School Choice and Performance: Private Education in France

Giuseppe Bertola (and Pierre Courtioux)

(EDHEC Business School)

April 2015, **Preliminary**

This paper proposes an empirical model of the determinants and effects of school choices and applies it to private schooling enrolment in a large and detailed French dataset. Estimation methods that allow individual ability to be relevant to private school choices indicate that in France State-provided education is more suitable for students who are talented and helped by a culturally privileged family background, while private schooling is chosen by students with lower learning ability from richer families. In a country that values equality of opportunity and provides abundant public funding to privately organized schools, these results have interesting policy implications.

Keywords: Skills, family background.

JEL codes: I21, I24.

1. Introduction

Expensive private education generally provides attractive amenities and networking opportunities, but it is not always associated with better results. Results, in fact, depend not only on available educational resources but also on the characteristics of their student population. If the costly resources financed by private school fees are complementary to students' talents, then they attract better students. This appears to be the case in Anglo-Saxon and other countries where the public school system provides only basic education, and private schools are attended by more talented students (De Fraja, 2002). In Italy and other countries with relatively demanding and selective State schools, conversely, private schools are chosen by somewhat less talented pupils, who benefit from expensive but essentially remedial educational services (Bertola and Checchi, 2013; Brunello and Rocco, 2008; Bertola, Checchi and Oppedisano, 2007).

For France, the PISA 2009 survey studied by e.g. OECD (2012) and Bertola and Checchi (2013) does not identify privately funded and/or operated schools. The French State pays the salaries of teachers at private schools of the *sous contrat* type that, while privately run, commit to employ only State-certified teachers and to abide by the same academic standards as State school. Almost all French private schools accept State funding and these constraints. The substantial cost of the former makes it all the more interesting to characterize how they use student fee revenues and the substantial degrees of freedom allowed by the latter.

This paper exploits National data sources to assess the educational effectiveness of French private schools and the process that selects French students into private education. The data do not provide information on the pedagogical aspects that Bertola and Checchi (2012) interact with individual background information to characterize cross-country differences in private school sorting. In some key respects, however, French data are more informative than those available in the international PISA dataset and in other country-specific data sets. A large and detailed panel survey provides information about students' individual background and about educational achievements at different stages of their career, and individuals records can be geographically matched to the summary statistics of administrative data in order to exploit within-country variation across "local educational market" (LEMs).

Empirical analysis of these data faces identification issues similar to those confronted by studies of Catholic private schools in the United States (e.g. Coleman, Hoffer, and Kilgore, 1982; Evans and Schwab, 1995; Neal, 1997) and by Bertola, Checchi and Oppedisano's (2007) study of Italian private schools. In the US, Catholic schools deliver further education and labor

market outcomes that are better on average, and less tightly related to the students' background than those observed in other schools. In Italy, private school students are on average less successful than State school students. Some of these differences result from selection across schools of heterogeneous students, whose performance depends on their characteristics directly as well as through their school enrollment choice. To identify the determinants and effects of private school enrolment, instrumental variables should determine school choices but not be directly relevant to results. Enrolment in American Catholic schools is more likely for families who are Catholic or reside near such schools, and richer Italian families are more likely to enroll their children in private schools. If religion, residence, and wealth do not directly influence a student's further progress, then it is possible disentangle the effects of schooling from those of other background characteristics. While debatable (see the references in Altonji et al, 2005, footnote 2) such identification assumptions typically yield results suggesting that schooling outcomes are less influenced by family background at American Catholic schools, and less influenced by individual students' ability to learn at Italian private schools.

Unlike the models estimated and reviewed by Altonji et al (2005), where schools types may only have a level effect on individual success probabilities while control variable coefficients are restricted to be the same across schools, our empirical specifications allow observable and unobservable variation to influence outcomes differently across State and private schools. We obtain interesting and fairly robust results exploiting the abundance in French data of sources of variation that influence enrolment but do not directly influence school outcomes. To disentangle the determinants of school choice from those of school performance, identifying assumptions can exploit information in the available data about each family's financial situation, past school choices, and school-choice criteria, as well as regional variation of educational market condition (following Martínez-Mora, 2006, who finds that low-quality private schools are more prevalent in US localities where property tax revenues support only low-quality State schools).

2. Institutional structure

In France, secondary school includes a common lower level (four years of *collège*) and a differentiated upper level (*lycée*) that in three years can lead to achievement of a *baccalauréat* exit degree in academic (scientific, economic-social, or literary) or technical tracks, with a common first year and two years of more specialized studies. It can also be obtained in four years by enrolling in an upper-secondary vocational track that, like technical three-year tracks,

is employment-oriented. After obtaining the *baccalauréat* students can enter tertiary education, which in France is mainly supplied by universities but also by *Grandes Ecoles* and by *lycées*: some general ones select students into *Classes Préparatoires aux Grandes Ecoles* (CPGE), some professional or technical ones offer selective vocational programs (BTS). Any *baccalauréat* also makes it possible to apply for admission to selective vocational programs run by universities (IUT), which are attractive for the best students from technical and vocational tracks as well as for academic track students who fail to obtain admission to longer and more prestigious *Grandes Ecoles.*

About 20% of students in secondary education attend private schools. This percentage has been stable since the mid 1990's (Maettz, 2004) and is somewhat higher than the 17% observed in primary (Vasconcellos and Bongrand, 2013). An overwhelming majority of the private schools are of the *sous contrat* type introduced by the 1959 Debré Act in order to include the private sector in the country's massive investment in primary education primary level. Since the 1977 Germeur Act, at *sous contrat* schools teacher salaries are entirely funded by the State and, at given seniority, are similar to those of State school teachers. In 2011, 93.5% of primary private schools and 85% of private secondary schools were *école privée sous contrat*. Only 2.8% of the students attending a private school were in totally autonomous private schools (*école privée hors contrat*) such as Montessori establishments (Vasconcellos and Bongrand 2013, p.13). As a condition for State funding of teachers' salaries, private schools have to teach the same curriculum as State ones, and employ only teachers who have passed a national competition (*concours*).

The substantial cost of State funding of teachers' wages at private schools makes it all the more interesting to find out how private schools use their limited autonomy. The academic curriculum should in principle be the same in State and in *sous contrat* private schools, but private schools may of course teach the material differently, with different teachers, and to a different student body.

State and private teachers are managed differently, and self-selected because, from their point of view, the State career is more attractive. State school teachers are civil servants (*fonctionnaire*), and they are assigned to jobs by a strictly administrative procedure: a vacant place must be assigned to the applicant who ranks highest in terms of a score based on *concours* results, seniority, and some career features (such as serving in administrative or managerial roles). While private school teachers must have passed an exam that is similar to that of State school teachers, if they have also passed the latter they are not likely to seek employment in the private sector, where career paths and working conditions are much less

tightly regulated. Private school teachers' contracts are not written under private labor law because the 1992 Coupet-Lang agreement recognized that private education provide public services (*mission de service public*; it is for this reason, as well as to avoid confusion with the private schools that the British call "public", that we use "State" to refer to government-run French schools). Like that of local government workers, their employment is subject to a form of public law (*contractuel de droit public*) that does not administratively restrict job assignments: for each of the private school teaching positions funded and assigned to schools by the State, school managers may freely choose any *concours*-qualified applicant.

Private schools are not only staffed by different teachers, but also attended by different students: private schools cater to richer families, because they are allowed to charge fees to cover the cost of facilities, amenities, and support staff (see Appendix 1), and because they can set their own admission criteria (which, as a condition of *sous contrat* funding, must not include religious allegiance). In the State sector, instead, students generally must attend a specific school within their area of residence (see Appendix 2 for details of this residence-based *sectorisation* constraint, which had been relaxed in previous times but was uniformly very tight over the period covered by our data: only a few elite *lycée*s in Paris could admit some excellent students from a broader area). They may instead choose to apply, and pay, for enrolment at private schools.

2. Available data

The *Panel d'élèves du second degré* 1995-2006 selects by birthday and follows over time a random 1/40th sample of the students who entered French lower secondary school (*collège*) in 1996.[1] The data includes information collected in the initial year about the student's math and reading skills at entry in secondary school: an assessment by the principal on a 0-10 scale, relative to what is typical in France in general rather than to that school or class, and test scores. The progress of each student is followed during secondary school, and for up of 9 years after upper secondary school completion. Both administrative records and survey answers are available. About 94% of the initial sample is retained up to 2006, the year when they were expected to complete their secondary studies. Attrition is much stronger afterwards, and this leads us to focus on school achievement at the end of secondary school.

[1] The dataset is documented and studied by e.g. Anamian et al. (2005), Nakhili (2005), Cayouette-Remblière and de Saint Pol (2013). Reforms, in particular of the *sectorisation* rule discussed in Appendix 2, would make it very difficult to merge these data with the similar panels that started in 1980 and 1989.

Information about each student's socio-economic situation and primary school experience was collected in 1998 from the family through mail questionnaires and follow-up phone interviews. This background information is available for 15290 of the 17830 individuals in the first year of the panel: 12981 families replied to mail questionnaires, and 2309 more answered similar questions by phone. Beginning in 2002, when students could have successfully completed secondary school, annual surveys were administered to the students; answers are available for 16701 individuals. Undocumented sampling weights are provided for responders to either family survey phases (pond1), responders to the postal survey phase (pond2), and for responders to the student survey (jpond). We use the appropriate variables as Stata pweights in our estimation, and we check that the results we report are substantially the same in all available selected samples.

The data record the school attended by each student in every school year. In 1995-96, the 17830 panel individuals are enrolled in 5686 distinct lower-secondary schools; in 2001-02, 1032 of them have dropped out of the sample (and probably also out of secondary school), and the rest are observed attending 4594 distinct upper-secondary schools. Each school is identified anonymously, but we know its location up to cells defined by size of town and *département* local government units (there are 96 *département* in Continental France and Corsica). The student's residence need not be in the same locality as the school's. Neither it nor school grades are recorded in the dataset, which however does report various school performance indicators. Family background information is collected in 1998, when panel students should be finishing the third or starting the fourth and final year of lower secondary school. It includes a summary indicator of the family's socio-economic indicator; the educational achievements and employment situation of the parents, retrospective information about pre-secondary schooling; and a battery of subjective questions that we further discuss and document below.

Detailed information is available about the content of studies: for each year from 1995 to 2006, we know in which academic or vocational upper secondary track each student is enrolled at the school attended (each school offers many, possibly all tracks). The data report whether each surveyed individual attends a private school. Consistently with the aggregate statistics reported above, for the overwhelming majority of private school students the school type is "*Contrat d'association toutes classes*"; only a few dozen attend "*Hors contrat*" private schools. No information is available on the amount of school fees and availability of the facilities and pedagogical aids that are typically offered by private schools (see Appendix 1).

We complement the information available in the individual survey data set with that available in the *Base Centrale de Scolarité* (BCS), the administrative database of the French education

system. The data made available to us cover all the 11,123 French secondary schools in 2004, 2005, and 2006. They record whether each school is attended by lower secondary (*collège*) or upper secondary (*lycée*) students; for the latter, the academic or professional curricula offered by the school are recorded. The BCS also reports whether the school is a private or State establishment and, in the latter case, whether it belongs to a ZEP ("*zone d'éducation prioritaire*"), or a REP ("*réseau d'éducation prioritaire*"), or is classified as "*établissement sensible*": all indicate that the school is attended by relatively troublesome and socially underprivileged students and is granted some additional resources that are supposed to improve the quality of State education but, by making it evident that the school is attended by underprivileged student, might increase the appeal of private education for families who reside in poor or mixed areas. For each of the about 5.5 million students attending these schools in each year, the BCS records age, gender, school(s) attended in the current and in the previous year, and an indicator of socio-economic status similar to that recorded in the panel.

The anonymous identifier of the (also about 11000) schools attended by the panel data set students cannot be linked to BCS data, but the geographical location of the school is coded similarly in both data sets. City sizes are recorded in 7 cells (from "rural" to "urban area between 200 000 and 2 000 000 inhabitants") and the Paris urban area is coded separately. Of course not all city sizes are present in each *département*; it is possible to identify 456 city size x *département* localities.

Aiming to characterize variation of education demand and supply within France at the level of "Local Educational Markets" (LEMs), we will link locally aggregated BCS administrative data to individual panel students. To better approximate the determinants of choices made in the panel's earlier period, in doing so we classify BCS schools on the basis of the information observed when they first appear in the administrative dataset. [2]

3. Data generating process

We specify the process generating these data as a Roy selection mechanism similar to that used in earlier related work (see Bertola, Checchi and Oppedisano, 2007, and their references). The

[2] This is almost always 2004, but 101 schools appear only in one or two of the three available years of data. No schools are observed switching in 2004-2006 between State and private, or into/out of the REP classification; only a few schools move into the ZEP (6 in 2005 and 1 in 2006) or *établissement sensible* categories (1 in 2005, 1 in 2006).

relatively high quality and fine detail of the available French data make it possible to use somewhat more sophisticated methods.

Let $y_i \in \{0,1\}$ indicate school failure or success by student i, and suppose i's probability of success is related to a vector X_i of individual characteristics (and a constant) by a vector of school-specific coefficients β_j: if i attends school j, then

$$P_j(y_i = 1) = \Pr\left(X_i \beta_j + \gamma_j \eta_i + \varepsilon_i > 0\right). \quad (1)$$

The individual characteristics X_i are observable, and predetermined along with η_i at the time when school choices aim at maximizing the expected value of the welfare $u_{ij}(y_i)$ realized when the student succeeds ($y_i = 1$) or fails ($y_i = 0$)

$$P_j(y_i = 1 \mid \eta_i)\, u_{ij}(1) + \left(1 - P_j(y_i = 1 \mid \eta_i)\right) u_{ij}(0). \quad (2)$$

From the family's point of view, the realization of η_i is already observed when choosing the school, and ε_i represents subsequent random events that prevent school outcomes from being deterministic. From the econometrician's point of view, neither of these variables is observable: individual i is observed attending a private school ($j = p$) when that is better in terms of (2) than a State school ($j = s$), i.e.

$$P_p(y_i = 1 \mid \eta_i)\left(u_{ip}(1) - u_{ip}(0)\right) + u_{ip}(0) > P_s(y_i = 1 \mid \eta_i)\left(u_{is}(1) - u_{is}(0)\right) + u_{is}(0) \quad (3)$$

(similar inequalities also determine whether individuals are observed in different curricula or geographic locations).

The notation in (3) allows the welfare implications of school performance to vary across families indexed by i and schools indexed by j. The different relevance of success for different students may reflect the more serious consequences of failure for families whose risk aversion is strengthened by financial circumstances, or to the asymmetric implications of parents' aversion to seeing their children achieve less than their own educational level (Breen and Goldthorpe, 1997). This implies that the welfare gain generated by academic success differs across families, and may be a function of observable variables that may include the same X_i that influence academic success as well as others, denoted Z_i, that do not:

$$u_{is}(1) - u_{is}(0) = u_{ip}(1) - u_{ip}(0) = f\left(X_i, Z_i\right). \quad (4)$$

Private schools may also be appealing or unappealing for financial, practical, or cultural reasons rather than because of their role in determining academic results. To represent this, let welfare differ across schools types in ways that again depend on family characteristics:

$$u_{ip}(1) - u_{is}(1) = u_{ip}(0) - u_{is}(0) = g(X_i, Z_i). \tag{5}$$

Then (3) reads

$$P_p(y_i = 1 \mid \eta_i) - P_s(y_i = 1 \mid \eta_i) > \frac{g(X_i, Z_i)}{f(X_i, Z_i)}.$$

It depends on observable family characteristics not only through the right-hand term but also, recalling (1) and writing the choice criterion as

$$P_p(\varepsilon_i > -X_i\beta_p - \gamma_p\eta_i) - P_s(\varepsilon_i > -X_i\beta_s - \gamma_s\eta_i) > \frac{f(X_i, Z_i)}{g(X_i, Z_i)},$$

on the parameters of school-specific success determination.

It is natural to let η_i and ε_i (and their sum) be normally distributed but, as long as $\gamma_p \neq 0$ or $\gamma_s \neq 0$, probit estimation of (1) yields inconsistent estimates. This is because variation of the η_i performance-relevant information that is unobservable to us, but known by the family when the school where we observe the student was chosen, implies that the unobservable $\gamma_p\eta_i + \varepsilon_i$ and $\gamma_s\eta_i + \varepsilon_i$ determinants of individual academic success are correlated with unobservable determinants of school choice.

To correct for self-selection bias the specification may include a control function based on approximating of the nonlinear criterion (6) with another probit equation. Denoting with μ_X and μ_Z the coefficients of observables in that approximation, noting that each i's characteristics are observed either in a private school as X_{pi} or in a state school as X_{si}, and denoting with $\{\zeta_{0i}, \zeta_{pi}, \zeta_{si}\}$ the realization of a trivariate normal random variable, the resulting system of probit equations is

$$\begin{aligned}
&j_i = p \text{ for } X_i\mu_x + Z_i\mu_z + \zeta_{0i} > 0,\ j_i = s \text{ otherwise;} \\
&\text{if } j_i = p \text{ then } y_1 = 1 \text{ for } X_{pi}\beta_p + \zeta_{pi} > 0,\ y_1 = 0 \text{ otherwise,} \\
&\text{if } j_i = s \text{ then } y_1 = 1 \text{ for } X_{si}\beta_s + \zeta_{si} > 0,\ y_1 = 0 \text{ otherwise,}
\end{aligned} \tag{7}$$

Because no individual is observed both in a private and a State school, the correlation between ζ_{pi} and ζ_{si} is not identified; other parameters (up to a scale factor) and predicted probabilities can be estimated with the switch_probit Stata command (Lokshin and Sajaia, 2011).

The data can be informative about cross-school differences not only through comparisons of properly estimated within-school relationships between individual characteristics and academic performance, but also through the school choices of heterogeneous individuals. Writing the choice criterion as

$$\cdot \Phi\left(-\frac{X_i\beta_p + \gamma_p\eta_i}{\sigma_p}\right) - \Phi\left(-\frac{X_i\beta_s + \gamma_s\eta_i}{\sigma_s}\right) > \frac{f(X_i, Z_i)}{g(X_i, Z_i)} \tag{6}$$

we see that the mean of η_i in each school type, for given parameters and covariates, is influenced by truncation of its school-specific distribution.

It is interesting to note that if γ_p is positive but smaller than γ_s then, at a given level of other variables, a small η_i makes private school enrolment preferable and academic success less likely. The choice criterion, however, also depends on the variances of the unpredictable disturbances, and on the observables variables and parameters that determine the point at which success probabilities are evaluated, and the preference-related right-hand side of (6). Whenever the X_i individual characteristics are differently relevant to success across schools, they should appear as explanatory variables in the first-stage school choice equation. If they happen not to be not directly relevant to school choice, so that the right-hand side of (6) only depends on Z_i, then the sign and significance of the first-stage coefficients or marginal effects of X_i variables provides information on differences across schools of each variable's influence on the probability of success. In particular, observable indicators of ability to succeed in school should, like unobservable ones, imply selection into schools where ability is less relevant. In general, however, the first-stage equations μ_x parameters do not have a structural interpretation, because some or all of the performance-relevant individual characteristics may also influence the welfare relevance of school success and/or the relative appeal of different schools. The selection equation's probit functional form is not structural, and identification of the other equations' parameters requires observable sources of choice variation that do not directly matter for success, i.e., variables that populate the Z_i vector.

Exclusion restrictions can also make it possible to estimate by two-stage least squares linear probability approximations of (1) and (6). Dummies for observed success, $y_i \in \{0,1\}$, and private school enrolment, $p_i \in \{0,1\}$ can be the dependent variables of linear regressions that include $X_{pi} = p_p X_i$, the interaction between individual success-relevant characteristics and the private school attendance dummy. In the resulting linear system

$$y_i = X_i \beta_s + X_{pi}(\beta_p - \beta_s) + \upsilon_{1i} \qquad (8)$$
$$p_i = X_i \mu_x + Z_i \mu_z + \upsilon_{2i}$$

the disturbance $\upsilon_{1i} = \gamma_j \eta_i + \varepsilon_i$ (plus the specification error induced by the linear approximation) is correlated with the private school choice that is embedded in X_{pi} and, as discussed above, is influenced by η_i through υ_{2i}. It is possible to estimate β_p and $\beta_p - \beta_s$ by two-stage least squares if sufficiently many elements of Z_i are available to instrument the X_i variables for which the coefficients are allowed to differ across State and private schools.[3]

In what follows we further discuss functional form and identification issues, in the process of applying these estimation techniques to the available French data.

5. Sample and variables

The data report whether the student is attending a private school during each year of the secondary curriculum. The information in `secteur1995` can be coded into a dummy variable `priv1995` that equals unity if the student is a private school at the start of the panel, i.e., during the first *collège* year, that is an empirical counterpart of observed school choices in specifications in the forms (7) or (8). Almost 19% of the panel students are observed to choose private school enrolment, in line with aggregate statistics. Students older than the normal age of *collège* entry, having repeated school years in primary school, can be identified because variable `datenai` reports the year of birth. Only 13224 students were born in 1984: more than a quarter of the panel individuals have repeated at least one primary school year; in the data, they appear less capable and less successful, and somewhat less likely to enrol in private school (only 16% of them do). We refrain from investigating them in this paper, and restrict

[3] It would also be possible to use a linear probability specification only for (6), and estimate interacted specifications of (1) by the Stata command `iv_probit`.

attention to the more homogeneous sample of students who have not repeated any year by the time they begin to be observed.

Available data record a variety of success indicators that may serve as empirical counterparts of the y_i outcome that in the model can be differently relevant for families with different socio-economic characteristics. A question in the post-*baccalauréat* surveys elicits information about the higher education track in which each panel individual is enrolled after exiting secondary school. The students reach that stage at different ages, depending on whether they have repeated some years in secondary school. In what follows we single out students who enter any higher education track by the time the post-*baccalauréat* survey is administered: `he_2002_4` equals one for students who start general or selective higher education tracks between 2002 and 2004. This is observed to be the case only for 7052 (53%) of the 13224 students who enter *collège* at normal age; success is more frequent at 58% for the 2597 (19.6%) of them who enrol in a private *collège.*[4]

Variation in terms of characteristics that predict academic success, included in vector X_i above, is observable though indicators of cognitive skills at the beginning of the secondary school curriculum. We focus in particular on the school principal's assessment of the student's skills at the time of entry in secondary school: variable `talent` is the average of the student's level in mathematics and in reading French at the entry of secondary school, on a 0-10 scale that is supposed to refer to France in general rather than to what is observed in the specific school. Its average is 6.84 among State school students and almost identical (6.85) among private school students, but observable skills are widely heterogeneous within each group, with standard deviations of 1.74 and 1.58 respectively, and positively correlated with academic success: Figure 1 plots all observations of `talent` and `he_2002_4` along with an unsurprisingly and strongly increasing nonparametric estimate of the relationship between these indicators of initial skills and final success.

Information on the family's socio-economic status and cultural level and can be used to assess how variation in those respects influences the objectives and constraints of educational choices. The panel data reports a classification in of the family head's occupation, reconstructed by the data provider on the basis of information from both school records and the family survey. Figure 2 shows that socio-economic status is strongly related both to the skills of students at the

[4] Non-response to the student survey is coded as a zero, which might mislabel as a failure the performance of students who e.g. drop out of the panel because they enrol in foreign higher education institutions..

beginning of secondary school (assessed by the group-specific mean of `talent` on the horizontal axis) and to the probability of academic success (assessed by the group-specific mean of `he_2002_4` on the vertical axis). It is also clear that and that within each group initial skills are widely heterogeneous, and that these two measures of cognitive skills are strongly and almost linearly positively correlated on a group-average basis, with some interesting and interpretable outliers: children of arts and media workers are on average smarter at age 10 than one would think on the basis of their observed later academic achievement, and the opposite is true for children of farmers and foremen.

This descriptive evidence need not reflect a direct influence of socio-economic status on educational outcomes. Some of it is mediated by the school choices that we aim to model and may be shaped by financial conditions, and some is spuriously driven by the correlation of job market outcomes with the family's cultural climate. The educational level attained by the parents, at given socio-economic status, may convey information that is relevant to school performance and school choice: Better educated parents indicate both that the child's background is likely to make learning easier, and that help is available from the parents when difficulties arise. The base specification we report includes dummy variables that take value 1 when the father or the mother obtained a tertiary degree: `he_father=`1 when A16P=8 or 9, `he_mother=1` when A16M=8 or 9.

Another observable exogenous determinant of school choice and achievement is gender, which in the data is strongly correlated with school choices and outcomes. Figure 3 shows that female students are somewhat differently likely to choose upper secondary tracks that are more likely to lead to tertiary education, and more likely to enroll in tertiary education within each track. While in this paper we do not model such choices as endogenous elements of strategies aimed at achieving higher education and life objectives, it would be inappropriate to exclude gender from the determinants of the higher-education outcomes and private school choices we do model.[5]

[5] Gender may be relevant also because parents particularly value the socially selected and well-disciplined student population of private schools for their female children.

6. Results

We proceed to implement the specifications outlined above, aiming to assess whether observed skills and background influence school outcomes differently in private and State schools (and, therefore, also influence school choice).

6.1 Linear models

An indication of these differences is offered by estimates of the approximate linear specification (8), which has two important advantages over the probit system (7). First, its linearity makes it straightforward to control for within-country variation including fixed geographic effects: individual characteristics are allowed to matter only to the extent they differ from those observed in a group of similar individuals that is likely to be faced by common educational supply conditions and influenced similarly by other socio-economic factors, and only heterogeneous in terms of financial and cultural conditions.[6] Second, the estimates provide a sharp snapshot of differences between State and private school on a sample-average basis: because they do not allow estimated effects to differ across students and schools other than through explicit interactions, the estimates they deliver do not need to account for other observable or unobservable heterogeneity.

For identification the system requires restrictions on the number of coefficients that may differ across private and State schools. In specifications that allow only the intercept and the coefficient of assessed initial skills (variable `talent`) to differ across State and private schools, at least two instruments are needed, and the panel's family survey data contain two individual variables that are plausible empirical counterparts for the data generating model's Z instruments.

After accounting for variation of factors that influence private school enrolment through its contribution to educational outcomes, an indicator of the particular relevance of other factors that generate demand for private education (such as geographical proximity, sporting facilities, and ideology) is attendance of private pre-primary or primary schools. Relevant information is gathered by two questions in the 1998 survey of families: we code the dummy `priv_b` to take value zero if replies to B3 and B7 are both 1="entirely in State school" or missing, so that

[6] Fixed geographical effects play a similar role in Burgess et al's (2015) study of school choices in England, where families may apply for enrolment in specific public primary schools; the preferences of the roughly 5% of English families who choose private schools are not estimated in that paper.

`priv_b=1` indicates that the student had at least some pre-secondary private schooling. The association between this and further private schooling may capture the cultural inclination of the family, or inertia, as well as the relevance to secondary school choices (at given other observable sources of variation) of features – such as full-time attendance or low likelihood of strikes – that do not directly matter for school results.

Household financial resources are plausibly relevant to the choice of paying the fees levied by private schools, and not directly relevant to school outcomes. Income, wealth, and tuition fees are not reported in the data, but the relevance of financial consideration is captured (at given levels of parents' education and occupational status) by a 1998 survey question. We code this information in the dummy `fic`, taking value 1 if A26=1 indicates that family's resources are very far from sufficient to allow the child to pursue his or her studies for as long as (s)he wishes. The identifying assumption is that, after controlling for other covariates, this information does not directly matter for school results, and is of course debatable in that (everything else equal) better financial resources might signal ability rather than luck. [7]

Table 1a reports linear regression results with fixed effects at the level of LEM cells defined by Department and city-size indicator, the most detailed geographical location information available in the data, to control for all local confounding factors. The OLS estimates of the interacted regression, reported for comparison purposes, find enrolment in a private lower-secondary school to be significantly and positively related to tertiary education enrolment after controlling for initial skills, parental education, and gender (all extremely significant), and do not detect any significant interaction with initial skills. Treating private secondary school enrolment as an endogenous variable and instrumenting it with private primary school enrolment and financial constraints indicators (either or both of which are missing for about 3% of the observations) only slightly reduces the relevance of parental education; but it has a dramatic and interesting effect on the estimated private school and individual skills interaction: the impact of initial skills on eventual success is strongly and negatively influenced by private school enrolment. This indicates that private schooling can remedy individual shortcomings, and consistently with this in the first stage it is not only predicted by primary school enrolment and prevented by financial constraints, but also negatively associated with initial assessed skills.

[7] A dummy `nofic` valued 1 if A26=4 indicates that financial resources are more than sufficient turns out to be a very weak instrument for private school choice, consistently with the fact that French private schools are not very expensive.

The regressions reported in Table 1b allows dummies for all socio-economic status categories to enter the OLS specification and either as included or excluded variables of 2SLS specifications (coefficients not reported). In the first two columns this has a very small effect on the coefficient estimates of the more interpretable variables already included in Table 1a, which continue to indicate that problematic students find private schooling attractive and choose it if they can afford it. Interestingly, in the 2SLS regressions the parental education is a more strongly significant determinant of school success when socio-economic status categories are excluded from the second stage; in the first stage, the financial constraints indicator is no longer significant when socio-economic categories are included as determinants of private school enrolment. These results suggest that (at given levels of parental education) socio-economic status categories capture financial heterogeneity that is relevant to private school enrolment, but is only mildly and collinearly directly related to school achievement.

6.2 Summary indicators of socio-economic and geographical variation

To provide a more interpretable indication of these and other effects, the following specifications replace the complete set of socio-economic categories dummy coefficients with a binomial indicator of privileged family background: the dummy `ses_h` equals unity for the socio-economic categories that in Figure 2 on average issue relatively skilled and successful children.[8] We do not report the very similar result this yields for the regressions reported in Table 1b. Table 1c displays the results obtained when `ses_h` appears as an instrument along with its interaction with a LEM-level variable computed from BCS administrative information: `badLEM` is the proportion of the student population that is enrolled in problematic `ZEP`, REP, or *établissement sensible* problematic schools. This variable cannot appear in levels when LEM fixed effects are present, but its interaction with `ses_h` is a significantly positive predictor of private school enrolment. This sensibly indicated that private school enrolment is chosen not only by families with slow-learning children (as indicated by the negative coefficient of `talent` in the first stage regression) and no stringent financial constraints (as indicated by the negative coefficient of `fic`), but also by high socio-economic status families that otherwise would have to send their children to underprivileged local public schools.

[8] In the sample of normal-age students with survey information, 3301 (37%) students belong to these socio-economic categories, which are: 23.Entrepreneur with >9 employees; 31.Self-employed professional; 32.Public sector executive; 33.Teacher, secondary and tertiary; 34.Scientist; 37.Private sector executive, administrative; 38.Private sector executive, technical; 42.Teacher, preschool and primary; 43.Paramedic or social worker; 44.Clergy; 45.Public sector supervisor, administrative; 46.Private sector supervisor, administrative; 47.Technician.

To similarly try and interpret the geographical fixed effects captured by location dummies in the specifications reported above, in linear regressions specifications fixed effects can be replaced by variables that capture relevant geographical variation. In order to exploit variation across the "Local Educational Market" environments in which individual panel students make their choices, we further characterize with administrative data the schools observed in that environment. LEM characteristics may plausibly change both across town sizes and across towns of the same size within each department: for each of the 456 cells defined by the *département* and size of town where the individual students surveyed in the panel are observed, we compute `p_hses`, the share of students who belong to the privileged socio-economic categories coded as `ses_h` for panel individuals, and `p_priv`, the share of students enrolled in private schools. To give a sense of the extent to which these indicators indeed vary across French localities, Figures 4 and 5 display their variation across departments and city sized separately.

Because these frequencies are computed across all school in the LEM, whether or not they were chosen by panel individuals, they characterize the environment in which those individual choices were made. The information is imprecise, because schools might have been chosen from narrower or broader area of feasible commuting, but arguably relevant. Including the area's average socio-economic characteristics along with the individual family's cultural and socio-economic status captures the idea that the latter is relevant to school choices when measured relative to that of classmates the family would like to have or avoid for their children. In the first stage, `p_hses` also approximate the availability and appeal of higher-education opportunities, both likely to be stronger in richer areas. After controlling for socio-economic characteristics, the LEM-level incidence of private schooling may plausibly capture local supply effects and cultural characteristics. Under the identifying assumption that these are not directly related to school outcome, the incidence of private schooling can be an instrument along with private primary school (capturing cultural characteristics specific to the family within the LEM) and financial constraints.

Table 2 reports the results, which are similar to those displayed in previous tables when the variables are the same and offer additional insights. The LEM proportion of privileged students has a moderately significant and sensibly positive coefficient as a determinant of tertiary school enrolment. It is insignificant in the first stage, where secondary private school enrolment is determined by the regional prevalence of private schooling as well as by the other instruments introduced above. While these interpretable results offer an interesting perspective on economic

and cultural determinants of school choice and performance, the testable restriction that the fixed effects estimated by the previous specification are well approximated by observed regional variation is unsurprisingly rejected by the data in the first and second stage. Of course, other characteristics of the location directly affect higher education and other outcomes, and there must be unobservable geographical variation in the characteristics of private education and in various factors that influence outcomes directly (such as the presence and quality of universities and other tertiary institutions).

All the estimates reported so far deliver a consistent and intriguing message: private secondary schooling has different effects on academic performances along the distribution of talents assessed when entering secondary school. The estimated main and interaction effects of private schooling and skill assessment are very similar in all specifications and, if taken at face value, suggest that private schooling is beneficial only for below-average students, and has a negative effect for students assessed above about 7/10 at the beginning of secondary school.

While this is an intriguing and fairly robust result, the linear approximation has shortcomings that are very apparent in the application proposed here. First, the linear specification requires an a priori choice of which coefficients may differ across State and private schools, the number of which is severely constrained by availability of suitable instrumental variables: allowing the intercept and k slopes to differ requires $k+1$ instruments that should predict rather collinear interaction effects in suitably distinct ways, and it is difficult for the linear specification to reliably estimate other interaction effects.[9] Second, sharpness of the estimates comes at a heavy cost in terms of approximating a more complex reality; all the specifications we reported so far predict success probabilities that are often very negative, or largely exceed unity.

6.3 Probit specifications

As usual, the linearized system "*is easy to interpret, providing parameter estimates that do not require transformation to learn the effects of a regressor on the mean of the dependent variable. It has a role to play in empirical work in summarizing the data as regards the conditional mean function and for initial explorations of the data, and virtually all practitioners use it for this purpose. But beyond this it has no defense*." (Moffitt, 2001).

[9] The coefficients of parental education may plausibly differ across private and State schools if they might if they capture availability of help at home, rather than ability to learn. Many instrumental variables are available in the regression displayed in the third column of Table 1b, which excludes all socio-economic categorical dummies from the second stage; but including as endogenous regressor to include interactions of private school choice with parental education and with gender estimates insignificant coefficients for all of them.

We can and do move beyond and estimate the nonlinear selection model (6), which allows all coefficients of performance-relevant variable to differ across State and private schools, and report results of specifications that approximate LEM variation with BCS indicators offers results that are complementary to and consistent with those or linear specifications.

Table 3a reports results of separate probit estimation on the subsample of private and State panel students: like the OLS estimates reported in previous tables, these are biased by selection effects but useful for comparison purposes. In particular, it is interesting to note that the constant term of single-equation probit specifications is more negative in the private than in the State sample, which would superficially suggest that (given observable characteristics) private schools deliver worse educational outcomes.

The opposite relationship between the constant terms is however observed in Table 3b, where the two outcome probit equations are estimated controlling for selection bias with another probit where private school enrolment is determined not only by outcome-relevant variables but also, with significant and sensibly signed coefficients, by the individual student's previous and LEM-level private enrolment, by the family's financial constraints, and by the prevalence of problematic schools and its interaction with socio-economic status. The intercept is now less negative in private schools: their students appear less likely to succeed in Table 3a (given observables) not because private schools are bad, but because the students are unobservably bad in ways that lead them to choose private schooling. Consistently with this observation, the estimated correlation `rho1` between unobservable determinants of success in private school and unobservable determinants of private school choice is significantly negative, while `rho0` is positive (if insignificantly so). As discussed in Section 3, so in the data failure for unobservable reasons is more likely when private school was chosen for unobservable reasons, because not only observably (as suggested by the negative coefficient of `talent` in the choice probit) but also unobservably weak students tend to enroll in private schools when it is possible for them to do so.

Extension and further interpretation of these results is more difficult and complicated than in the linear specifications above. It is interesting to see in Table 3b that the father's education does not appear to significantly determine success in private schools, assessing the robustness of this result would require additional investigation. Specifications with fixed-effects estimation are numerically challenging for nonlinear functional forms; while the simple binomial indicator `ses_h` may not adequately represent the relevant financial and cultural socio-economic factors, and is not a significant determinant of private school choices, it does

attract significantly positive coefficients as a determinant of success in similar regressions that include it in the second stage along with parental education indicators - which, like the other included variables, have very similar coefficients and significance levels in this specification.

As to interpretation, differences of coefficients across private and State schools cannot be detected reliably by testing whether $\beta_p = \beta_s$, because the variance of unobservable factors is not identified and may plausibly differ across the groups of individuals observed in different types of schools (Allison, 1999). It is instead possible and interesting to compare the predicted probabilities of academic success across different schools, which in the choice criterion (3) contribute to determine school choice regardless of whether they are implied by differences of slope parameters or variances. The difference between success probabilities (i.e., the treatment effect of private schooling) provides information on cross-school heterogeneity that is more precise than that of the linear probability approximations, but also more complicated to convey. Since probabilities depend nonlinearly on covariates, it is necessary to compare them at specific valued of the latter. The statistical significance of their difference can be assessed by a linear approximation of their covariance matrix (Scott, 2009), or by evaluating the normal probability distribution at the confidence bounds of the predicted latent variable; the 95% confidence intervals plotted in the Figures are computed with the latter method.

Figure 6 shows, using the estimates reported in Table 3b, that for a student with sample-mean levels of all covariates except gender (set to "male" for concreteness) tertiary enrolment is more likely at private schools, significantly so at intermediate levels of assessed initial skills. Figure 7 considers two less hypothetical individuals, both male, one with highly educated parents and one without any tertiary-educated parent. For the former, learning is likely to be easier and help at home is available if needed, and there is no significant difference between the effects of enrolment in State or private school. Private education significantly improves outcomes for the latter, who is likely to need help and (depending on financial and local condition) may or may not be able to get it from a private establishment.

7. Conclusion [preliminary]

The empirical approach and results of this paper suggest that France is more similar to Italy than to Anglo-Saxon countries within the spectrum of school system configurations identified by Bertola and Checchi (2013), in that French private schools provide educational resources that appear to substitute rather than complement their students' ability to learn. The rich

information available in French data makes it possible to characterize this feature in a variety of specifications.

The strong role of socio-economic status and financial conditions in determination of school and life results is remarkable in a country that traditionally values equality of opportunity, and organizes its State school system to pursue it. Should France truly wish to level uneven family backgrounds, the results of this paper indicate that its State schools should try and provide more of the educational services supplied by private schools. In light of the institutional information reviewed in Section 2, achieving this may involve a reform of State personnel management.

Appendix 1: The cost of attending a private school in France

According to Merle (2012), little reliable information is available on the cost of attending private schools in France. Notably, the Fogec (*Fédération nationale des organismes de gestion de l'enseignement catholique*) does not report such information for its Catholic private school members.

Within a given private school, total cost differs across students depending on the services they use: one of the Merle (2012) case studies reports that, in a region characterized by a traditionally high share of private (Catholic) education, the basic cost of a secondary school year is 770 €, but the family may also pay 216€ for supervised afterschool classes (*études surveillées*), or 553€ for availability of a teacher to help with homework (*étude encadrée*); access to a locker costs 59€, parents pay for the children lunch more than in State schools, and additional fees are charged for specific curricula (such as international sections). The total cost is therefore much higher than the basic fee, and tends to increase with the education level attended (the option 'international section' costs only 212€ in primary school, but 1006€ for the last year of secondary school), and can easily exceed 2500€ for a final-year (*terminale*) student preparing for the *baccalauréat*.

Private schools also typically offer discounted fees to the families with more than two children (*familles nombreuses*); Merle (2012) reports discounts of about 20% of the fees for the third child and 40% for the fourth one. Poor families are also entitled to lower fees if they apply for access to the solidarity fund, which they may refrain from doing for reputational reasons.

The cost of private education also varies considerably across schools, and across geographical areas, with schools located in high-income areas charging higher fees. Merle (2012) finds that fees to be higher in regional capitals, and are particularly high in Lyon and in Paris (where the annual basic fees amount to 1826€ and all the other services supplied are far more expensive that those exemplified above for provincial schools: in Paris, after-school classes cost 1826 € and lunches 1156 € per school year).

Appendix 2: State school choice constraints in France

The *sectorisation* system introduced in 1963 assigned each student finishing primary school to a single lower secondary school (*collège*), on the basis of residence in that school's catchment area, and students finishing each *collège* to a single *lycée*. This made it easier to manage the massive expansion of French education at the time, and was consistent with the aim of ensuring that the quality of State education would be homogeneous across all schools.

In 1984, it became possible for parents to ask the Education Administration for an exception to this assignment rule, and at the same time schools were given wider autonomy in offering specific pedagogical projects such as remedial help (especially in underprivileged geographical areas), exotic language courses, and curricula that allowed intense sport activity. This led to a more polarized schooling system, with stronger concentration in each school of students with similar social background (Merle, 2012). Some 10% of the parents applied for an exceptional reassignment on average, but up to 20% in some areas. Reassignment was granted (to 60% to 80% of the applications in each area) if a slot was available in the targeted school and enough students remained in the default school.

In 1987 the assignment process was relaxed by the introduction of "free to choose" zones in places where the default assignment did not seem to generate too many and uneven exception applications. Initially, only about 11% of *collège*s were in such zones, and only five of twenty administrative districts in Paris. By the end of 1990, half of the *collège*s and a quarter of the *lycées* were allowed to enroll students from outside their area without a formal application for exception from the default assignment; in 1993 the Paris area was divided in only six districts, within each of which any school could in principle be chosen by students: 75% of the students of each school had to reside locally, but up to 25% could be enrolled from elsewhere, making it possible for prestigious schools to selectively admit good students. Along with increasingly loose criteria for formal derogation of remaining constraints, relaxation of *sectorisation* constraints generated intense strategic behavior by parents (who aimed to enroll their children in schools they perceived to be better) and by schools (that launched attractive pedagogical project aiming to attract good students).

In 1997 a strict automatic assignment principle was reintroduced, and remained in place until 2007.

References

Allison, Paul D. (1999) “Comparing logit and probit coefficients across groups” Sociological Methods & Research 28:2 186-208

Altonji, Joseph G.; Elder, Todd E.; Taber, Christopher R. (2005) “Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools” *Journal of Political Economy*, 113:1 pp.151

Anamian S., Bonnaud A., Lambertyn A., Vercambre M.-N. (2005) “Les disparités d’orientation au Lycée ” Education & Formation 72, 101-111.

Base centrale scolarité (BCS) - 2004, DEP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

Base centrale scolarité (BCS) - 2005, DEP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

Base centrale scolarité (BCS) - 2005, DEP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

Base centrale scolarité (BCS) - 2006, DEPP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

Bertola G., Checchi D., Oppedisano V. (2007) “Private School quality in Italy ” Giornale degli Economisti e Annali di Economia, 66 (3) 375-400.

Bertola, Giuseppe, and Daniele Checchi (2013) “Who Chooses Which Private Education? Theory and International Evidence” Labour 27 (3) 249–271.

Breen, Richard and John H. Goldthorpe (1997) “Explaining Educational Differentials: Towards a Formal Rational Action Theory” Rationality and Society 9(3) 275-305.

Brunello, Giorgio, and Lorenzo Rocco (2008) “Educational Standards in Private and Public Schools" Economic Journal 118(533) 1866-1887.

Burgess, Simon, Ellen Greaves, Anna Vignoles and Deborah Wilson (2015) “What Parents Want: School Preferences and School Choice” Economic Journal, in press.

Cayouette-Remblière Joanie, and Thibaut de Saint Pol (2013) “Le sinueux chemin vers le baccalauréat : entre redoublement, réorientation et décrochage scolaire” Economie et Statistique, 459, 59-88.

Coleman, J.S., T.Hoffer and S.Kilgore. (1982). High School Achievement: Public, Catholic, and Private Schools Compared, New York: Basic Books Inc.

De Fraja, Gianni (2002) “The design of optimal educational policies”, Review of Economic Studies 69 437-466.

Evans, William N., and Robert M. Schwab (1995) “Finishing High School and Starting *Collège*: Do Catholic Schools Make a Difference?” Quarterly Journal of Economics 110(4) November 1995, pp. 941-74

Felouzis, G. (2005) “Performance et `valeur ajoutée’ des lycées: le marché scolaire fait la différence”, Revue française de sociologie 46, 3-36.

Lokshin, Michael and Zurab Sajaia (2011) “Impact of interventions on discrete outcomes: Maximum likelihood estimation of the binary choice models with binary endogenous regressors” The Stata Journal 11:3, 368–385.

Long, J. Scott (2009) “Group comparisons in logit and probit using predicted probabilities” working paper, Indiana University

Maetz, I. (2004) “Public et privé : flux, parcours scolaires et caractéristiques des élèves” Education & Formation 69, 23-35.

Martínez-Mora, F. (2006). “The existence of non-elite private schools” Journal of Public Economics 90, 1505–1518.

Merle, P. (2009) *La démocratisation de l’enseignement*, Paris : La découverte.

Merle, P. (2012) *La ségrégation scolaire*, Paris : La découverte.

Moffitt, Robert A. (2001) “Comment“ on Angrist, Joshua D., “Estimation of limited dependent variables with dummy endogenous regressors : Simple strategies for empirical practice“ Journal of Business & Economic Statistics 19 :1, 20-23.

Nakhili N. (2005) “Impact du contexte scolaire dans l’élaboration des choix d’études supérieures des élèves de terminale” Education & Formation 72, 155-167.

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

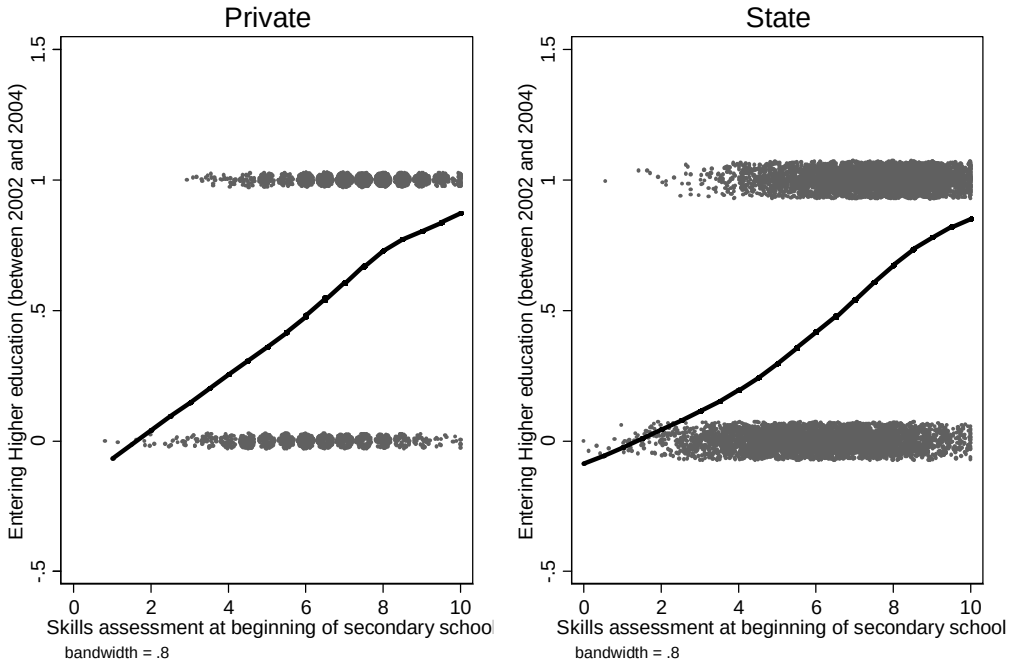
OECD (2012) *Public and Private Schools. How management and funding relate to their socio-economic profile*, Paris: Organization for Economic Cooperation and Development.

Panel d'élèves du second degré, recrutement 1980 - 1980-1989, DEP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

Panel d'élèves du second degré, recrutement 1989 - 1989-2003, DEP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

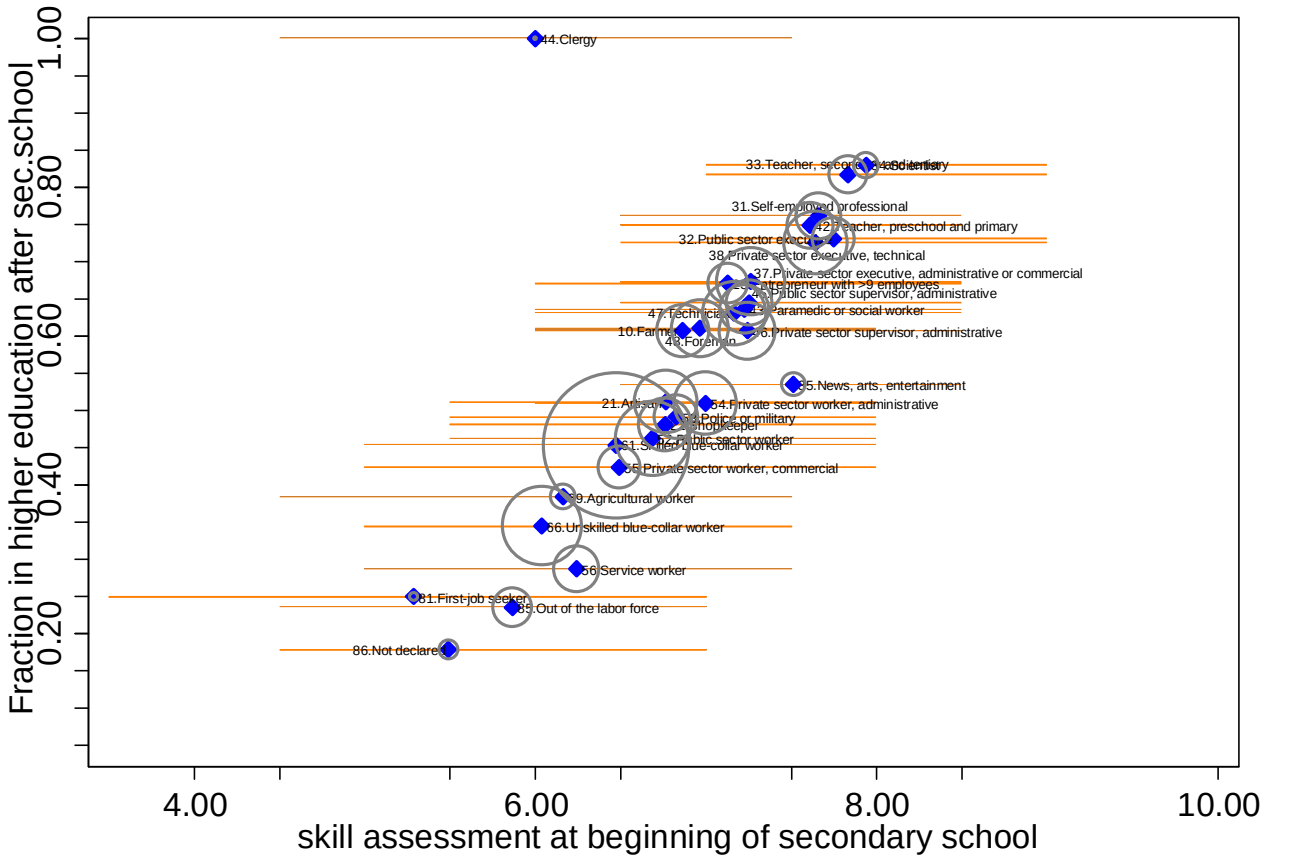
Panel d'élèves du second degré, recrutement 1995 - 1995-2011, DEPP - Ministère de l'Éducation [producers], ADISP-CMH [distributor].

Vasconcellos M., Bongrand P. (2013) *Le système éducatif*, Paris : La découverte.



dots: outcomes; line: nonparametric smoother

Figure 1



label: Socio-economic classification of the family head; ball size: number of obs; line: interquartile range

Figure 2

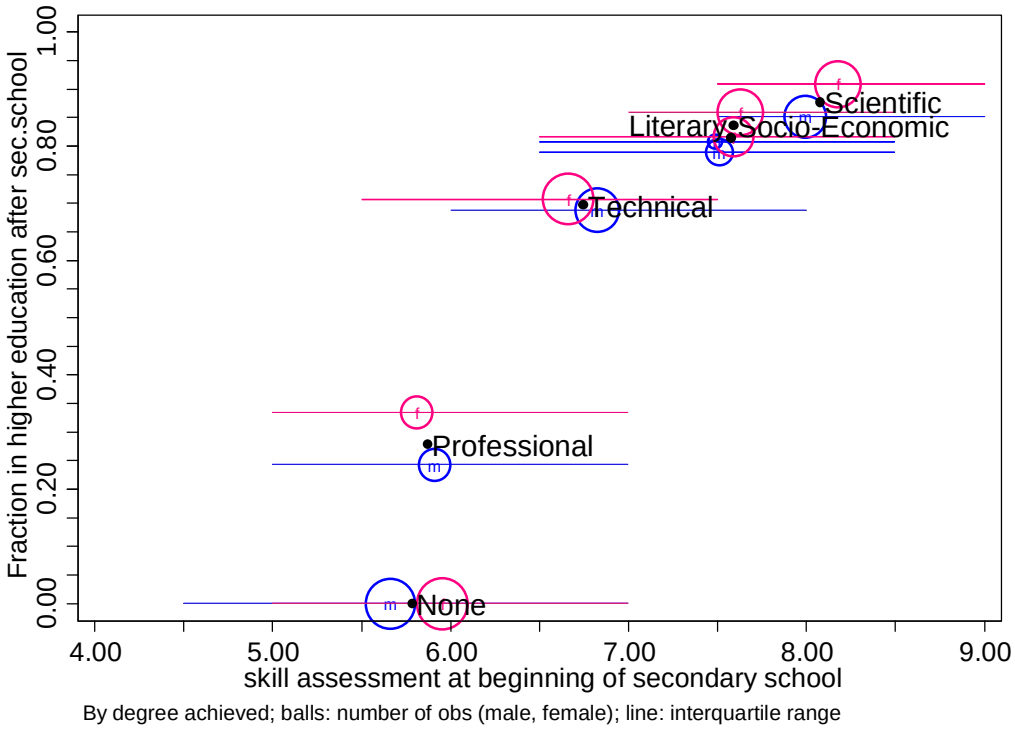

Fraction in higher education after sec.school
0.00
0.20
0.40
0.60
0.80
1.00
4.00
5.00
6.00
7.00
8.00
9.00
skill assessment at beginning of secondary school
Scientific
Literary
Socio-Economic
Technical
Professional
None
f
m
By degree achieved; balls: number of obs (male, female); line: interquartile range

Figure 3

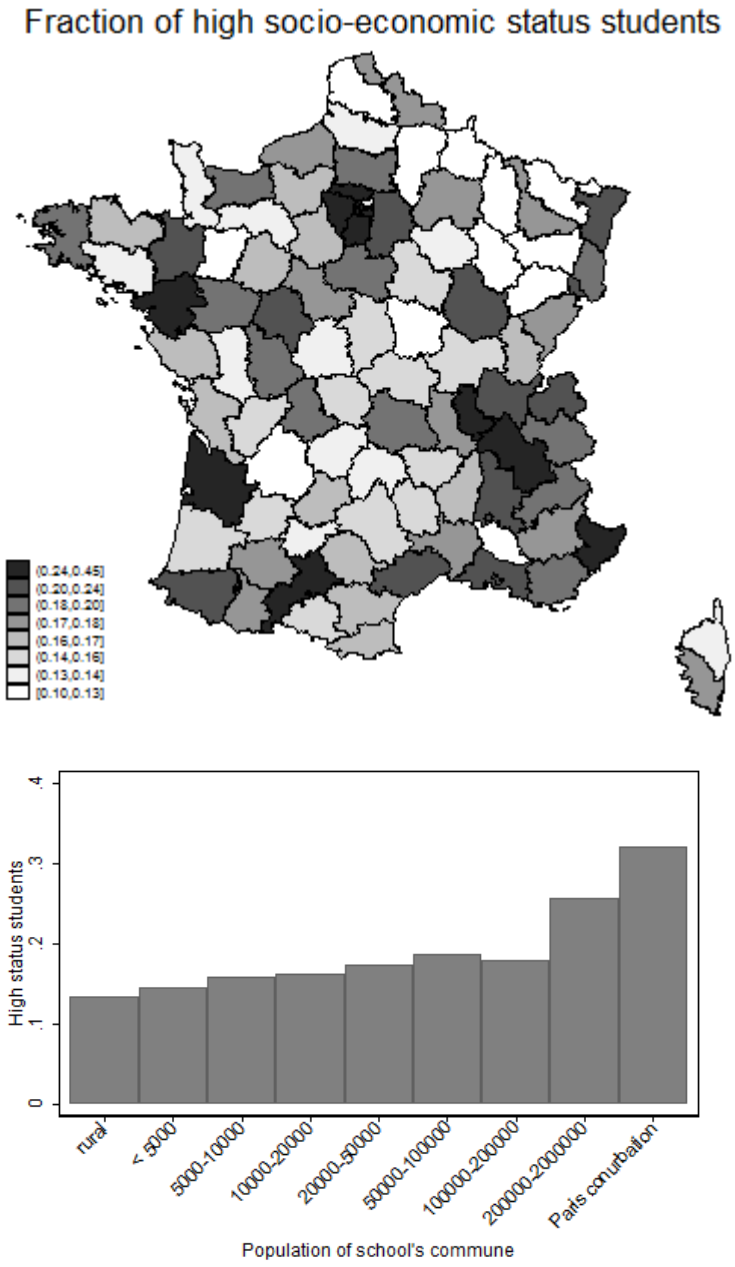

Fraction of high socio-economic status students
(0.24,0.45]
(0.20,0.24]
(0.18,0.20]
(0.17,0.18]
(0.16,0.17]
(0.14,0.16]
(0.13,0.14]
[0.10,0.13]
High status students
0
.1
.2
.3
.4
rural
< 5000
5000-10000
10000-20000
20000-50000
50000-100000
100000-200000
200000-2000000
Paris conurbation
Population of school's commune

Figure 4

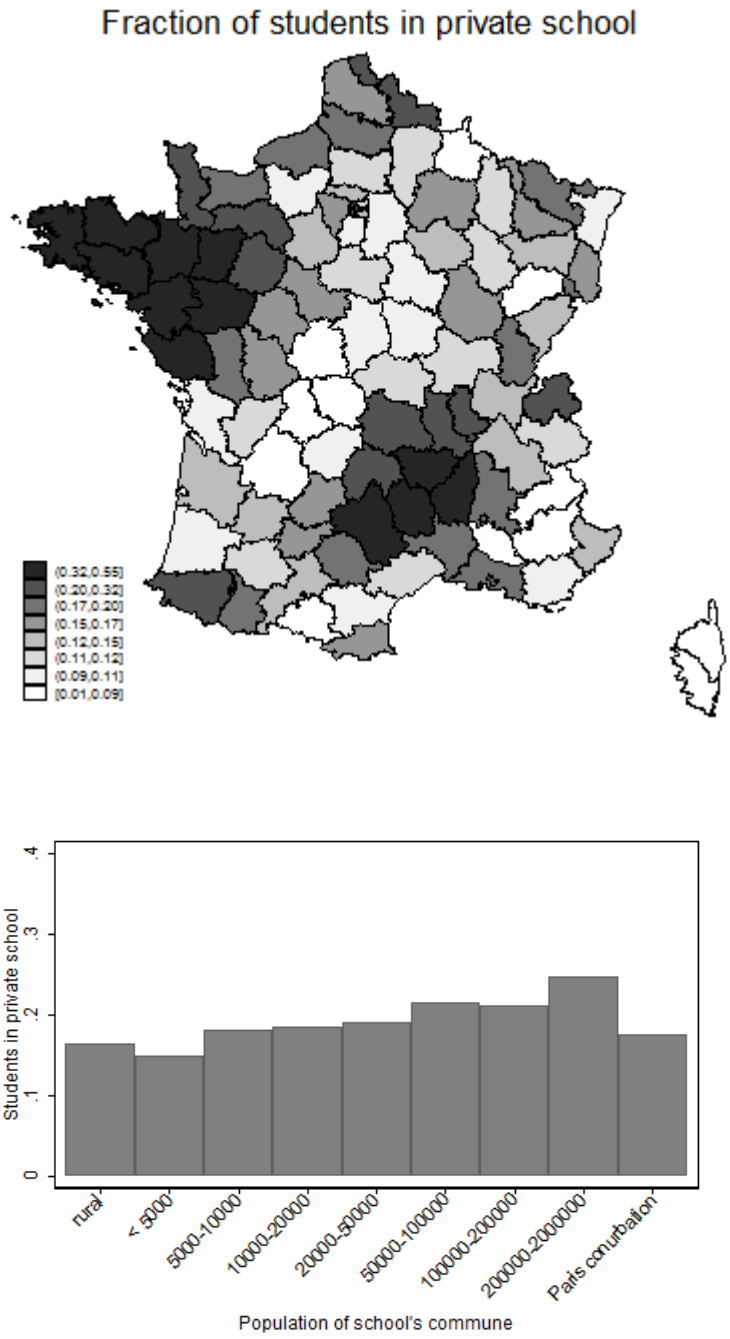

Fraction of students in private school
(0.32,0.55]
(0.20,0.32]
(0.17,0.20]
(0.15,0.17]
(0.12,0.15]
(0.11,0.12]
(0.09,0.11]
[0.01,0.09]
Students in private school
0
.1
.2
.3
.4
rural
< 5000
5000-10000
10000-20000
20000-50000
50000-100000
100000-200000
200000-2000000
Paris conurbation
Population of school's commune

Figure 5

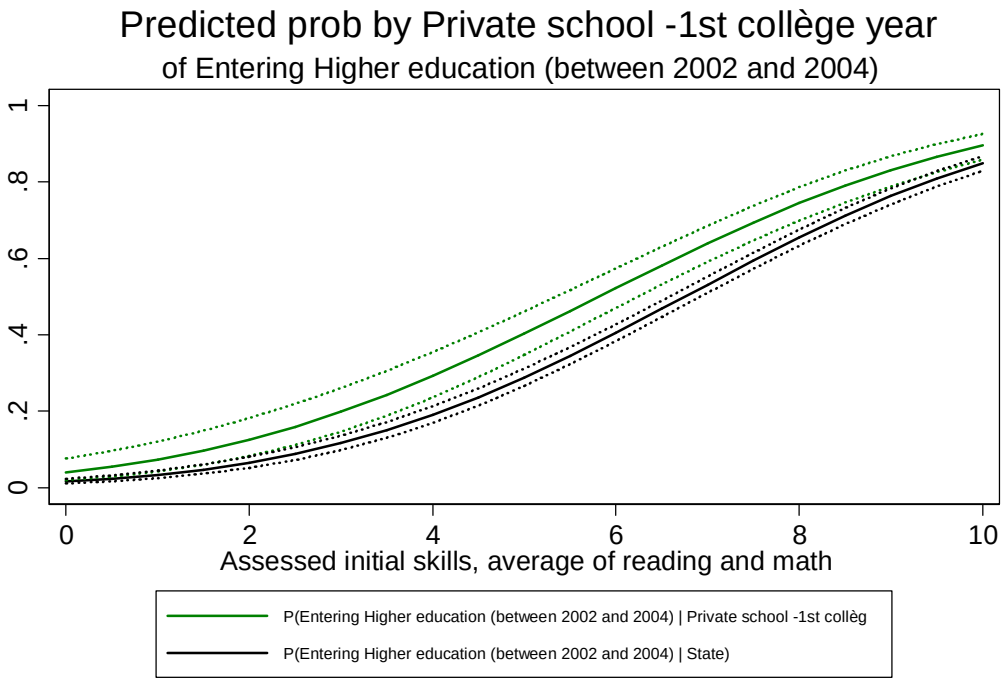
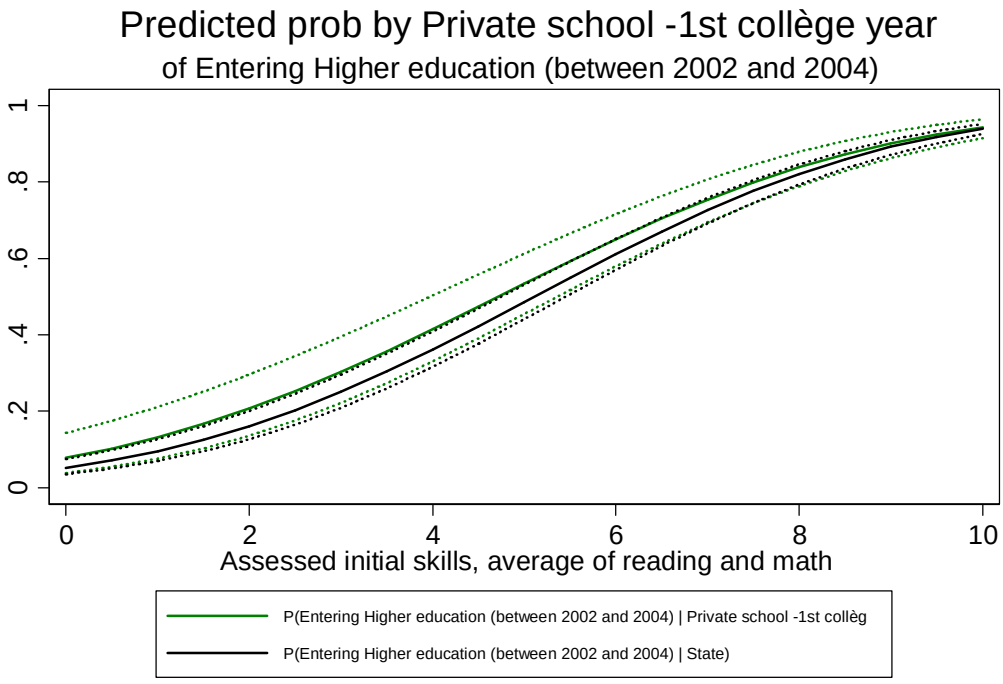
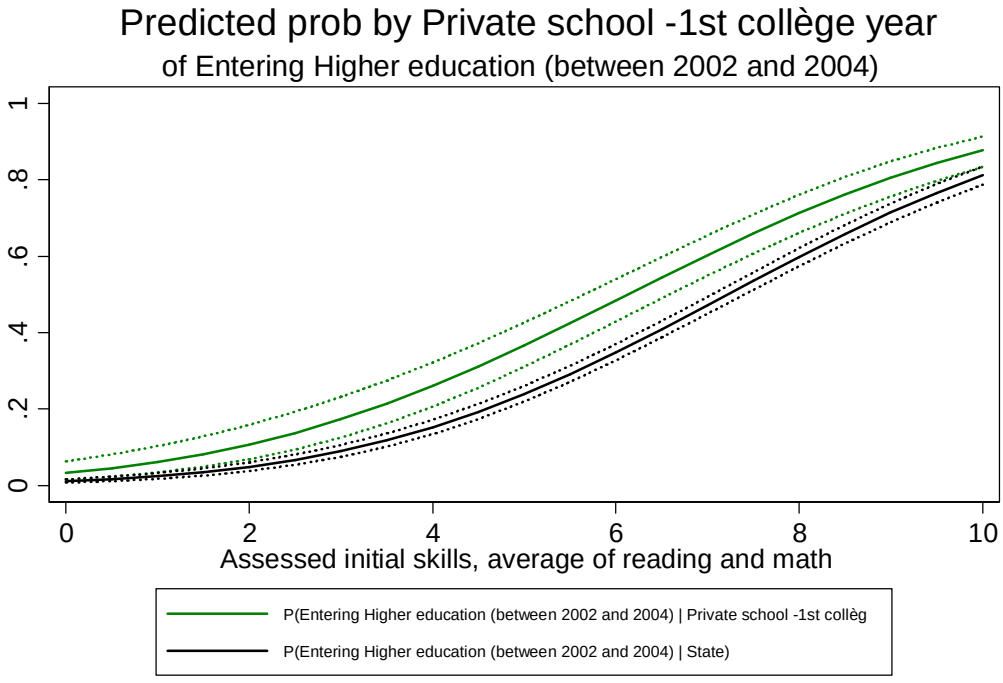

Predicted prob by Private school -1st collège year
of Entering Higher education (between 2002 and 2004)
0
.2
.4
.6
.8
1
0
2
4
6
8
10
Assessed initial skills, average of reading and math
P(Entering Higher education (between 2002 and 2004) | Private school -1st collèg
P(Entering Higher education (between 2002 and 2004) | State)
gender=male, other variables at sample mean

Figure 6



gender=male, both parents highly educated, other variables at sample mean



gender=male, neither parent highly educated, other variables at sample mean

Figure 7

Table 1a Dependent variable: in Higher Education by 2004

	LEM fixed effects	
	OLS	2sls
inpriv	0.108*	7.799***
	(0.05)	(2.08)
talprv	-0.010	-1.109***
	(0.01)	(0.30)
talent	0.109***	0.303***
	(0.00)	(0.05)
he_father	0.088***	0.081**
	(0.01)	(0.03)
he_mother	0.104***	0.085**
	(0.01)	(0.03)
fem	0.112***	0.142***
	(0.01)	(0.02)
		Instruments: fic priv_b
df_m	436.000	445.000
N	9224.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

First-stage regressions

	inpriv	talpriv
fic	-0.025**	-0.090
	(0.01)	(0.07)
priv_b	0.553***	3.881***
	(0.01)	(0.09)
talent	-0.005*	0.142***
	(0.00)	(0.01)
he_father	-0.000	0.002
	(0.01)	(0.08)
he_mother	0.009	0.056
	(0.01)	(0.08)
fem	-0.005	-0.009
	(0.01)	(0.05)
df_m	434.000	434.000
N	8941.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

Table 1b Dependent variable: in Higher Education by 2004

	LEM and SES fixed effects		LEM f.e.
	OLS	2sls	2sls
inpriv	0.084+	6.798**	4.458***
	(0.05)	(2.16)	(0.86)
talprv	-0.008	-0.968**	-0.631***
	(0.01)	(0.31)	(0.12)
talent	0.104***	0.275***	0.218***
	(0.00)	(0.05)	(0.02)
he_father	0.045**	0.049+	0.083***
	(0.02)	(0.03)	(0.02)
he_mother	0.085***	0.076**	0.093***
	(0.01)	(0.02)	(0.02)
fem	0.114***	0.139***	0.129***
	(0.01)	(0.02)	(0.01)
	Instruments: fic priv_b		fic priv_b SES f.e.
df_m	462.000	471.000	445.000
N	9224.000	8941.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

First-stage regressions

	inpriv	talprv
fic	-0.013	-0.017
	(0.01)	(0.07)
priv_b	0.544***	3.817***
	(0.01)	(0.09)
talent	-0.006**	0.134***
	(0.00)	(0.01)
he_father	-0.012	-0.076
	(0.01)	(0.09)
he_mother	0.004	0.024
	(0.01)	(0.08)
fem	-0.003	0.001
	(0.01)	(0.05)
LEM f.e.	YES	YES
df_m	460.000	460.000
N	8941.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

Table 1c Dependent variable: in Higher Education by 2004

	LEM fixed effects 2sls
inpriv	8.682***
	(2.01)
talprv	-1.235***
	(0.29)
talent	0.325***
	(0.05)
he_father	0.081**
	(0.03)
he_mother	0.083**
	(0.03)
fem	0.145***
	(0.02)
Instruments:	fic priv_b ses_h sesh_badLEM
df_m	445.000
N	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

First-stage regressions

	inpriv	talprv
fic	-0.022*	-0.078
	(0.01)	(0.07)
priv_b	0.552***	3.875***
	(0.01)	(0.09)
ses_h	0.003	-0.002
	(0.01)	(0.08)
sesh_badLEM	0.080*	0.508+
	(0.04)	(0.28)
talent	-0.006**	0.139***
	(0.00)	(0.01)
he_father	-0.009	-0.043
	(0.01)	(0.09)
he_mother	0.006	0.036
	(0.01)	(0.08)
fem	-0.004	-0.005
	(0.01)	(0.05)
df_m	436.000	436.000
N	8941.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

Table 2 Dependent variable: in Higher Education by 2004

	ols	2sls
inpriv	0.113*	8.877***
	(0.05)	(1.99)
talprv	-0.009	-1.259***
	(0.01)	(0.29)
talent	0.108***	0.329***
	(0.00)	(0.05)
he_father	0.089***	0.083**
	(0.01)	(0.03)
he_mother	0.102***	0.075**
	(0.01)	(0.03)
fem	0.112***	0.150***
	(0.01)	(0.02)
p_hses	0.162**	0.194+
	(0.05)	(0.11)
		Instruments: priv_b fic p_priv ses_h badLEM sesh_badLEM
df_m	7.000	7.000
N	9224.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

First-stage regressions

	inpriv	talprv
	b/se	b/se
priv_b	0.543***	3.817***
	(0.01)	(0.09)
fic	-0.026**	-0.104+
	(0.01)	(0.06)
p_priv	0.485***	3.344***
	(0.03)	(0.20)
ses_h	-0.004	-0.037
	(0.01)	(0.08)
badLEM	0.062**	0.474***
	(0.02)	(0.14)
sesh_badLEM	0.090*	0.515+
	(0.04)	(0.28)
talent	-0.005*	0.144***
	(0.00)	(0.01)
he_father	-0.009	-0.045
	(0.01)	(0.08)
he_mother	0.007	0.042
	(0.01)	(0.08)
fem	-0.004	-0.002
	(0.01)	(0.05)
p_hses	0.005	0.117
	(0.04)	(0.28)
df_m	11.000	11.000
N	8941.000	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

Table 3a Sector-specific probits
Outcome: Entering Higher education (between 2002 and 2004)

Starts secondary in	Private	State
p_hses	0.376	0.532**
	(0.37)	(0.18)
he_father	0.145	0.343***
	(0.09)	(0.05)
he_mother	0.326***	0.336***
	(0.09)	(0.05)
talent	0.304***	0.320***
	(0.02)	(0.01)
fem	0.286***	0.358***
	(0.06)	(0.03)
_cons	-2.060***	-2.432***
	(0.17)	(0.08)
df_m	5.000	5.000
N	1836.000	7388.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

Table 3b Switching probit

Outcome: Entering Higher education (between 2002 and 2004)

Starts secondary in	Private	State
p_hses	0.366	0.485**
	(0.37)	(0.19)
he_father	0.111	0.339***
	(0.09)	(0.05)
he_mother	0.314***	0.333***
	(0.09)	(0.05)
talent	0.301***	0.319***
	(0.02)	(0.01)
fem	0.281***	0.357***
	(0.07)	(0.03)
_cons	-1.921***	-2.402***
	(0.18)	(0.09)
df_m	5.000	5.000
N	1836.000	7388.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

rho1	-0.118
	(0.05)
rho0	0.036
	(0.06)
chi2_c	5.620 p-level 0.06

Outcome: Private school in 1st collège year

p_hses	0.281
	(0.20)
he_father	-0.049
	(0.06)
he_mother	0.024
	(0.05)
priv_b	1.736***
	(0.04)
fic	-0.165**
	(0.06)
p_priv	2.454***
	(0.14)
badLEM	0.537***
	(0.12)
SESh_badLEM	0.361+
	(0.20)
ses_h	-0.013
	(0.06)
_cons	-2.094***
	(0.06)
df_m	9.000
N	8941.000

(robust standard errors), p-levels +<0.1 <*0.05 **<0.01 ***<0.001

Table 3c Switching probit, allowing ses_h to appear in second stage

Outcome: Entering Higher education (between 2002 and 2004)

Starts secondary in Private		State	
p_hses	0.209	0.396*	
	(0.38)	(0.19)	
he_father	0.029	0.222***	
	(0.10)	(0.05)	
he_mother	0.288**	0.290***	
	(0.09)	(0.05)	
talent	0.298***	0.312***	
	(0.02)	(0.01)	
fem	0.290***	0.363***	
	(0.07)	(0.03)	
ses_h	0.176*	0.233***	
	(0.08)	(0.04)	
_cons	-1.920***	-2.386***	
	(0.18)	(0.09)	
rho1	-0.111		
rho0	0.052		
chi2_c	5.420	p-level	0.07
p_c	0.067		
df_m	9.000		
N	8941.000		