, results appear to confirm previous findings from Bertola and Checchi (2002) on the fact that private schools attract students from richer families and with relatively lower ability compared to public schools ones, whereas general high schools attract higher performance individuals from richer families.

The remaining columns in Table A1 illustrate that both parental background and indicators of previous school performance are associated with outcomes; in particular, a favorable parental background, in terms of both education and occupation, and good performances at previous educational levels, tend to be associated with high academic participation, low economic participation, and high exams passing speed.

## 4. High school choices and early labour market outcomes

The raw correlations presented in Section 2 indicate the existence of a strong link between the type of high school attended and behaviors at labor market entry; in this section the extent to which those associations can be imputed a causal interpretation is investigated. As in Section 3, endogeneity of high school types is controlled for by explicitly modeling their determinants and by allowing any residual component to freely correlate with the ones of the outcomes of interest. Endogenous selection issues are also investigated; differently from Section 3, in the present section a double selectivity process is modeled: selection into labor market participation (potentially endogenous for the estimation of employment probabilities) and selection into employment (potentially endogenous for the estimation of low pay probabilities). Finally, the impact of high school type on economic outcomes is evaluated while controlling for endogenous academic participation.

### 4.1 The model of early labour market outcomes

The probability of graduating from a high school of a given type is modeled in the same fashion of Section 3, but on a partly different sample (see Section 2 for the differences in sample selection rules between this and the previous section); the same remark applies to university enrolment and labor market participation. Therefore, the first four equations of the model are equations (1) to (4) of Section 3.

Depending upon the outcomes of the labor market participation equation, the employment status at the date of interview can be observed (recall from Section 2 that self-employment is excluded from the analysis and that seasonal or occasional jobs are considered as unemployment). Assume that it represents some unobserved employment propensity $e^*_i$, which depends upon personal attributes $m_e$ , school choices and academic participation:

$$e^*_i=\beta_e' m_{ei}+\theta_u u_i+\theta_g g_i+\theta_p p_i+\varepsilon_{ei}\ ; \quad \textit{observed if } r_i=1 \qquad (6)$$

and let $e_i$=I($e^*_i$>0) indicate whether an individual is employed or not at the interview date. The employment process is truncated for those who do not participate in the labor market after graduation.

Besides employment probabilities, another relevant dimension of the school-to-work transition is the wage level that individuals obtain on entry into the labor market. Clearly, entry levels are not fully informative of future earnings prospects, since the correlation between initial and later earnings varies depending upon which model better describes labor market functioning. Here the focus is placed on the probability of low pay for two reasons. First, from a substantive point of view, in recent years low pay has emerged as a policy issue in many countries, and research on this topic stresses that as long as there are discontinuities in the earnings process across the distribution, the relevance of earnings determinants might vary according to the earnings quantiles considered. Second, from a modeling point of view, estimating an earnings equation in the multivariate probit context of this paper would have required to assume normality or log-normality of earnings, whereas by focusing on low pay probabilities, normality is required only up to any unspecified transformation of earnings (as will be shown below). Finally, it has been chosen to derive the low pay cut-off from the sample earnings distribution, i.e. to focus on individual earnings relative to the cohort of high school graduates.

Let individual earnings $w_i$ depend upon personal and job attributes $q_i$, high school choices and academic participation according to the following equation:

$$f(w_i)=\psi' q_i + \omega_u u_i+\omega_g g_i+\omega_p p_i+v_i \;;\; \textit{observed if } e_i=1 \qquad (7.a)$$

where $f(\ )$ is an increasing unspecified transformation such as the error term $v_i$ is distributed as standard normal; let also $p_{25}$ indicate the first quartile of the earnings distribution. Note that the earnings distribution is censored upon employment (and, *a fortiori*, upon participation). An individual is counted as low paid whenever $w_i < p_{25}$, i.e. when $f(w_i)< f(p_{25})$. By subtracting each side of (7.a) from $f(p_{25})$, the earnings process can be rewritten as:

$$l^*_i=\beta_l' q_i+\varphi_u u_i+\varphi_g g_i+\varphi_p p_i+\varepsilon_{li} \;;\; \textit{observed if } e_i=1 \qquad (7.b)$$

where $l^*_i\equiv f(p_{25})-f(w_i)$, the first element of $\beta_l$ subsumes the difference between $f(p_{25})$ and the constant term in $\psi$, remaining coefficients parameterize the association between covariates and low pay propensities and the error term is distributed as standard normal. Observing an individual in low pay means that the corresponding latent variable $l^*_i$ is positive; let $l_i$=I($l^*_i$ >0) be a dichotomous indicator of the low pay event.

As in the case of academic performance, endogeneity of high school types and (double nested) selection is tackled by means of a multivariate probit model, which emerges if one assumes that the vector of error terms $\varepsilon^L{}_i=(\varepsilon_{gi}, \varepsilon_{pi}, \varepsilon_{ri}, \varepsilon_{ui}, \varepsilon_{ei}, \varepsilon_{li})$ follows the six-variate normal distribution: $\varepsilon^L_i \sim N_6(0,\Sigma)$, where the correlation matrix $\Sigma$ has unit diagonal elements and extra-diagonal elements equal to $\lambda_{jk}$, $j,k=g,p,r,u,e,l$. Overall, the model specified is a six-variate probit with nested truncation of two equations. Test of the endogeneity of high school types and selection can be performed by testing the significance of the cross-equation correlation coefficients of unobservables.

Specification of school types, university enrolment and labor market participation equations is the one adopted in Section 3. Variables that are –plausibly– not taken into account by employers when screening job candidates might serve for the purpose of identifying the employment equation. For example, if firms screen job applicants mainly on the basis of the type of diploma obtained (general *versus* technical), school marks indicators may serve as instruments for employment probabilities. Empirically, it was found that high school marks, not junior high schools ones, were insignificant in shifting employment probabilities, suggesting that employers do not value the informative content of marks, but limit themselves at considering school types.[11] In addition to that, also parental education and number of siblings (significant in explaining college enrolment and participation) were found insignificant in the employment equation and excluded from the preferred specification. Finally, indicators of parental occupations and junior high school marks could be excluded from the low pay equation after their impact on schooling choices and employment probabilities has been controlled for. Job characteristics (occupation, type of contract, industry, firm size and tenure), mode of search of the current job, high school marks, parental education and number of siblings complete specification of the low pay equation.

[11] Boero et al. (2001) find that high school and university marks have no explanatory power in wage regressions for their sample of Italian college graduates.

4.2 Results

Table 3 reports selected results obtained estimating the economic performance model, the full set of estimates for labor market outcomes being provided in Table A.2. As for the academic performance model, the adopted identification strategy is supported by the data – see the p-values for instruments significance in the various equations reported at the bottom of the table. Looking at correlation coefficients already present in the academic performance model, results tend to reproduce the patterns emerged in Section 3. In addition, a significant dimension of heterogeneity is the positive correlation between unobservables in the general high school and employment equations. Furthermore, low pay propensities are larger among those who participate to the labor market compared to the economic inactive, evidence which might reflect lower reservation wages or the selection of low-productivity individuals into the pool of participants; the opposite is true for academic participation, although the estimate is not very precise in this case.

Marginal effects of high school types and college enrolment on employment and low pay probabilities are reported in panel (A) of the table (estimated effects on participation – similar to those obtained in Section 3—are reported in Table A.2, while marginal effects for equations (1) – (3) –again similar to those shown in Table A1—are not reported). General schools, private schools and college enrolment all reduce employment probabilities. Considering general schools, the negative causal impact could emerge from the skill content of their curriculum, which could be less valued by employers compared to technical schools. Alternatively, it could be the symptom of bad signaling, i.e. the labor market does not expect general school students to be engaged in job search activities and interprets their participation as a sign of low-ability. To the extent that the skill content argument does not apply to the effect of private schools, its negative sign can be interpreted resorting to the signaling explanation. Finally, the negative effect of academic participation can be interpreted saying that it drags resources from job search, similarly to what observed in Section 3 with the negative effect of economic participation on academic outcomes. No significant effects, on the other hand, can be detected on low pay probabilities.

Column (2) of panel (A) provides estimates obtained by assuming that all equations are independent. Major differences with column (1) relate to the effect of private and general schools on employment probabilities. Ignoring endogeneity leads to underestimate the negative impact of school type, reflecting the positive employment propensity characterizing both types (see the correlation coefficients between the equations).

The third column of panel (A) reports results from estimation of a restricted model that ignores academic participation. Estimates of high school type effects are sensitive to this restriction, although less than the estimates of the academic performance model were to the exclusion of the economic participation equation. In particular, the substantive change refers to the negative marginal effect of general schools on employment, which rises in absolute value by 5 percentage points when academic behavior is ignored, thereby picking up some of the negative effect of college enrolment on employment probabilities. Another estimate that is sensitive to the restriction is the correlation coefficient between participation and employment probabilities, which, in this case, has an estimate of -0.41 (0.079), suggesting the presence of endogenous selection of low-employability individuals into the pool of labor market participants. Given the evidence about the correlation structure of the full model, such a misleading conclusion can be ascribed to the omission of the college enrolment process.

The other estimates of employment and low pay determinants shown in Table A.2 tend to conform to the patterns one could expect *a priori*. For example, marked penalties on both fronts characterize southern youths. Also, job characteristics affect low pay probabilities in the expected direction. Finally, the job search mode matters for earnings levels and, in particular, those who found the job through informal networks have the largest chance to be low paid: either heterogeneous ability or heterogeneous job productivity across search modes can explain such an outcome.

## 6. Concluding remarks

The high school types chosen by young Italians largely depend on individual ability and family background. This paper has shown that better performance individual and those with favorable educational and occupational family background tend to select into general high schools. On the other hand, the probability of graduating from a private school rises with the availability of resources in the family of origin but decreases with school performance prior to high school.

The paper has provided evidence on the effects of high school types on subsequent academic performances. Graduating from a general high school substantially increases the probability that individuals go on to college rather than becoming active on the economic front. On the other hand, no shift in such probabilities due to private schools can be found in the data after a proper allowance is made for simultaneity issues. Once at university, the speed in passing exams – a crucial performance indicator for today's Italian students—is found to depend negatively on private schools and labor market participation.

Results suggest that controlling for economic behavior when measuring academic outcomes is relevant, omitting such controls inducing to impute some of its effect effects to school types. Put another way, school types impact on academic outcomes either directly and through their effects on economic behavior, and overlooking the latter might induce to consider any effect as a direct one. For example, when labor market participation is ignored, estimates show that general schools increase exams passing speed, an effect that disappears once economic behavior is taken into account: while in the first case one would conclude that general schools increase performance, in the second such a statement would not find support in the data.

The impact of school types also spreads to the economic sphere; in particular, general and private schools lower employment chances. In the case of general high schools, the evidence might reflect the skill content of the curricula, which is not attractive for employers compared to that of technical schools. An alternative interpretation is that school types act as ability signals and that general school students engaged in economic (rather than academic) activities are seen as low ability types. Finally, an employment penalty characterizes college students engaged in labor market participation.

The picture emerging from this paper extends previous findings for Italy. Family background plays a central role in determining the choice of high school types, while such choices have relevant effects on subsequent educational and economic paths. By deepening the separation between generalist and vocational tracks, the reform of secondary high education which will be implemented in the near future might have the effect of increasing the role of parental backgrounds in shaping individuals' lives. An effective functioning of the mechanisms designed in order to guarantee the 'equal dignity' of tracks –such as the possibility of track changes after the initial choice— appears as a crucial feature of the implementation phase for preventing intergenerational persistence and social segmentation to increase. As for the other topical dimension of the school debate, i.e. school provision, the negative effect of private schools estimated in the academic performance equation indicates that, on average, private schools are less effective than public ones in enhancing subsequent educational outcomes.

References


Altonji, G.J., Elder, T.E. and Taber, C.R. (2000): Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic School, National Bureau of Economic Research, Working Paper 7831.

Bertola,G. and Checchi,D. (2002): Sorting and Private Education in Italy, Centre for Economic Policy Research, discussion paper no. 3198.

Boero, G., McKnight, A., Naylor, R.A., and Smith, J. (2001): Graduates and Graduate Labour Markets in the UK and Italy, Centro Ricerche Economiche Nord Sud, Università di Cagliari, Contributo di Ricerca no. 11.

Brunello, G. and Giannini, M. (2000): Stratified or Comprehensive? The Economic Efficiency of School Design, Institute for Social and Economic Research, University of Essex, Working Paper 32.

Büchel, F., Frick, J. Krause, P. and Wagner, G. (2001): The impact of poverty on children's school attendance. Evidence from West Germany, in K. Vleminckx and T. Smeeding (eds.), Child Well-being, Child Poverty, and Child Policy in Modern Nations, The Policy Press, Bristol.

Checchi, D., Ichino, A. and Rustichini, A. (1999): More Equal but Less Mobile ? Intergenerational Mobility and Inequality in Italy and in the US, Journal of Public Economics, 74 (3), pp. 351-393.

Checchi, D. and Jappelli, T. (2002): School Choice and Quality, unpublished manuscript available at: http://www.economia.unimi.it/checchi/index.html.

Dustmann, C. (2001): Parental background, primary to secondary school transitions, and wages, Institute for the Study of Labor, Bonn, Discussion Paper 367.

Ermisch, J. and Francesconi, M. (2000): Educational Choice, Families, and Young People's Earnings, The Journal of Human Resources, XXXV (1), pp. 143-176.

Evans, W.N. and Schwab, R.M. (1995): Finishing High School and Starting College: Do Catholic Schools Make a Difference?, Quarterly Journal of Economics, 110(4), pp. 941-974.

Figlio, D.N. and Stone, J. (1999): Are Private Schools Really Better? Research in Labor Economics, pp. 115-140.

Figlio, D.N. and Ludwig, J. (2000): Sex, Drugs, and Catholic Schools: Private Schooling and Non-Market Adolescent Behaviors, National Bureau of Economic Research, Working Paper 7990.

Jenkins, Stephen P., and Schluter, C. (2002) The effect of family income during childhood on later-life attainment: evidence from Germany, Institute for Social and Economic Research, University of Essex, Working Paper 20.

Light, A. and Strayer, W. (2000): Determinants of College Completion: School Quality or Student Ability?, The Journal of Human Resources, XXXV (2), pp. 299-332.

Margolis, D.N. and Simonnet, V. (2003): Educational Track, Networks and Labor Market Outcomes, Institute for the Study of Labor, Bonn, Discussion Paper 699.

Neal, D. (1997): The Effects of Catholic Secondary Schooling on Educational Achievement, Journal of Labor Economics, 15(1), pp. 98-123.

Neal, D. (2002): How Vouchers Could Change the Market for Education, Journal of Economic Perspectives, 16(4), pp. 25–44.

Stinebrickner, R. and Stinebrickner, T. R. (2003): Working during School and Academic Performance, Journal of Labor Economics, 21(2), pp. 473-491.

Table 1: Description of school types, academic performance and labor market outcomes in the ISTAT sample.

| | (A) Distribution of high school types | |
|---|---|---|
| General | 31.13 | |
| Private | 12.32 | |
| General if has attended private | 48.49 | |
| Private if has attended general | 19.19 | |

| | (B) High school types and academic performance | |
|---|---|---|
| | College attendance | Exams speed |
| Whole sample | 40.81 | 45.64 |
| Technical school | 22.34 | 38.62 |
| General school | 81.67 | 49.88 |
| Public school | 40.02 | 46.13 |
| Private school | 46.42 | 42.68 |

| | (C) High school types and labor market outcomes | | |
|---|---|---|---|
| | Participation | Employment | Incidence of low pay |
| Whole sample | 67.71 | 42.77 | 24.76 |
| Technical school | 80.95 | 49.18 | 24.23 |
| General school | 40.47 | 16.4 | 31.31 |
| Public school | 68.41 | 43.24 | 25.17 |
| Private school | 62.65 | 39.01 | 21.13 |

Note: Number of observations= 15,295 in Panels (A) and (B) and 14,420 in Panel (C). See the text for the sample selection rules adopted.

Table 2: Results from estimation of academic performance models

(A) Effects of high school types and labor market participation on academic performance

| | (1) Full model | | | (2) As 1, but without controlling for unobserved heterogeneity | | | (3) As 1, but without controlling for participation | |
|---|---|---|---|---|---|---|---|---|
| | Labor market participation | College enrollment | Exams Speed | Labor market participation | College enrollment | Exams Speed | College enrollment | Exams Speed |
| Predicted probability | 0.751 | 0.391 | 0.437 | 0.750 | 0.396 | 0.450 | 0.394 | 0.521 |
| 'Marginal effects' | | | | | | | | |
| General high school | -0.230 | 0.373 | -0.063 | -0.281 | 0.468 | 0.070 | 0.385 | 0.193 |
| | (0.069) | (0.069) | (0.089) | (0.059) | (0.011) | (0.018) | (0.071) | (0.083) |
| Private high school | -0.023 | -0.012 | -0.211 | 0.017 | -0.039 | -0.060 | -0.054 | -0.082 |
| | (0.064) | (0.064) | (0.091) | (0.012) | (0.015) | (0.021) | (0.060) | (0.086) |
| Labor market participation | | | -0.314 | | | -0.113 | | |
| | | | (0.081) | | | (0.021) | | |

(B) Estimated correlations of unobservables, tests of ignorability and tests of instruments validity for model (1) of panel (A)

| | | | | | |
|---|---|---|---|---|---|
| ρ(general,private) | 0.328 | (0.019) | Ignorability of general high school | 316.21 [4] | *0.0000* |
| ρ(general,participation) | -0.080 | (0.093) | Ignorability of private high school | 319.63[4] | *0.0000* |
| ρ(general,college) | 0.143 | (0.104) | Significance of correlation structure | 4179.19 [10] | *0.0000* |
| ρ(general,speed) | 0.156 | (0.109) | | | |
| ρ(private, participation) | 0.053 | (0.092) | Significance of instruments in high school equations | 28.06[4] | *0.0000* |
| ρ(private, college) | -0.013 | (0.092) | Significance of instruments in college equation | 91.34 [26] | *0.0000* |
| ρ(private, speed) | 0.267 | (0.109) | Significance of instruments in participation equation | 96.53 [26] | *0.0000* |
| ρ(participation, college) | -0.680 | (0.012) | Exclusion of instruments from college equation | 3.02 [3] | *0.3879* |
| ρ(participation, speed) | 0.315 | (0.134) | Exclusion of instruments from participation equation | 3.10 [3] | *0.3770* |
| ρ(college, speed) | -0.032 | (0.090) | Exclusion of instruments from speed equation | 25.05 [26] | *0.5162* |

Note: Number of observations=15,295. Results of full model derived from estimation of five-variate probit model with endogenous truncation of one equation. Multivariate normal distributions are computed using a GHK simulator with 130 random draws. Standard errors in parentheses, degrees of freedom in square brackets, p-values in italic. Test statistics are derived from Wald tests. Predicted probabilities and 'marginal effects' are evaluated at sample means of explanatory variables.

Table 3: Results from estimation of labor market models

(A) Effects of high school types and college enrolment on labor market outcomes

| | (1) Full model | | (2) As 1, but without controlling for unobserved heterogeneity | | (3) As 1, but without controlling for college enrolment | |
|---|---|---|---|---|---|---|
| | Employment | Low pay | Employment | Low pay | Employment | Low pay |
| Predicted probability | 0.386 | 0.148 | 0.392 | 0.198 | 0.406 | 0.221 |
| 'Marginal effects' | | | | | | |
| General high school | -0.281 | -0.072 | -0.165 | 0.053 | -0.330 | -0.019 |
| | (0.047) | (0.135) | (0.016) | (0.030) | (0.042) | (0.077) |
| Private high school | -0.088 | -0.032 | -0.026 | -0.024 | 0.011 | -0.053 |
| | (0.063) | (0.079) | (0.018) | (0.021) | (0.075) | (0.093) |
| College enrolment | -0.348 | -0.250 | -0.375 | 0.045 | | |
| | (0.038) | (0.297) | (0.015) | (0.031) | | |

(B) Estimated correlations of unobservables, tests of ignorability and tests of instruments validity for model (1) of panel (A)

| | | | | | |
|---|---|---|---|---|---|
| λ(general,private) | 0.338 | (0.019) | Ignorability of general high school | 335.24 [5] | *0.0000* |
| λ(general,college) | 0.081 | (0.100) | Ignorability of private high school | 324.53 [5] | *0.0000* |
| λ(general,participation) | -0.149 | (0.112) | Significance of correlation structure | 4014.01 [15] | *0.0000* |
| λ(general,employment) | 0.169 | (0.076) | | | |
| λ(general,low pay) | 0.187 | (0.117) | | | |
| λ(private,college) | 0.053 | (0.093) | | | |
| λ(private,participation) | -0.069 | (0.102) | | | |
| λ(private,employment) | 0.125 | (0.089) | Significance of instruments in high school equations | 26.34 [4] | *0.0000* |
| λ(private,low pay) | 0.066 | (0.147) | Significance of instruments in college equation | 841.25 [49] | *0.0000* |
| λ(college,participation) | -0.671 | (0.012) | Significance of instruments in participation equation | 502.85 [49] | *0.0000* |
| λ( college,employment) | -0.069 | (0.070) | Significance of instruments in employment equation | 69.82 [29] | *0.0000* |
| λ( college,low pay) | -0.282 | (0.186) | Exclusion of instruments from college equation | 2.26 [3] | *0.5197* |
| λ(participation,employment) | -0.004 | (0.086) | Exclusion of instruments from participation equation | 3.63 [3] | *0.3045* |
| λ(participation,low pay) | 0.309 | (0.160) | Exclusion of instruments from employment equation | 15.43 [20] | *0.7510* |
| λ( employment,low pay) | -0.206 | (0.210) | Exclusion of instruments from low pay equation | 32.70 [29] | *0.2900* |

Note: Number of observations=14,420. Results of full model derived from estimation of six-variate probit model with endogenous truncation of two equations. Multivariate normal distributions are computed using a GHK simulator with 130 random draws. Standard errors in parentheses, degrees of freedom in square brackets, p-values in italic. Test statistics are derived from Wald tests. Predicted probabilities and 'marginal effects' are evaluated at sample means of explanatory variables.

Table A1: Results from estimation of academic performance model ('marginal effects', standard errors in parentheses)

| | General high school | | Private high school | | Labor market participation | | University enrolment | | Exams speed | |
|---|---|---|---|---|---|---|---|---|---|---|
| Predicted probability | 0.255 | | 0.106 | | 0.752 | | 0.392 | | 0.438 | |
| General high school | | | | | -0.230 | (0.070) | 0.373 | (0.070) | -0.064 | (0.089) |
| Private high school | | | | | -0.024 | (0.064) | -0.013 | (0.064) | -0.211 | (0.092) |
| Labor market participation | | | | | | | | | -0.315 | (0.081) |
| Female | 0.096 | (0.010) | 0.031 | (0.006) | 0.048 | (0.012) | -0.026 | (0.012) | 0.066 | (0.018) |
| Resides in the north east | -0.012 | (0.012) | -0.019 | (0.007) | 0.006 | (0.015) | 0.000 | (0.015) | -0.011 | (0.020) |
| Resides in the centre | 0.005 | (0.012) | -0.056 | (0.007) | 0.001 | (0.015) | 0.017 | (0.015) | -0.151 | (0.042) |
| Resides in the south | 0.010 | (0.012) | -0.039 | (0.007) | -0.012 | (0.016) | 0.067 | (0.016) | -0.251 | (0.065) |
| Born before 1970 | -0.099 | (0.038) | 0.141 | (0.038) | 0.210 | (0.116) | -0.149 | (0.116) | -0.130 | (0.285) |
| Born between 1970 and 1973 | -0.144 | (0.021) | 0.114 | (0.023) | 0.139 | (0.055) | -0.125 | (0.055) | -0.052 | (0.115) |
| Born in 1974 | -0.074 | (0.016) | 0.022 | (0.012) | 0.101 | (0.034) | -0.071 | (0.034) | -0.062 | (0.071) |
| Born in 1975 | -0.029 | (0.011) | 0.007 | (0.007) | 0.043 | (0.020) | -0.011 | (0.020) | -0.033 | (0.040) |
| Born in 1977 | 0.177 | (0.021) | -0.023 | (0.010) | -0.034 | (0.024) | -0.019 | (0.024) | -0.024 | (0.015) |
| Junior high school mark=Buono (C) | 0.146 | (0.016) | -0.020 | (0.006) | -0.032 | (0.017) | 0.093 | (0.017) | 0.033 | (0.026) |
| Junior high school mark=Distinto (B) | 0.330 | (0.019) | -0.032 | (0.007) | -0.062 | (0.024) | 0.179 | (0.024) | 0.077 | (0.034) |
| Junior high school mark=Ottimo (A) | 0.530 | (0.016) | -0.051 | (0.008) | -0.105 | (0.034) | 0.262 | (0.034) | 0.116 | (0.044) |
| Junior high school mark missing | 0.162 | (0.024) | 0.022 | (0.012) | -0.027 | (0.024) | 0.107 | (0.024) | 0.096 | (0.036) |
| High school enrollment at 13 | -0.014 | (0.016) | 0.035 | (0.012) | 0.013 | (0.021) | 0.048 | (0.021) | 0.055 | (0.025) |
| High school enrollment at 15 | -0.022 | (0.018) | 0.015 | (0.011) | 0.033 | (0.023) | -0.080 | (0.023) | 0.004 | (0.037) |
| High school enrollment at 16 or more | 0.020 | (0.039) | 0.026 | (0.020) | 0.003 | (0.039) | -0.034 | (0.039) | -0.151 | (0.097) |
| High school mark from 41 to 45 | | | | | -0.045 | (0.015) | 0.080 | (0.015) | 0.035 | (0.023) |
| High school mark from 46 to 50 | | | | | -0.112 | (0.018) | 0.183 | (0.018) | 0.066 | (0.028) |
| High school mark from 51 to 55 | | | | | -0.164 | (0.017) | 0.296 | (0.017) | 0.163 | (0.034) |
| High school mark from 56 to 59 | | | | | -0.248 | (0.025) | 0.336 | (0.025) | 0.168 | (0.040) |
| High school mark= 60 | | | | | -0.299 | (0.037) | 0.420 | (0.037) | 0.237 | (0.041) |
| Has failed during high school | | | | | 0.020 | (0.019) | -0.068 | (0.019) | -0.065 | (0.034) |
| Has changed high school | | | | | 0.026 | (0.021) | -0.056 | (0.021) | 0.008 | (0.038) |
| Age at end of high school | | | | | -0.038 | (0.016) | 0.009 | (0.016) | 0.048 | (0.044) |
| Has 1 sibling | -0.069 | (0.011) | -0.044 | (0.007) | 0.053 | (0.016) | -0.063 | (0.016) | | |
| Has 2 siblings | -0.082 | (0.012) | -0.052 | (0.008) | 0.073 | (0.020) | -0.086 | (0.020) | | |
| Has 3 siblings | -0.085 | (0.016) | -0.060 | (0.009) | 0.060 | (0.025) | -0.088 | (0.025) | | |
| Has 4 or more siblings | -0.149 | (0.019) | -0.054 | (0.011) | 0.105 | (0.034) | -0.115 | (0.034) | | |
| Father's degree = Junior high school | 0.035 | (0.013) | 0.016 | (0.008) | 0.000 | (0.014) | 0.003 | (0.014) | -0.045 | (0.028) |
| Father's degree = High school | 0.117 | (0.018) | 0.031 | (0.011) | -0.039 | (0.021) | 0.096 | (0.021) | 0.007 | (0.023) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Father's degree = University | 0.295 | (0.029) | 0.061 | (0.019) | -0.138 | (0.033) | 0.233 | (0.033) | 0.025 | (0.028) |
| Father's degree not reported | 0.043 | (0.036) | 0.031 | (0.023) | -0.010 | (0.038) | -0.045 | (0.038) | 0.022 | (0.037) |
| Mother's degree =Junior high school | 0.039 | (0.013) | 0.014 | (0.008) | -0.002 | (0.014) | 0.032 | (0.014) | -0.033 | (0.061) |
| Mother's degree =High school | 0.140 | (0.018) | 0.057 | (0.013) | -0.045 | (0.022) | 0.094 | (0.022) | 0.038 | (0.022) |
| Mother's degree =University | 0.279 | (0.033) | 0.050 | (0.021) | -0.112 | (0.039) | 0.156 | (0.039) | 0.055 | (0.028) |
| Mother's degree not reported | 0.162 | (0.046) | 0.050 | (0.029) | -0.068 | (0.047) | -0.029 | (0.047) | 0.058 | (0.038) |
| Father's occupation | | | | | | | | | | |
| Self employed shop seller/retailer | 0.125 | (0.025) | 0.128 | (0.024) | -0.112 | (0.027) | 0.102 | (0.027) | | |
| Craft | 0.044 | (0.021) | 0.034 | (0.015) | -0.079 | (0.023) | 0.050 | (0.023) | | |
| Farmer | -0.014 | (0.026) | 0.016 | (0.019) | -0.059 | (0.030) | 0.044 | (0.030) | | |
| Entrepreneur | 0.086 | (0.028) | 0.134 | (0.027) | -0.129 | (0.031) | 0.118 | (0.031) | | |
| Professional | 0.126 | (0.028) | 0.079 | (0.022) | -0.150 | (0.030) | 0.148 | (0.030) | | |
| Manager | 0.146 | (0.027) | 0.055 | (0.019) | -0.087 | (0.030) | 0.122 | (0.030) | | |
| Teacher | 0.067 | (0.034) | -0.045 | (0.015) | -0.089 | (0.039) | 0.083 | (0.039) | | |
| White collar high level | 0.074 | (0.020) | 0.006 | (0.012) | -0.087 | (0.022) | 0.106 | (0.022) | | |
| White collar low level | 0.055 | (0.021) | 0.005 | (0.013) | -0.084 | (0.024) | 0.087 | (0.024) | | |
| Blue collar high level | -0.034 | (0.015) | -0.006 | (0.010) | -0.027 | (0.018) | 0.004 | (0.018) | | |
| Not reported | 0.026 | (0.034) | 0.019 | (0.022) | -0.116 | (0.038) | 0.105 | (0.038) | | |
| Mother's occupation | | | | | | | | | | |
| Self employed shop seller/retailer | 0,011 | (0.026) | 0.047 | (0.021) | -0.006 | (0.029) | -0.017 | (0.029) | | |
| Craft | 0.021 | (0.039) | 0.053 | (0.030) | 0.019 | (0.042) | 0.016 | (0.042) | | |
| Farmer | -0.047 | (0.041) | -0.023 | (0.025) | 0.058 | (0.048) | -0.026 | (0.048) | | |
| Entrepreneur | 0.004 | (0.062) | 0.026 | (0.042) | -0.122 | (0.072) | 0.007 | (0.072) | | |
| Professional | 0.050 | (0.053) | 0.017 | (0.032) | 0.008 | (0.059) | -0.023 | (0.059) | | |
| Manager | 0.187 | (0.061) | 0.006 | (0.027) | -0.055 | (0.068) | 0.139 | (0.068) | | |
| Teacher | 0.044 | (0.026) | -0.006 | (0.014) | -0.003 | (0.029) | 0.025 | (0.029) | | |
| White collar high level | 0.074 | (0.027) | 0.028 | (0.018) | -0.016 | (0.028) | 0.054 | (0.028) | | |
| White collar low level | 0.038 | (0.024) | 0.016 | (0.015) | -0.002 | (0.025) | -0.003 | (0.025) | | |
| Blue collar high level | -0.061 | (0.024) | 0.012 | (0.018) | 0.069 | (0.031) | -0.088 | (0.031) | | |
| Not reported | -0.007 | (0.017) | 0.011 | (0.022) | 0.006 | (0.019) | -0.046 | (0.019) | | |
| One grandparent has high school degree | 0.068 | (0.015) | 0.016 | (0.009) | | | | | 0.038 | (0.019) |
| Mother retired when individual was 14 | | | -0.034 | (0.016) | | | | | 0.00001 | (0.043) |
| Mother housewife when individual was 14 | | | -0.011 | (0.020) | | | | | -0.056 | (0.070) |

Note: Number of observations=15,295, Log-likelihood= -28192.359. Multivariate normal distributions are computed using a GHK simulator with 130 random draws. Predicted probabilities and 'marginal effects' are evaluated at sample means of explanatory variables. The excluded category is: technical or vocational high school, public high school, male, resides in the north west, has no siblings, has both parents with no or elementary degree in low level manual occupations, reported a mark of D at junior high school, enrolled at high school at 14, is born in 1976, reported a mark between 36 and 40 at high school, has never failed at high school, has never changed high school, has no grandparents with at least high school degree and his mother was not retired or housewife when the individual was 14.

Table A2: Results from estimation of labor market outcomes model ('marginal effects', standard errors in parentheses)

| | Labor market participation | | Employment | | Low pay | |
|---|---|---|---|---|---|---|
| Predicted probability | 0.709 | | 0.386 | | 0.148 | |
| General high school | -0.193 | (0.081) | -0.281 | (0.047) | -0.072 | (0.135) |
| Private high school | 0.044 | (0.061) | -0.088 | (0.063) | -0.032 | (0.079) |
| College enrolment | | | -0.348 | (0.038) | 0.250 | (0.297) |
| Female | 0.052 | (0.016) | -0.081 | (0.012) | 0.117 | (0.173) |
| Resides in the north east | 0.003 | (0.013) | 0.064 | (0.018) | -0.050 | (0.081) |
| Resides in the centre | 0.012 | (0.013) | -0.167 | (0.014) | 0.091 | (0.129) |
| Resides in the south | 0.010 | (0.012) | -0.398 | (0.014) | 0.317 | (0.307) |
| Born before 1970 | 0.218 | (0.122) | -0.007 | (0.051) | -0.039 | (0.231) |
| Born between 1970 and 1973 | 0.138 | (0.065) | -0.047 | (0.029) | -0.056 | (0.148) |
| Born in 1974 | 0.109 | (0.046) | -0.016 | (0.021) | -0.002 | (0.047) |
| Born in 1975 | 0.041 | (0.023) | -0.002 | (0.015) | 0.003 | (0.024) |
| Born in 1977 | -0.035 | (0.025) | 0.018 | (0.032) | 0.010 | (0.050) |
| Junior high school mark=Buono (C) | -0.028 | (0.015) | | | | |
| Junior high school mark=Distinto (B) | -0.068 | (0.026) | | | | |
| Junior high school mark=Ottimo (A) | -0.118 | (0.040) | | | | |
| Junior high school mark missing | -0.037 | (0.023) | | | | |
| High school enrollment at 13 | 0.007 | (0.017) | | | | |
| High school enrollment at 15 | 0.042 | (0.020) | | | | |
| High school enrollment at 16 or more | 0.018 | (0.035) | | | | |
| High school mark from 41 to 45 | -0.044 | (0.015) | | | -0.011 | (0.022) |
| High school mark from 46 to 50 | -0.112 | (0.025) | | | -0.026 | (0.043) |
| High school mark from 51 to 55 | -0.162 | (0.030) | | | -0.024 | (0.044) |
| High school mark from 56 to 59 | -0.238 | (0.034) | | | -0.084 | (0.143) |
| High school mark= 60 | -0.295 | (0.033) | | | -0.101 | (0.179) |
| Has failed during high school | 0.019 | (0.016) | | | 0.015 | (0.031) |
| Has changed high school | 0.015 | (0.019) | | | 0.015 | (0.033) |
| Age at end of high school | -0.036 | (0.022) | | | -0.017 | (0.012) |
| Has 1 sibling | 0.060 | (0.018) | | | 0.007 | (0.022) |
| Has 2 siblings | 0.082 | (0.024) | | | 0.044 | (0.065) |
| Has 3 siblings | 0.071 | (0.025) | | | 0.044 | (0.068) |
| Has 4 or more siblings | 0.107 | (0.038) | | | 0.066 | (0.094) |
| Father's degree = Junior high school | 0.002 | (0.013) | | | 0.017 | (0.029) |
| Father's degree = High school | -0.043 | (0.020) | | | 0.035 | (0.054) |
| Father's degree = University | -0.153 | (0.039) | | | 0.041 | (0.086) |
| Father's degree not reported | -0.022 | (0.035) | | | 0.081 | (0.113) |
| Mother's degree =Junior high school | -0.004 | (0.012) | | | -0.0005 | (0.016) |
| Mother's degree =High school | -0.063 | (0.022) | | | -0.021 | (0.039) |
| Mother's degree =University | -0.141 | (0.039) | | | 0.065 | (0.119) |
| Mother's degree not reported | -0.100 | (0.048) | | | 0.089 | (0.131) |
| Father's occupation | | | | | | |
| Self employed shop seller/retailer | -0.138 | (0.033) | 0.008 | (0.029) | | |
| Craft | -0.093 | (0.028) | 0.007 | (0.024) | | |
| Farmer | -0.060 | (0.031) | -0.023 | (0.031) | | |
| Entrepreneur | -0.156 | (0.037) | 0.080 | (0.036) | | |
| Professional | -0.177 | (0.037) | 0.022 | (0.037) | | |
| Manager | -0.091 | (0.030) | 0.036 | (0.035) | | |
| Teacher | -0.080 | (0.035) | -0.007 | (0.055) | | |
| White collar high level | -0.087 | (0.026) | 0.005 | (0.022) | | |
| White collar low level | -0.092 | (0.027) | 0.013 | (0.025) | | |
| Blue collar high level | -0.019 | (0.017) | 0.049 | (0.019) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Not reported | -0.137 | (0.042) | -0.053 | (0.042) | | |
| Mother's occupation | | | | | | |
| Self employed shop seller/retailer | -0.037 | (0.029) | 0.024 | (0.036) | | |
| Craft | -0.006 | (0.039) | 0.095 | (0.052) | | |
| Farmer | 0.069 | (0.043) | -0.066 | (0.049) | | |
| Entrepreneur | -0.170 | (0.076) | -0.099 | (0.097) | | |
| Professional | -0.004 | (0.050) | -0.089 | (0.073) | | |
| Manager | -0.074 | (0.052) | -0.124 | (0.084) | | |
| Teacher | 0.000 | (0.024) | -0.085 | (0.031) | | |
| White collar high level | -0.025 | (0.025) | -0.034 | (0.031) | | |
| White collar low level | -0.009 | (0.023) | -0.022 | (0.029) | | |
| Blue collar high level | 0.081 | (0.031) | 0.024 | (0.032) | | |
| Not reported | -0.001 | (0.017) | -0.017 | (0.020) | | |
| Search mode current job | | | | | | |
| Knew the employer personally | | | | | -0.009 | (0.024) |
| Contacted by the employer | | | | | -0.071 | (0.123) |
| Job advertisement | | | | | -0.034 | (0.058) |
| Sending CVs to employers | | | | | -0.036 | (0.059) |
| Public competition | | | | | -0.140 | (0.283) |
| Family firm | | | | | 0.032 | (0.055) |
| Employment agencies | | | | | -0.051 | (0.087) |
| Part-time | | | | | -0.114 | (0.207) |
| Work and training contract | | | | | 0.045 | (0.072) |
| Fixed term contract | | | | | -0.050 | (0.078) |
| Occupation | | | | | | |
| White collar high level | | | | | -0.092 | (0.173) |
| White collar low level | | | | | -0.006 | (0.017) |
| Blue collar high level | | | | | -0.006 | (0.018) |
| Employee of family firm | | | | | 0.066 | (0.101) |
| Apprenticeship | | | | | 0.204 | (0.214) |
| Public sector | | | | | 0.003 | (0.020) |
| Industry | | | | | | |
| Agriculture | | | | | 0.011 | (0.040) |
| Retail trade | | | | | 0.022 | (0.037) |
| Transport & communication | | | | | 0.002 | (0.031) |
| Financial services | | | | | -0.082 | (0.149) |
| Public administration. Education. Health | | | | | -0.012 | (0.033) |
| Housing or IT services | | | | | 0.057 | (0.085) |
| Other services | | | | | 0.086 | (0.117) |
| 6≤Firm size<15 | | | | | -0.074 | (0.123) |
| 15≤Firm size<50 | | | | | -0.109 | (0.189) |
| 50≤Firm size<100 | | | | | -0.131 | (0.253) |
| Firm size≥100 | | | | | -0.173 | (0.301) |
| Job started in 1995 | | | | | -0.063 | (0.105) |
| Job started in 1996 | | | | | -0.066 | (0.108) |
| Job started in 1997 | | | | | -0.039 | (0.061) |

Note: Number of observations=14.420. Log-likelihood= -30851.587. Estimates for high school types and university enrolment equations omitted. Multivariate normal distributions are computed using a GHK simulator with 130 random draws. Predicted probabilities and 'marginal effects' are evaluated at sample means of explanatory variables. The excluded category is: technical or vocational high school, public high school, male, resides in the north west, has no siblings, has both parents with no or elementary degree in low level manual occupations, reported a mark of D at junior high school, enrolled in high school at 14, born in 1976, reported a mark between 36 and 40 at high school, has never failed during high school, has never changed high school, has no grandparents with at least high school degree and his mother was not retired or housewife when the individual was 14, found current job through informal networks, works full time on a long term contract as low level blue collar, in a small firm of the private manufacturing sector, started current job in 1998.