CHOOSING THE FIELD OF STUDY IN POSTSECONDARY EDUCATION: DO EXPECTED EARNINGS MATTER?

Magali Beffy, Denis Fougère, and Arnaud Maurel*

Abstract—This paper examines the determinants of the choice of the college major when the length of studies and future earnings are uncertain. We estimate a three-stage schooling decision model, focusing on the effect of expected earnings on major choice. We control for dynamic selection through the use of mixture distributions. Exploiting variations across the French business cycle in the relative returns to the majors, our results yield a very low, though significant, elasticity of major choice to expected earnings. This suggests that at least for the French university context, nonpecuniary factors are a key determinant of schooling choices.

I. Introduction

OVER recent years, the French postsecondary education system has been the subject of much debate and sharp criticism. In a report for the French Council of Economic Analysis, Aghion and Cohen (2004) emphasize the main difficulties that this system, and especially the French university, has to cope with. Pointing out, among others, the high dropout rate in French universities, they argue that the French postsecondary education system urgently needs to be reformed. In particular, French universities do not impose any selection at entry, which may partly explain the high failure rates prevailing at the college level. In this context, it seems especially worthwhile to understand how individuals choose their major when entering college.

In this paper, we focus on the effect of expected labor market income on the choice of the postsecondary field of study. In particular, we assess the sensitivity of students' major choices to expected earnings by estimating from French data a semistructural model of postsecondary educational choices. First, after graduating from high school, students entering university are assumed to choose a major. Then they keep on studying in this major until they reach a given level of education, before finally entering the labor market. We control for the dynamic selection occurring in the model through the use of finite mixture distributions. We try to disentangle the simultaneous effects of, on the one hand, preferences and abilities, and on the other hand, expected returns, on the choice of major. From a policy point of view, this question is also related to the skill composition of the labor force, and in particular to the efficiency of financial incentives as a solution to the scientific skills shortage prevailing in Europe.

In the existing applied literature, several papers explicitly consider the impact of expected labor market earnings on schooling choices. In a seminal paper, Willis and Rosen (1979) allow the demand for college education to depend on expected future earnings.[1] Assuming that students form rational expectations, these authors show that the expected flow of posteducation earnings is a strong determinant of college attendance. Berger (1988) also focuses on the impact of expected earnings on the individual demand for postsecondary education: his results show that when choosing college majors, students are more influenced by the expected flow of future earnings than by their expected initial earnings.[2] Then, following Keane and Wolpin (1997), several econometricians have estimated structural dynamic models of schooling decisions (see Eckstein & Wolpin 1999; Keane & Wolpin, 2001; Belzil & Hansen, 2002; Lee, 2005). Their papers assume that students form rational earnings expectations conditional on schooling decisions and that the expected earnings affect in turn their educational choices. More recently, Arcidiacono (2004, 2005) has considered sequential models of college attendance, accounting for both the demand and the supply side of schooling, in which the value of each major depends on the corresponding expected flow of earnings. However, in the literature noted above, Berger (1988) and Arcidiacono (2004, 2005) are the only ones focusing on the effect of expected earnings on the choice of major and not that of the educational level.[3]

Our paper contributes to the literature on the effects of expected earnings on schooling choices in several ways. First, unlike the previous papers, our approach concentrates on the effects of expected earnings on the choice of the major, in a framework in which the student does not know exactly the length of postsecondary studies when choosing a major.

Received for publication June 24, 2009. Revision accepted for publication September 20, 2010.

* Beffy: CREST-INSEE, France; Fougère: CNRS, CREST-INSEE, France, CEPR, United Kingdom and IZA, Germany; Maurel: Duke University, CREST-INSEE, France, and IZA, Germany.

We thank Christian Belzil, Moshe Buchinsky, Nicolas Chopin, Xavier d'Haultfoeuille, Francis Kramarz, Guy Laroque, Robert Miller, Jean-Marc Robin, Gerard J. van den Berg, participants in the IZA Workshop on "Heterogeneity in Micro Econometric Models" (Bonn, June 2007), in the Fourth Symposium of the CEPR Network Economics of Education and Education Policy in Europe (Madrid, October 2007), in the 11th IZA European Summer School in Labor Economics (Buch-Ammersee, May 2008), in the 63rd European Meeting of the Econometric Society (Milan, August 2008), in the 2008 EALE Annual Conference (Amsterdam, September 2008), in the 2009 North American Winter Meeting of the Econometric Society (San Francisco, January 2009), and in seminars at CREST-INSEE (Paris, January 2008), Université du Mans (January 2008), at Université Paris 1-Sorbonne (Paris, June 2008), and at DEPP (Ministry of Education, Paris, December 2008), for very helpful discussions and comments. We also thank the editor and two anonymous referees for their valuable remarks and suggestions.

[1] On a related ground, Altonji (1993) estimates a sequential model in which schooling decisions depend on expected returns to education, without explicitly considering the choice of major.

[2] Several other articles have shown some large differences in earnings across majors in the United States (see, for instance, James et al., 1989; Loury & Garman, 1995; Brewer, Eide, Ehrenberg, 1999). However, none of these papers models the choice of the major itself as a function of expected earnings.

[3] Freeman (1971, 1975), Manski (1993), and Boudarbat and Montmarquette (2009) depart from the rational expectations assumption by proposing models assuming that individuals have myopic expectations relatively to their potential labor market earnings. Within such a framework, students are assumed to form their wage expectations by observing the earnings of comparable individuals who are currently working.

The Review of Economics and Statistics, February 2012, 94(1): 334–347

Stylized facts appear to be consistent with such a framework.[4] Another interesting feature of our paper lies in the fact that we exploit the arguably exogenous variation across the business cycle in the relative returns to each major in order to identify the model parameters.

Using the parameter estimates of our model, we calculate the elasticities of major choices to expected earnings by simulating exogenous variations of the earnings distribution. These elasticities are found to be very low, which means that the choice of a major is mainly driven by nonpecuniary factors.

By allowing for heteroskedacticity in the variance of log earnings and by imposing a constant relative risk aversion (CRRA) utility function with a fixed risk-aversion parameter, we also estimate an additional specification of our baseline model, which yields similar results. Note that we ignore the possibility for the student to switch majors during postsecondary studies. Such a switch is potentially an endogenous event whose treatment would make the model much more complicated. Stylized facts actually show that this makes sense for the broad majors that we consider (see table 7, appendix A).[5]

The remainder of the paper is organized as follows. Section II describes our econometric model. The specification of the model and the likelihood function are discussed in section III. Section IV describes the data and presents some preliminary statistics, while section V presents the identification strategy. Section VI contains the estimation and simulation results. Finally, section VII summarizes and concludes.

II. The Econometric Model

After graduating from high school, individuals are assumed to choose their field of study (major) in which they complete a given (partly random) level of education (note that we restrict our analysis to individuals who attend college).[6] Once they leave postsecondary education, they are supposed to enter the labor market. Thus, we consider a sequence of three events:

Stage 1: When entering college, each student chooses her postsecondary major.

Stage 2: She keeps on studying in the major chosen in stage 1 until she reaches an endogenously determined level of education.

Stage 3: She leaves postsecondary education and participates in the labor market.

Following Heckman and Singer (1984), we assume that there are R types of individuals, Π_r denoting the proportion of type r in the population of students.[7] Individuals are supposed to know their type, which the econometrician does not observe. Within this framework, unobserved heterogeneity (that is, unobserved preferences for each major, unobserved schooling ability and unobserved labor market productivity) is type-specific.

A. *Stage 1: Choice of Major*

After graduating from high school (and getting the final high school diploma, the *baccalauréat* in France), the individual who decides to continue studying must choose a college major, hereafter indexed by j^*.[8] We assume that this choice is made among a set of M majors. Furthermore, we assume that the chosen major j^* depends on the individual's expectations concerning both the education level that she will achieve within this major and her future labor market earnings, which are assumed to depend on her educational level. An important underlying assumption is that future earnings, as well as the highest level of education reached in major j^*, are uncertain to the individual.[9]

For a student of type r, let us denote by V_j^r the value function associated with the choice of major j $(j = 1, \ldots, M)$. This value function is assumed to be composed of two additive elements, respectively denoted by v_{0j}^r and v_{1j}^r. The first term v_{0j}^r represents the intrinsic value (the consumption value) of the major, while v_{1j}^r is the investment value of a postsecondary education in major j. v_{1j}^r is proportional to the sum of the expected future labor market earnings, conditional on each educational level k $(k = 0, \ldots, L)$, weighted by the probabilities $\Pr(K = k \mid J = j)$ to reach each educational level within major j. Here $k = L$ denotes the highest educational level that can be reached within major j, and $k = 0$ corresponds to the case where the student drops out from the major before getting an associate degree. Then, for a student of type r, the value V_j^r of major j can be written as

$$V_j^r = v_{0j}^r + v_{1j}^r, \text{ for } j = 1, \ldots, M, \tag{1}$$

where

$$\begin{aligned} v_{1j}^r &= \alpha E\big((T-K)V_{e(j,K)}^r \mid r, J=j\big) \\ &= \alpha E_{K|r,J=j}\big((T-K)E\big(V_{e(j,K)}^r \mid r, J=j, K\big)\big) \\ &= \alpha \sum_{k\in\{0,1,\ldots,L\}} \Pr(K=k \mid r, J=j)(T-k) \\ &\quad \times E\big(V_{e(j,k)}^r \mid r, J=j, K=k\big), \end{aligned}$$

[4] Indeed, descriptive statistics from the French Panel 1989 database (DEPP, French Ministry of Education) show that most students complete a final level of education that is different from the level they wanted to reach when entering college (see Appendix A, table 8).

[5] This evidence is suggestive of the existence of high switching costs across majors.

[6] The argument justifying our choice to focus on individuals attending college is detailed in section IV.

[7] Examples of econometric models of schooling decisions relying on a similar assumption can be found in Keane and Wolpin (1997, 2001), Eckstein and Wolpin (1999), Cameron and Heckman (1998, 2001), Belzil and Hansen (2002, 2004), Arcidiacono (2004, 2005), and Lee (2005).

[8] We omit the individual subscript for simplicity.

[9] We suppose that each individual has an idiosyncratic propensity to achieve a high level of education. This propensity is partly affected by random factors, such as her own health status and unexpected changes in her family environment. These factors are ex ante unknown by the individual when choosing her majoring and then revealed when attending college.

with $V^r_{e(j,k)}$ denoting the average earnings associated with education (j, k), for a student of type r, T denoting the maximum length of working life (set equal to forty years), and α being an unknown sensitivity parameter to be estimated.[10]

The subcomponent v^r_{0j} can be interpreted as the nonpecuniary value of major j for a student of type r. It relates to the social gratification brought by studying in major j and to the individual's taste as well as ability for this major. Given that we do not model explicitly the postschooling dynamics, v^r_{0j} may also capture preferences for nonpecuniary aspects of future jobs associated with major j. We assume that v^r_{0j} is a linear function of a set of observable individual covariates X_1 that affect the attractiveness of major j (such as gender, place of birth, parents' nationality and profession, and the student's educational history, including the cumulated delay when entering junior high school). It also includes a type-specific intercept $\alpha^r_{(1,j)}$ and a random term u_j. Consequently, v^r_{0j} is specified as

$$v^r_{0j} = \alpha^r_{(1,j)} + X_1'\beta^j_1 + u_j,$$

where β^j_1 is a parameter vector associated with X_1 and specific to major j. The individual chooses the major j^* that corresponds to the highest-value function:

$$j^* = \arg \max_{j \in \{1,\ldots,M\}} V^r_j.$$

B. Stage 2: Determination of the Length of Studies

Once a student of type r has chosen her major j^*, she studies until she reaches a level k^*_j of education within major j. We assume that this level k^*_j is an element of a set of $L + 1$ possible levels that may be reached in each major; $k = 0$ corresponds to a dropout, which occurs when a student leaves college without any postsecondary degree; $k = 1$ refers to an associate degree (DEUG in France), which is obtained after two years of college; $k = 2$ corresponds to a bachelor's degree (*licence*); $k = 3$ corresponds to the first year of a master's degree (*maîtrise*); and $k = L = 4$ refers to higher levels—namely master's degree and Ph.D. (*post-maîtrise*).

The length of studies k^*_j within major j is supposed to be determined by the individual propensity $\widetilde{k}_j$ to succeed in long postsecondary studies within this major. More precisely, we assume that the length of studies k^*_j is such that $\forall k \in \{0, 1, \ldots, L\}, k^*_j = k \Leftrightarrow s_k < \widetilde{k}^r_j \leq s_{k+1}$, where $\{s_0, \ldots, s_{L+1}\}$ are thresholds that correspond to the minimum ability levels required to obtain the different degrees, with $s_0 = -\infty$ and $s_{L+1} = +\infty$. The latent propensity $\widetilde{k}^r_j$ is assumed to depend linearly on observable covariates $X_{2,j}$. It also depends on a type-specific intercept α^r_2 and an independent term η that is unknown ex ante by the student when she decides to enter college. Thus, the propensity $\widetilde{k}^r_j$ is defined as

$$\widetilde{k}^r_j = \alpha^r_2 + X'_{2,j}\beta_2 + \eta, \tag{2}$$

where α^r_2 and β_2 are unknown parameters to be estimated. In this expression, $X_{2,j}$ is a vector of exogenous regressors, including individual characteristics but also covariates that are specific to the major j. Namely, we allow the major and university-specific proportion of college students enrolled in the first two years of college to affect the length of studies.[11] In the absence of variables plausibly affecting the choice of major but not the length of studies, we choose not to include major-specific dummies in $X_{2,j}$ since the related coefficients would be identified only through nonlinearities.[12]

C. Stage 3: Labor Market Earnings

Having reached the educational level k^*_j in major j^*, the student then enters the labor market. We assume that the labor market is an absorbing state: individuals do not resume studies after entering the labor force. When making her post-secondary schooling decision in the first stage, the individual is assumed to anticipate the impact of the major and the length of the studies on her future labor market earnings. In order to take both employment and nonemployment spells into account, we refer to average earnings as the sum of wages weighted by employment spell durations and unemployment benefits weighted by unemployment spell durations.[13] Hence, the logarithm of the average monthly earnings received over a period of length T_{obs} (in months) by a worker with education (j, k) and of type r is given by

$$\overline{\ln w}^r_{jk} = \ln \frac{\sum_{s=1}^{N_e} w_{s,jk} l^e_s + \sum_{s'=1}^{N_u} b_{s',jk} l^u_{s'}}{T_{obs}} \tag{3}$$

with

$$T_{obs} = \sum_{s=1}^{N_e} l^e_s + \sum_{s'=1}^{N_u} l^u_{s'},$$

where N_e (respectively, N_u) is the number of observed employment (unemployment) spells in the individual labor market history, $w_{s,jk}$ is the monthly wage in the sth employment spell, $b_{s',jk}$ is the monthly unemployment benefit in the s'th unemployment spell, l^e_s (respectively, l^u_s) are durations of the sth employment (respectively, unemployment) spell, and T_{obs} is the total length of the observed labor market history of the individual.[14] We define

[10] The term $T - K$ allows capturing the opportunity cost of schooling. Note that we do not account for tuition fees since they were very low in France over the period of interest.

[11] This variable is calculated using information coming from the SISE database provided by the French Ministry of Education.

[12] In our framework, the length of studies is not the number of years spent effectively in postsecondary education, but the terminal level of education that the student reaches, whatever the time spent in college. Note also that we do not account for the selection of applicants made by the university administration at college entry; this last assumption is consistent with the functioning of the French university system.

[13] Unemployment benefits are assumed to be equal to a constant times the former wage received when employed. This constant is taken equal to 0.7, as often done in the literature.

[14] Up to the constant $\ln(T_{\text{obs}})$, our specification based on the log of average earnings is similar to the one used by Arcidiacono (2005). We implicitly normalize the discount factor to be equal to 1.

$$V^r_{e(j,k)} = \overline{\ln w}^r_{jk}. \tag{4}$$

Thereafter, we focus on this aggregate notion of labor market earnings without modeling separately wages and individual probabilities of employment. This appears to be consistent with the students' behavior when they make their postsecondary schooling decisions: most individuals anticipate future labor market conditions as a whole, without separately taking into account the effects of their educational choices on wages and employment probabilities.

Labor market earnings depend on the postsecondary educational field and level, namely, on the pair (j^*, k^*_j). Note that our framework accounts for the earnings gaps not only across schooling levels (within a given major) but also across majors (for a given educational level). Earnings are also supposed to be a function of exogenous and predetermined individual characteristics. For a student of type r, the log average earnings equation is assumed to be given by

$$\overline{\ln w}^r_{jk} = \alpha^r_3 + X'_{3(j,k)}\beta_3 + \varepsilon, \tag{5}$$

where $X_{3(j,k)}$ is a vector of observed characteristics that may affect labor market earnings, including postsecondary education; α^r_3 represents the type-specific intercept; and ε denotes an independent random factor that affects the individual's earnings. This error term is unobserved by both the econometrician and ex ante by the agent.

III. Model Specification

Let us recall that the type-specific intercepts are mass points of a discrete distribution with probabilities $(\Pi_1, \ldots, \Pi_R)$ verifying $\sum_{r=1}^R \Pi_r = 1$, and that the residuals of the three stages are stochastically independent of these heterogeneity terms.[15]

A. Stochastic Assumptions

Residuals are supposed to be normally distributed. We assume that the random vector $(u_1, \ldots, u_M)$ affecting the choice of major and the residuals η and ε entering the two other equations are independently distributed.[16] The whole vector of residuals is assumed to be distributed as[17]

$$\begin{pmatrix} \eta \\ u_2 - u_1 \\ u_3 - u_1 \\ \cdots \\ u_M - u_1 \\ \varepsilon \end{pmatrix} \sim \mathcal{N}(0, \Sigma),$$

where Σ is the $(M+1) \times (M+1)$ covariance matrix of the residuals, with $\Sigma[1,1] = 1$ and $\Sigma[2,2] = 1$ for identifiability reasons.

B. The Likelihood Function

Under our stochastic assumptions, the contribution to the likelihood function of an individual of type r who chooses major j^*, reaches the educational level k^*_j, and gets the average labor market log-earnings $\overline{\ln w}^r_{jk}$ is

$$l\big(j^*, k^*_{j^*}, \overline{\ln w}^r_{jk} | r\big) = \Pr\left[\bigcap_{j' \neq j^*} (u_{j'} - u_{j^*} \leq f_r(j^*) - f_r(j'))\right] \times \Pr[s_{k^*_{j^*}} - \widetilde{h}_r < \eta \leq s_{k^*_{j^*}+1} - \widetilde{h}_r] \times g(\varepsilon), \tag{6}$$

where

$$\widetilde{h}_r = \alpha^r_2 + X'_{2,j}\beta_2$$

$$f_r(j) = \alpha^r_{(1,j)} + X'_1\beta^j_1 + \alpha \sum_{k=0}^{L} \left(X'_{3(j,k)}\beta_3 + \alpha^r_3\right) \times [\Phi(s_{k+1} - \widetilde{h}_r) - \Phi(s_k - \widetilde{h}_r)] \times (T - k)$$

$$g(\varepsilon) = \frac{1}{\sqrt{\Sigma[M+1, M+1]}} \times \varphi\left(\frac{\varepsilon}{\sqrt{\Sigma[M+1, M+1]}}\right)$$

with

$$\varepsilon = \overline{\ln w}^r_{j,k} - \alpha^r_3 - X'_{3(j,k)}\beta_3$$

and

$$\Pr[s_k - \widetilde{h}_r < \eta \leq s_{k+1} - \widetilde{h}_r] = \Phi(s_{k+1} - \widetilde{h}_r) - \Phi(s_k - \widetilde{h}_r),$$

φ and Φ being, respectively, the density and cumulative distribution functions of the standard normal distribution $\mathcal{N}(0,1)$.[18] Note that the first stage of the econometric model corresponds to the estimation of a multinomial probit model. Within this framework, the choice probabilities $\Pr(j|r)$ do not have a closed-form expression. As indicated in the next section, estimations are based on $J = 3$ aggregated majors. Thus, in stage 1, each choice probability is expressed as a double integral, which can be evaluated using usual integration procedures (such as quadrature methods) without the need to rely on a GHK probit simulator.

[15] Some covariates introduced in the equations may not be independent of the individual's type. It applies especially to the high school graduation track, which may be related to unobserved preferences for each major. When the type probabilities are assumed to depend on the high school graduation track, the parameter estimates are not significantly modified (these results are not reported here, but are available from us on request). The estimates reported in the paper are obtained without conditioning on the high school track.

[16] Correlated unobserved heterogeneity across equations is captured by the type-specific random intercepts $(\alpha^r_{(1,j)})_{j=1,\ldots,M}$, α^r_2, and α^r_3.

[17] Only differences in utility levels matter in random utility models.

[18] We refer readers to Beffy et al. (2009) for details on the estimation procedure.

Unconditional on the type, the individual contribution to the likelihood function follows a finite mixture distribution:

$$l\big(j^*, k^*_{j^*}, \overline{\ln w}_{j^*,k^*_{j^*}}\big) = \sum_{r=1}^{R} \Pi_r l\big(j^*, k^*_{j^*}, \overline{\ln w}^{r}_{j^*,k^*_{j^*}} | r\big), \quad (7)$$

where $l(j^*, k^*_{j^*}, \overline{\ln w}^{r}_{j^*,k^*_{j^*}} | r)$ denotes the contribution to the likelihood given the student's type r.

C. Estimation

In order to present our estimation strategy, we introduce some further notations: θ_F denotes the parameters of the major choice equations, θ_L those of the equation for the length of studies, and finally θ_W those of the earnings equation. These vectors do not include the type-specific intercepts denoted by $(\alpha^W_r)_r, (\alpha^L_r)_r, (\alpha^F_r)_r$. As is usual for a finite mixture of Gaussian distributions, we use the expectation-maximization (EM) algorithm (see, for instance, Dempster, Laird, & Rubin, 1977) to estimate our model.

Due to the partial separability of the conditional completed log-likelihood function (Arcidiacono & Jones, 2003), we get three sequential optimization problems since residuals are assumed to be independent across the three equations. Henceforth, denoting by $\pi^{(n)}_{i,r}$ the posterior probability, computed at the nth iteration of the EM algorithm, for the individual i to be of type r,[19]

$$\begin{aligned}
&\sum_{i=1}^{N}\sum_{r=1}^{R} \pi^{(n)}_{i,r} \ln l(j_i, k_i, \ln w_i | \text{Type}_i \\
&\quad = r, (\Pi_r)_r, (\alpha_r)_r, \theta_F, \theta_L, \theta_W) \\
&= \sum_{i=1}^{N}\sum_{r=1}^{R} \pi^{(n)}_{i,r} \ln l(\ln w_i | \text{Type}_i \\
&\quad = r, k_i, j_i, (\Pi_r)_r, \big(\alpha^W_r\big)_r, \theta_W) \\
&+ \sum_{i=1}^{N}\sum_{r=1}^{R} \pi^{(n)}_{i,r} \ln l(k_i | \text{Type}_i \\
&\quad = r, j_i, (\Pi_r)_r, \big(\alpha^W_r\big)_r, \big(\alpha^L_r\big)_r, \theta_W, \theta_L) \\
&+ \sum_{i=1}^{N}\sum_{r=1}^{R} \pi^{(n)}_{i,r} \ln l(j_i | \text{Type}_i \\
&\quad = r, (\Pi_r)_r, \big(\alpha^W_r\big)_r, \big(\alpha^L_r\big)_r, \big(\alpha^F_r\big)_r, \theta_W, \theta_L, \theta_F).
\end{aligned}$$

Thus, we first maximize the log-earnings and length of studies contributions. Then, given the previous estimates, we maximize the last term, which relates to the choice of major. Although this procedure does not yield full information maximum likelihood estimates, Arcidiacono and Jones (2003) show that this method produces consistent estimates of the parameters, with large computational savings.[20]

IV. Data

The model presented above is estimated using French data coming from the Génération 92 and Génération 98 surveys, which are collected by the Centre d'Etudes et de Recherches sur l'Emploi et les Qualifications (CEREQ, Marseille).[21] The Génération 92 survey consists of a large sample of 26,359 individuals who left the French educational system in 1992 and were interviewed five years later, in 1997. In the original sample, education levels range from the lowest to the highest one, respectively referred to as Level VI and Level I in the French nomenclature. This database contains information on both educational and labor market histories (over the first five years following the exit from schooling). Furthermore, the survey provides a set of individual covariates that are used as controls in our estimation procedure such as gender, place of birth, nationality, parental profession, and residence when entering the labor market. Most of the covariates observed in the Génération 92 survey are also provided by the Génération 98 survey, a sample of 22,021 individuals who left the French educational system six years later, in 1998, and were interviewed in 2003.[22] In this paper, we use the pooled data set, which contains information on 48,380 individuals entering the labor market in either 1992 or 1998.

Our subsample is made up of respondents who at least passed the national high school final examination. It is then restricted to 27,389 individuals. Furthermore, within this sample, we restrict our analysis to the individuals who attended university,[23] except medicine faculties and IUT (Institut Universitaire de Technologie, two-year vocational colleges). This sample restriction was made in order to keep a homogeneous set of postsecondary tracks in terms of selection at entry and possible length of studies. Missing covariate values leave us with a sample of 7,346 individuals.[24]

Postsecondary studies are aggregated into three broad majors: sciences, humanities and social sciences (including art studies), and law, economics, and management. Tables 1

[19] For each type $r = 1, \ldots, R$, and each individual i, at each iteration n of the EM algorithm, the posterior probability is equal to $\pi^{(n)}_{i,r} = \frac{\pi^{(n)}_r l(j_i,k_i,\ln w_i|r)}{\sum_{r=1}^{R}\pi^{(n)}_r l(j_i,k_i,\ln w_i|r))}$, where $(\pi^{(n)}_r)_{r=1,\ldots,R}$ denote the mixing proportions at iteration n.

[20] In order to get standard error estimates, we rely on a parametric bootstrap procedure (with fifty replications), instead of a nonparametric one, since the latter method is unstable when applied to the EM algorithm.

[21] These data have been previously used by Brodaty, Gary-Bobo, and Prieto (2006), who estimate a structural model of educational investments accounting for attitudes toward risk.

[22] Although a longer observation window is available for each data set, the average log earnings are computed using only the observations from 1992 to 1995 for Génération 1992 (resp. 1998 to 2001 for Génération 1998). In particular, restricting to a four-year window allows limiting the number of individuals who have to be dropped because of missing earnings values, in addition to the fact that it permits us to work with two periods of contrasted macroeconomic conditions, which helps us to identify the elasticity of major choice to expected earnings.

[23] We exclude other postsecondary tracks with selective admissions, such as Classes Préparatoires aux Grandes Ecoles and Brevets de Technicien Supérieur.

[24] In order to prevent our estimates to be driven by outliers, we also drop individuals with average log-earnings below the 2.5 percentile (respectively, above the 97.5 percentile) of the log earnings distribution.

TABLE 1.—DESCRIPTIVE STATISTICS: MAJORS AND LEVELS OF POSTSECONDARY EDUCATION

	Total	Dropout	Associate Degree	Bachelor's Degree	First Year of Master's Degree	Master's Degree and Ph.D.
Sciences	2,106	337	210	324	403	832
Humanities and social sciences	2,761	825	311	721	420	484
Law, economics, and management	2,479	600	211	355	663	650
Total	7,346	1,762	732	1,400	1,486	1,966
Génération 1992						
Sciences	1,094	103	99	207	230	455
Humanities and social sciences	1,174	191	116	358	223	286
Law, economics, and management	1,168	224	66	177	328	373
Total	3,436	518	281	742	781	1,114
Génération 1998						
Sciences	1,012	234	111	117	173	377
Humanities and social sciences	1,587	634	195	363	197	198
Law, economics, and management	1,311	376	145	178	335	277
Total	3,910	1,244	451	658	705	852

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE 2.—DESCRIPTIVE STATISTICS: COVARIATES

	Number	Percentage
Year of entry into the labor market		
1992	3,436	46.77
1998	3,910	53.23
Gender		
Male	3,197	43.52
Female	4,149	56.48
Born abroad		
No	7,164	97.52
Yes	182	2.48
Age in sixth grade		
≤10	858	11.68
11	6,109	83.16
≥12	379	5.16
High school graduation track		
Humanities	1,712	23.31
Economics and social sciences	1,733	23.59
Sciences	2,523	34.35
Vocational or technical	1,378	18.76
Father's profession (at the survey date)		
Farmer or tradesman	1,131	15.40
Executive	2,213	30.13
Intermediate occupation	898	12.22
White collar	1,468	19.98
Blue collar	1,237	16.84
Out of the labor force	399	5.43
Mother's profession (at the survey date)		
Farmer or tradesman	527	7.17
Executive	1,226	16.69
Intermediate occupation	508	6.92
White collar	3,269	44.50
Blue collar	508	6.92
Out of the labor force	1,308	17.81

Source: Surveys Générations 1992 and 1998 (CEREQ).

and 2 provide basic descriptive statistics for the selected subsample.[25]

We first focus on the choice of the major, which is related to gender, age at the entry into junior high school (sixth grade),[26] and parental profession. Noteworthy, male students are more likely to choose majors in sciences (39.40% among male versus 16.42% among females) while female students are more likely to have majors in humanities and social sciences (29.32% among males versus 48.93% among females). The student's age in sixth grade and the chosen major are highly correlated: individuals who were above the "normal" age in sixth grade are less likely to choose a major in science and more likely to choose one in law, economics, and management.

Parental characteristics also seem to play an important role on the choice of the major. Individuals from a higher parental background have a higher probability of studying sciences. For instance, individuals whose father is a blue-collar worker are more likely to study humanities and social sciences and less likely to study sciences.[27] There is also a strong correlation between the chosen major and the length of studies. Only one-quarter of individuals with a master's degree or a Ph.D. completed their degree in humanities and social sciences. In contrast, half of the college dropouts majored in the humanities.

Finally, as expected, earnings are positively correlated with the educational level (see table 3): individuals with a master's degree or a Ph.D. earn on average 1.7 times more than dropouts. There are significant differences in average earnings associated with the different majors: sciences ranks first, followed by law, economics, and management, and finally humanities and social sciences. The discrepancy between majors is greater in 1998 than in 1992. This feature will be used to identify the effect of expected earnings on major choice. Unlike humanities and social sciences, sciences, as well as law, economics, and management, benefited from the macroeconomic expansion that occurred in the late 1990s. Following a referee's suggestion, we have also computed the average earnings growth over the first five years of labor market history. These computations show that (a) on average, labor market earnings increase by about 30% over this five-year period, which corresponds to a 6% annual earnings

[25] Additional descriptives are reported in Beffy et al. (2009).

[26] This variable can be seen as a proxy for individual schooling ability. Indeed, in France, most of its variation stems from grade repetition.

[27] Mother's profession is associated with the major in a similar way.

TABLE 3.—AVERAGE MONTHLY EARNINGS (1992 FRENCH FRANCS) ACCORDING TO LENGTH AND FIELD OF STUDIES

	Mean	Dropout	Associate Degree	Bachelor's Degree	First Year of Master's Degree	Master's Degree and Ph.D.
Sciences	7,277	5,072	6,056	6,393	6,873	9,019
Humanities and social sciences	5,942	4,912	5,695	5,893	6,245	7,669
Law, economics, and management	6,666	4,847	6,336	6,572	6,971	8,193
Mean	6,569	4,921	5,983	6,181	6,739	8,414
Génération 1992						
Sciences	6,833	4,279	5,828	6,121	6,684	8,029
Humanities and social sciences	6,088	4,160	6,123	5,946	6,454	7,255
Law, economics, and management	6,318	4,209	6,283	6,311	6,536	7,403
Mean	6,404	4,205	6,057	6,082	6,556	7,621
Génération 1998						
Sciences	7,758	5,422	6,259	6,876	7,124	10,214
Humanities and social sciences	5,835	5,139	5,441	5,840	6,009	8,267
Law, economics, and management	6,976	5,227	6,360	6,831	7,397	9,257
Mean	6,715	5,219	5,938	6,292	6,942	9,450

Source: Surveys Générations 1992 and 1998 (CEREQ).

growth, and (b) we cannot reject (at the 10% level) the hypothesis of equality of earnings growth rates across majors. Thus, there are no significant differences in earnings profiles across majors within the first five years of labor market history. This suggests that in this context, relying on average earnings over the period of observation to compute the relative returns to each major is indeed reasonable.

V. Identification Strategy

For identifiability reasons, we impose the usual restrictions on the type-specific heterogeneity terms of equations (1) and (2). Namely, in the multinomial probit model for the choice of major, we set $\alpha^r_{(1,1)} = 0, \forall r \in \{1, \ldots, R\}$, and in the ordered probit model corresponding to the second equation, we set $\alpha^1_2 = 0$.

In order to identify our model, and in particular the effect of expected earnings on the probability of choosing each major, without relying on only distributional and functional forms assumptions, we take advantage of the variations in the relative wage returns induced by the business cycle. In other terms, we take into account the fact that these relative returns depend on the year of entry into the labor market.[28] Descriptive statistics reported in table 3 show that the relative returns associated with the majors change significantly between 1992 and 1998; these years correspond respectively a downturn and an expansion in the French business cycle.[29] After controlling for the change in the distribution of educational levels as well as for inflation, we find a relative increase of 13.5% (respectively, 10.4%) in the average earnings associated with majors in sciences (respectively, in law, economics, and management) between 1992 and 1998, while the average earnings associated with majors in humanities and social sciences decreased by 4.2% over the same period.[30] Besides, it seems reasonable to assume that the date of entry into the labor market has no direct influence on the choice of the major—in other words that, other observable things being equal, preferences for the majors were stable during this period.[31] In order to identify the elasticity of the choice of the major with respect to expected earnings, we exploit the fact that the returns to the different majors are unequally affected by the business cycle.[32] Hence, we introduce into the earnings equation interaction terms between the chosen major and an entry year dummy. This dummy variable is equal to 0 if the individual enters the labor market in 1992 and to unity otherwise (namely, if she enters the labor market six years later in 1998). Its interaction with the chosen major is assumed to affect only the earnings and not the two other outcomes. This exclusion restriction (over)identifies the parameter α associated with the expected returns in the choice equation. Moreover, the covariates indicating the father's and mother's professions (respectively, in 1992 and 1998), the age of the student in the sixth grade, and the high school major are included in the list of regressors affecting both the choice of the major and the determination of the length of studies, but they are excluded from the earnings equation. Similarly to Arcidiacono (2005; see section IV), these exclusion restrictions, in

[28] Berger (1988) also relies on exogenous variations in the wage returns to each major according to the date of entry into the labor market in order to identify the effect of expected earnings on choice of college major. Unlike ours, his framework does not take into account the determination of the length of studies. Besides, his results rely on the independence from irrelevant alternative assumption for the choice of the major, which is unlikely to hold in such a context.

[29] See figure A1 in appendix A.

[30] These relative variations between 1992 and 1998 are obtained by computing for each major the average of mean monthly earnings at each level of education, weighted by the proportion of students having reached this level.

[31] In particular, no reform concerning postsecondary education was implemented in France between 1992 and 1998. The progressive application of the Bologna process to the French postsecondary educational system began in 1999. Thus, it should not affect the decisions of the individuals in our sample who had already entered the labor market at that time. The purpose of the Bologna process was to create a European higher education area by making academic degree standards more comparable and compatible throughout Europe.

[32] On a related ground, in a paper examining the career effects of graduating in a recession, Oreopoulos, von Wachter, and Heisz (2008) show that Canadian college graduates are unequally affected by the recession according to their major.

addition to the assumed functional forms, help to identify the unobserved heterogeneity types. Finally, we assume that the proportion of two-year college students who have the same major in the same university as the individual may influence the length of her studies but not her choice of major or her earnings.

VI. Results

A. *Parameter Estimates*

Table B1 (see appendix B) reports the parameter estimates of the equations for the major choice. The results are obtained under the assumption that there are $R = 3$ types of individuals within the sample.[33]

Students whose mother has a white-collar occupation choose less frequently majors in humanities and social sciences, compared to sciences, than students whose mother is an executive. Noteworthy, students whose mother is a farmer, a tradeswoman, or a white-collar worker or whose mother has an intermediate profession, also less frequently choose majors in law, economics, and management compared to sciences. In all other cases, parental (and in particular father's) profession has generally no effect on the major choice.

The nationality of the student's parents has a significant and quantitatively large impact on the choice of a major in law, economics, and management, as well as in humanities and social sciences, compared to sciences. Students born abroad are significantly less likely to study law, economics, or management. Noteworthy, female students are significantly less likely to study sciences. As expected, students who obtained a baccalauréat in sciences are significantly more likely to choose a major in sciences too. Students who were older than expected (12 years old or above) at entry into junior high school less frequently choose a major in sciences. It is noteworthy that the expected returns in a given major have a statistically significant but rather small effect on the choice of the major (see the value for the estimate of the parameter α in table B1).

Most covariates have a significant impact on the length of postsecondary studies (see table B2). For instance, students whose parents are white-collar or blue-collar workers tend to get a lower level of education. Students whose both parents are French reach a generally higher level of postsecondary education. Students who were younger than expected (10 years old or below) at entry into junior high school reach a higher level of education. Those who obtained their baccalauréat in sciences are also more likely to reach a higher level of postsecondary education.

When the proportion of two-year college students who have the same major in the same university as the individual increases, which implies that the proportion of students preparing a bachelor's, a master's degrees, or a Ph.D. is lower in this major and in this university, the individual probability of reaching a high level of education in this major is lower. This may result from the selection imposed by the university after the end of the associate level and from the effects of peers, on interpretation set forth by Arcidiacono (2004, 2005). Finally, women are less likely to pursue long studies.

Table B3 gives the parameter estimates of the (log-)earnings equation. On average, earnings are lower for females and higher in Paris region (Ile-de-France). Mean earnings increase with the length of studies. However, this increase is lower above the bachelor's degree in the humanities and social sciences. Noteworthy, the marginal returns to each additional year of postsecondary education are also lower, up to the master's and Ph.D. levels, for those entering the labor market in 1998 than for those leaving university six years before. Consistent with the fact that individuals entering the French labor force in 1998 benefited from favorable economic conditions (as compared to those entering the labor market in 1992), earnings are substantially higher for those leaving the university in 1998. Finally, while controlling for selection on observables and unobservables renders statistically insignificant the relative returns to the majors for the individuals leaving university in 1992, those entering the labor market in 1998 after graduating in humanities and social sciences experience negative relative returns compared to other majors.

Tables B4 and B5 report the parameter estimates of the distribution of unobserved individual heterogeneity terms. The first group of individuals represents 38% of the population of students. Individuals in this group are characterized by the lowest unobserved type-specific preference for studying sciences, as well as the highest type-specific earnings intercept α_3. The second group represents approximatively 34% of the population of students. Individuals in this group are characterized by the lowest type-specific preference $\alpha_{(1.3)}$ for studies in law, economics, and management. They also have the lowest type-specific propensity (or ability) α_2 to undertake long postsecondary studies. Finally, the third group represents about 28% of the population; it is characterized by both the lowest type-specific earnings intercept term α_3 and the highest propensity to pursue lengthy postsecondary studies.

Table B6 reports the estimated proportions of students in each major, at each level of postsecondary education, according to each type, while table B7 reports the means and standard deviations of log-earning distributions by type. These tables are obtained by attributing to each observed individual the type that maximizes her posterior type probability. It follows from table B6 that individuals in the first group never choose a major in sciences, while around 60% of them study law, economics, and management. In contrast, individuals in the second group never choose to study law, economics,

[33] The number R of components is chosen on the basis of the BIC penalized likelihood criterion. Models with $R = 2$ and $R = 1$ heterogeneity types yield substantially larger BIC values. Models with more than three types were also estimated, yielding almost degenerate type distributions. Thus, we chose $R = 3$ types for our final set of estimations.

TABLE 4.—SIMULATION OF A 10% VARIATION IN EXPECTED EARNINGS ASSOCIATED WITH MAJORS IN SCIENCES (PERCENTAGES)

	Observed Proportion	Predicted Proportion	$\hat{\Delta}_p$	$\hat{\sigma}_{\hat{\Delta}_p}$
10% increase				
Sciences	28.67	27.97	0.251	0.019
Humanities and social sciences	37.59	41.17	−0.189	0.013
Law, economics, and management	33.75	30.86	−0.062	0.009
10% decrease				
Sciences	28.67	27.97	−0.276	0.021
Humanities and social sciences	37.59	41.17	0.209	0.014
Law, economics, and management	33.75	30.86	0.068	0.009

Source: Surveys Générations 1992 and 1998 (CEREQ).
$\hat{\Delta}_p$ (respectively, $\hat{\sigma}_{\hat{\Delta}_p}$) denotes the estimated variation (respectively, its standard error) in the predicted probability due to a 10% variation in the expected earnings.

TABLE 5.—SIMULATION OF A 10% VARIATION IN EXPECTED EARNINGS ASSOCIATED WITH MAJORS IN HUMANITIES AND SOCIAL SCIENCES (PERCENTAGES)

	Observed Proportion	Predicted Proportion	$\hat{\Delta}_p$	$\hat{\sigma}_{\hat{\Delta}_p}$
10% increase				
Sciences	28.67	27.97	−0.189	0.013
Humanities and social sciences	37.59	41.17	0.526	0.048
Law, economics, and management	33.75	30.86	−0.336	0.038
10% decrease				
Sciences	28.67	27.97	0.209	0.014
Humanities and social sciences	37.59	41.17	−0.580	0.053
Law, economics, and management	33.75	30.86	0.371	0.042

Source: Surveys Générations 1992 and 1998 (CEREQ).
$\hat{\Delta}_p$ (respectively, $\hat{\sigma}_{\hat{\Delta}_p}$) denotes the estimated variation (respectively, its standard error) in the predicted probability due to a 10% variation in the expected earnings.

and management, and about 45% of them choose a major in sciences. Finally, individuals in the third group are much less likely to study humanities and social sciences, since slightly more than 17% of them are enrolled in this major. This pattern suggests a substantial heterogeneity across types in the preferences for each major. Table B6 also shows that individuals differ a great deal in the length of their studies, with students from the second and third groups making opposite decisions in terms of the level of postsecondary education. Consistent with the results reported in table B5, table B7 suggests that the three types are much less heterogeneous in their earnings than they are in their preferences and schooling abilities. Finally, table B8 reports the R^2 obtained from simple OLS regressions of predicted individual schooling decisions (major and length of studies) on the type-specific endowments and the full set of regressors used when estimating our model, respectively. Overall, the results show that the relative importance of regressors is similar to the one of type-specific heterogeneity, albeit slightly higher. This result suggests that unobserved characteristics related in particular to taste for schooling, but also to unobserved schooling ability, are a key determinant of postsecondary schooling decisions.[34]

The model fit is quite good. Table 4 shows that the model slightly overestimates (respectively underestimates) the proportion of students in humanities and social sciences (respectively in law, economics, and management).

B. *Simulation Exercises*

To get a more precise view of the effect of expected earnings on the choice of the major, we run simulation exercises that consider a 10% increase or decrease in the expected earnings associated with a given major (see tables 4 to 6).[35]

Overall, the impacts are quantitatively small even though they are statistically significant. The lowest impacts concern the majors in sciences. A 10% increase in the expected earnings associated with majors in sciences leads to an increase of 0.25 percentage points in the proportion of students in this major. This increase is mainly compensated by a decrease of 0.19 percentage points in the proportion of students in humanities and social sciences (see table 4). A 10% decrease in the expected earnings associated with majors in sciences results in almost symmetric variations in allocations across majors.

Impacts resulting from a 10% increase or decrease in the expected earnings associated with majors in humanities and social sciences are substantially higher although still quantitatively small (see table 5). For instance, a 10% increase in the expected earnings associated with these majors results in an increase of about 0.53 percentage points in the proportion of students in these majors, this increase being mainly compensated by a decrease of about 0.34 percentage points in the proportion of students in law, economics, and management and, to a lesser extent, by a 0.19 points decrease in the proportion of students in sciences.

Finally, a 10% increase in the expected earnings associated with a postsecondary education in law, economics, and management majors results in an increase of 0.4 percentage points in the proportion of students in these majors, this increase being mainly compensated for by a decrease of 0.34 percentage points in the proportion of students in humanities and social sciences (see table 6). The effects are still symmetric for a 10% decrease in the expected earnings associated with this major.

These simulation exercises allow computing the sample earnings elasticities of major choice, which present the advantage of being easily interpreted. Namely, simulating a 10% increase in the expected earnings for each major yields

[34] A referee has rightly pointed out that it would be interesting to allow heterogeneity across types in the returns to major. Nevertheless, we think that to be consistent with such an approach, we should allow each slope parameter to vary across types, especially those associated with the field and also with the length of study, as well as the interactions among them. This would induce a substantial computational burden, probably leading to a loss of precision. Besides, the interpretation of the results would be more difficult. This extension is left for further research.

[35] Simulating both types of variation enables us to see whether the impacts on allocations across majors are symmetric.

TABLE 6.—SIMULATION OF A 10% VARIATION IN EXPECTED EARNINGS ASSOCIATED WITH MAJORS IN LAW, ECONOMICS, AND MANAGEMENT (PERCENTAGES)

	Observed Proportion	Predicted Proportion	$\hat{\Delta}_p$	$\hat{\sigma}_{\hat{\Delta}_p}$
10% increase				
Sciences	28.67	27.97	−0.062	0.009
Humanities and social sciences	37.59	41.17	−0.337	0.038
Law, economics, and management	33.75	30.86	0.399	0.042
10% decrease				
Sciences	28.67	27.97	0.068	0.009
Humanities and social sciences	37.59	41.17	0.371	0.042
Law, economics, and management	33.75	30.86	−0.439	0.046

Source: Surveys Générations 1992 and 1998 (CEREQ).
$\hat{\Delta}_p$ (respectively, $\hat{\sigma}_{\hat{\Delta}_p}$) denotes the estimated variation (respectively, its standard error) in the predicted probability due to a 10% variation in the expected earnings.

very low elasticities, respectively equal to 0.09 for the sciences, 0.14 for the humanities and social sciences, and 0.12 for law, economics, and management.[36]

The results were obtained relying on the econometric framework detailed in section III, which in particular does not account for log-earnings heteroskedasticity. In order to address the fact that major choices may also be driven by major and level-specific earnings dispersions, we also run additional estimations based on an extension of our model accounting for heteroskedasticity and risk aversion. Namely, we impose an exponential parametric form of heteroskedasticity, allowing the variance of log-earnings to depend on both the major and the level of education. More precisely, we assume that individuals value the major and level-specific earnings through a CRRA von Neumann–Morgenstern utility function, with a risk-aversion parameter taken equal to $\rho = 1.1$. This means that the expectations appearing in the expression of the terms v^r_{1j}, for $j = 1, \ldots, M$ become

$$\frac{1}{1-\rho} e^{(1-\rho)\mu^r_{j,k} + \frac{(1-\rho)^2 \sigma^2_{j,k}}{2}},$$

where $\mu^r_{j,k}$ and $\sigma^2_{j,k}$ denote the mean and variance of log-earnings, respectively. This alternative specification yields fairly similar earnings elasticities of major choice. The effects of expected earnings on major choice are still significant but quantitatively small, with point estimates of the same magnitude (see Beffy, Fougère, & Maurel, 2009, appendix C, for the detailed simulation results).

Note also that in our setting, the length of studies is modeled within a reduced-form framework. In particular, we do not explicitly account for the fact that within each major, the length of studies may be sensitive to expected earnings. We have checked for the robustness of our results by estimating an alternative specification in which the length of studies depends on the average expected returns to each additional year of postsecondary education, the first and last stages of the model being left unchanged.[37] More precisely, the latent variable generating the length of studies in a given major j is now written as (with a slight abuse of notation here since we do not include the dummy for the year of entry into the labor market in the set of regressors $X_{2,j}$, which helps to identify the parameter γ_2)

$$\begin{aligned} \widetilde{k}^r_j = \alpha^r_2 + X'_{2,j}\beta_2 \\ + \gamma_2 \Bigg(\frac{1}{L} \sum_{k=0}^{L-1} \big[E\big(V^r_{e(j,k+1)} \mid r, j, k+1\big) \\ - E\big(V^r_{e(j,k)} \mid r, j, k\big) \big] \Bigg) \\ + \eta. \end{aligned}$$

Results of simulations obtained with this alternative specification suggest that the elasticity of the length of studies to expected earnings is substantial, with the proportion of individuals with a master's degree or a Ph.D. increasing by 8% after a 10% increase in the expected earnings associated with this level of education. Nonetheless, although the length of studies appears to be quite sensitive to expected earnings, additional simulation exercises show that the elasticity of major choice with respect to expected earnings is robust to this alternative specification.[38]

VII. Conclusion

This paper considers the determinants of the choice of the postsecondary major, with a focus on the elasticity of this choice to expected earnings. We specify and estimate a three-stage schooling decision model. First, students choose their field of study according to the expected earnings returns, as well as their preferences and aptitude for each major. Then they continue studying in this major until they reach a given level of education, which is ex ante unknown to the students, before finally entering the labor market. We extend the literature on the effects of expected earnings on schooling choices by endogenizing the length of study. We also rely on an original identification strategy for recovering the key elasticity parameters, which makes the most of the variation across the French business cycle of the relative returns to each major.

[36] As a referee pointed out, our data concerning the first four years of labor market history do not allow us to identify whether these measured elasticities are with respect to a permanent change or a transitory change in major-specific earnings. Nevertheless, evidence from data extracted from the French Labor Force Surveys (1995–2008) suggests that relative wage differentials (in particular between sciences and humanities and social sciences) according to the year of entry into the labor market (1992 or 1998) are fairly persistent, at least over the first ten years after graduation. Besides, one can show that given the functional forms assumed in the model, even if the observed changes were transitory, a permanent increase in earnings would lead to a similar response in terms of major choice. These elements are available from us on request.

[37] An adequate strategy to deal with this issue would be to estimate a fully structural dynamic model treating the decision to stop schooling as an optimal stopping rule. Nevertheless, this specification would require a different estimation procedure since we would not be able to use the sequential EM algorithm anymore.

[38] Detailed simulation results are not reported in the paper for the sake of brevity. These results are available from us on request.

Our results suggest that the elasticity of major choices to expected earnings is very low. In general, the impact of expected earnings on these choices is quantitatively small while statistically significant. The lowest impact concerns majors in sciences. The impact of the expected earnings associated with majors in humanities and social sciences is substantially higher although still quantitatively small. Increases and decreases in the expected earnings result in almost symmetric variations in allocations across majors. Our main results are robust to an alternative specification accounting for log-earnings heteroskedasticity, across majors and levels of education, and for risk aversion. Explicitly allowing the length of studies to depend on the expected returns to each additional year of postsecondary education also yields similar elasticities of major choices to expected earnings.

Overall it appears that the choice of a major, which is made when entering college, is mainly driven by the consumption value of schooling, which is related to both schooling preferences and abilities rather than by its investment value. Thus, our paper provides strong evidence, in line with Carneiro, Hansen, and Heckman (2003), that nonpecuniary factors are a key determinant of schooling choices.

From a policy point of view, this paper suggests that the solution to the relative decrease in the number of students in science and technology does not lie in financial incentives.[39] Providing incentives, as often advocated, to implement gain and profit-sharing schemes appears to be unlikely to overcome skill shortages. The solution probably lies upstream on the development of preferences and abilities earlier in schooling.

[39] See, for instance, the OECD (2008) policy report on the evolution of student interest in Science and Technology.

REFERENCES

Aghion, Philippe, and Elie Cohen, *Education et croissance* (Paris: La Documentation Française, 2004).

Altonji, Joseph, "The Demand for and Return to Education When Education Outcomes Are Uncertain," *Journal of Labor Economics* 11 (2003), 48–83.

Arcidiacono, Peter, "Ability Sorting and the Returns to College Major," *Journal of Econometrics* 121 (2004), 343–375.

——— "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica* 73 (2005), 1477–1524.

Arcidiacono, Peter, and John B. Jones, "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica* 71 (2003), 933–946.

Beffy, Magali, Denis Fougère, and Arnaud Maurel, "Choosing the Field of Study in Post-Secondary Education: Do Expected Earnings Matter?" IZA discussion paper no. 4127 (April: 2009).

Belzil, Christian, and Jorgen Hansen, "Unobserved Ability and the Return to Schooling," *Econometrica* 70 (2002), 2075–2091.

——— "Earnings Dispersion, Risk Aversion and Education," *Research in Labor Economics* 23 (2004), 335–358.

Berger, Mark C., "Predicted Future Earnings and Choice of College Major," *Industrial and Labor Relations Review* 41 (1988), 418–429.

Boudarbat, Brahim, and Claude Montmarquette, "Choice of Fields of Study of University Canadian Graduates: The Role of Gender and Their Parents' Education," *Education Economics* 17 (2009), 185–213.

Brewer, Dominic J., Eric Eide, and Ronald G. Ehrenberg, "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings," *Journal of Human Resources* 34 (1999), 104–123.

Brodaty, Thomas, Robert J. Gary-Bobo, and Ana Prieto, "Risk Aversion and Human Capital Investment: A Structural Econometric Model," CEPR discussion paper no. 5694 (2006).

Cameron, Stephen V., and James J. Heckman, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy* 106 (1998), 262–333.

——— "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy* 109 (2001), 455–499.

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman, "Estimating Distributions of Counterfactuals with an Application to the Returns to Schooling and Measurement of the Effect of Uncertainty on Schooling Choice," *International Economic Review* 44 (2003), 361–422.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society, Series B* 39 (1977), 1–38.

Eckstein, Zvi, and Kenneth I. Wolpin, "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities and Abilities," *Econometrica* 67 (1999), 1295–1339.

Freeman, Richard B., *The Market for College-Trained Manpower* (Cambridge, MA: Harvard University Press, 1971).

——— "Supply and Salary Adjustments to the Changing Science Manpower Market: Physics, 1948–1973," *American Economic Review* 65 (1975), 27–39.

Heckman, James J., and Burton Singer, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica* 52 (1984), 271–320.

James, Estelle, Nabeel Alsalam, Joseph C. Conaty, and Duc-Le To, "College Quality and Future Earnings: Where Should You Send Your Child to College?" *American Economic Review, Papers and Proceedings* 79 (1989), 247–252.

Keane, Michael P., and Kenneth I. Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy* 105 (1997), 473–522.

——— "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment," *International Economic Review* 42 (2001), 1051–1103.

Lee, Donghoon, "An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice," *International Economic Review* 46 (2005), 1–34.

Loury, Linda D., and David Garman, "College Selectivity and Earnings," *Journal of Labor Economics* 13 (1998), 289–308.

Manski, Charles F., "Adolescent Econometricians: How Do Youths Infer the Returns to Schooling?" in Charles T. Clotfelter and Michael Rothschild (Eds.), *Studies of Supply and Demand in Higher Education* (Chicago: University of Chicago Press, 1993).

OECD, *Evolution of Student Interest in Science and Technology Studies* (Paris: OECD, 2008).

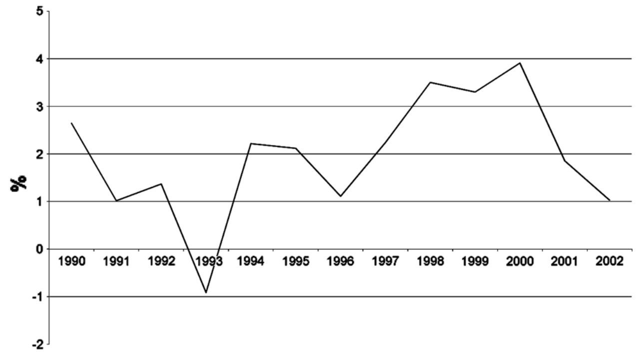
Oreopoulos, Philip, Till von Wachter, and Andrew Heisz, "The Short- and Long-Term Career Effects of Graduating in a Recession: Hysteresis and Heterogeneity in the Market for College Graduates," IZA discussion paper no. 3578 (2008).

Willis, Robert J., and Sherwin Rosen, "Education and Self-Selection," *Journal of Political Economy* 87 (1979), S7–S36.

APPENDIX A

Other Descriptive Statistics

FIGURE A1.—REAL GROWTH RATE OF THE FRENCH GDP, 1990–2002



Source: IMF.

TABLE 7.—PROPORTIONS OF STUDENTS WHO CHANGE MAJOR AFTER ONE YEAR OF COLLEGE (PERCENTAGE)

Major (First Year of College)	LEM	HSS	S
Major (second year of college)			
Law, economics, and management	94.95	1.45	0.69
Humanities and social science	4.89	97.78	3.70
Science	0.16	0.77	95.60

Source: Panel 1989 (DEPP).
Lines sum up to 100%.

TABLE 8.—EXPECTED AND EFFECTIVE LEVELS OF STUDIES (PERCENTAGE)

Effective Level of Studies	Less Than Associate	Associate	Bachelor's	Master's or More
Aspiration (first year of college)				
Less than associate	33.71	12.36	28.09	25.84
Associate	45.00	20.50	17.00	17.50
Bachelor's	32.49	16.40	24.61	26.50
Master's or more	23.06	13.97	25.40	37.57

Source: Panel 1989 (DEPP).
Lines sum to 100%.

APPENDIX B

Parameter Estimates

TABLE B1.—CHOICE OF MAJOR

Covariates	Estimate	Standard Error
Expected earnings (α)	0.019	0.001
Sciences	Ref	Ref
Humanities and Social Sciences		
Father's profession		
Executive	Ref	Ref
Farmer or tradesman	−0.103	0.095
Intermediate occupation	−0.083	0.090
White collar	−0.053	0.068
Blue collar	0.073	0.089
Unknown	0.438	0.129
Mother's profession		
Executive	Ref	Ref
Farmer or tradesman	−0.210	0.109
Intermediate occupation	−0.139	0.101
White collar	−0.134	0.057
Blue collar	−0.175	0.107
Unknown	−0.237	0.082
Born abroad	−0.190	0.123
Woman	0.920	0.051
Both parents are French	−0.303	0.062
Age in sixth grade		
≤10	−0.021	0.072
11	Ref	Ref
≥12	0.391	0.102
High school graduation track		
Sciences	Ref	Ref
Humanities	2.200	0.075
Economics and social sciences	2.287	0.082
Vocational or technical	1.164	0.064
Law, Economics, and Management		
Father's profession		
Executive	Ref	Ref
Farmer or tradesman	−0.030	0.111
Intermediate occupation	−0.094	0.110
White collar	−0.026	0.097
Blue collar	0.004	0.117
Unknown	0.477	0.145
Mother's profession		
Executive	Ref	Ref
Farmer or tradesman	−0.335	0.143
Intermediate occupation	−0.261	0.135
White collar	−0.179	0.073
Blue collar	−0.046	0.162
Unknown	−0.165	0.088
Born abroad	−0.335	0.180
Woman	0.900	0.072
Both parents are French	−0.343	0.084
Age in sixth grade		
≤10	−0.031	0.092
11	Ref	Ref
≥12 years	0.528	0.150
High school graduation track		
Sciences	Ref	Ref
Humanities	1.888	0.117
Economics and social sciences	3.065	0.150
Vocational or technical	1.587	0.105

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B2.—EQUATION FOR THE LENGTH OF STUDIES

Covariates	Estimate	Standard Error
Father's profession		
Farmer or tradesman	−0.232	0.043
Executive	Ref	Ref
Technician	−0.214	0.046
White collar	−0.424	0.040
Blue collar	−0.391	0.042
Unknown	−0.238	0.060
Mother's profession		
Farmer or tradesman	0.010	0.053
Executive	Ref	Ref
Technician	−0.143	0.064
White collar	−0.118	0.038
Blue collar	−0.236	0.049
Unknown	0.070	0.046
Born abroad	0.319	0.079
Woman	−0.063	0.031
Both parents are French	0.165	0.044

TABLE B2.—(CONTINUED)

Covariates	Estimate	Standard Error
Age in sixth grade		
≤10	0.192	0.047
11	Ref	Ref
≥12	−0.313	0.084
High school graduation track		
Sciences	Ref	Ref
Humanities	−0.484	0.036
Economics and social sciences	−0.267	0.031
Vocational or technical	−1.051	0.045
Proportion of students in the first two years of college	−1.306	0.063
Leaving postsecondary education in 1998	−0.446	0.035

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B3.—EARNINGS EQUATION

Covariates	Estimate	Standard Error
Both parents are French	0.004	0.016
Paris region	0.118	0.015
Female	−0.074	0.020
Born abroad	0.009	0.039
Leaving postsecondary education in 1998	0.366	0.030
Major		
Sciences	Ref	Ref
Humanities and social sciences	0.056	0.039
Law, economics, and management	−0.047	0.040
Level of studies		
Dropout	Ref	Ref
Associate degree	0.397	0.050
Bachelor's degree	0.496	0.037
First year of master's degree	0.534	0.046
Master's degree and Ph.D.	0.760	0.039
Interactions between the Major and Educational Level		
Humanities and social sciences		
Dropout	Ref	Ref
Associate degree	−0.087	0.053
Bachelor's degree	−0.155	0.043
First year of master's degree	−0.208	0.040
Master's degree and Ph.D.	−0.194	0.048
Law, economics, and management		
Dropout	Ref	Ref
Associate degree	0.044	0.049
Bachelor's degree	−0.003	0.047
First year of master's degree	−0.008	0.041
Master's degree and Ph.D.	−0.076	0.053
Interactions between Gender (Female) and Educational Level		
Dropout	Ref	Ref
Associate degree	0.058	0.037
Bachelor's degree	0.090	0.036
First year of master's degree	0.136	0.037
Master's degree and Ph.D.	0.018	0.036
Interactions between a Dummy for the year 1998 and Educational Level		
Dropout	Ref	Ref
Associate degree	−0.267	0.042
Bachelor's degree	−0.259	0.031
First year of master's degree	−0.231	0.034
Master's degree and Ph.D.	−0.054	0.051
Interactions between a Dummy for the year 1998 and the Major		
Sciences	Ref	Ref
Humanities and social sciences	−0.121	0.031
Law, economics, and management	0.016	0.035

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B4.—OTHER PARAMETERS

	Estimate	Standard Error
Thresholds		
s_2	−2.556	0.067
s_3	−2.154	0.070
s_4	−1.472	0.067
s_5	−0.710	0.069
Type probabilities		
Type 1	0.380	0.004
Type 2	0.337	0.004
Type 3	0.283	0.004

Covariance matrix of residuals (standard errors in parentheses):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1.053 & 0 \\ (-) & (-) & (\mathit{0.129}) & (-) \\ 0 & 1.053 & 2.379 & 0 \\ (-) & (\mathit{0.129}) & (\mathit{0.318}) & (-) \\ 0 & 0 & 0 & 0.516 \\ (-) & (-) & (-) & (\mathit{0.005}) \end{pmatrix}.$$

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B5.—TYPE-SPECIFIC HETEROGENEITY PARAMETERS

	Estimate	Standard Error
Type 1		
$\alpha_{(1.1)}$	0.000	–
$\alpha_{(1.2)}$	1.244	0.091
$\alpha_{(1.3)}$	1.037	0.093
α_2	0.000	–
α_3	8.192	0.038
Type 2		
$\alpha_{(1.1)}$	0.000	–
$\alpha_{(1.2)}$	−1.312	0.091
$\alpha_{(1.3)}$	−2.828	0.141
α_2	−0.363	0.043
α_3	8.111	0.032
Type 3		
$\alpha_{(1.1)}$	0.000	–
$\alpha_{(1.2)}$	−1.363	0.082
$\alpha_{(1.3)}$	−1.571	0.119
α_2	0.502	0.051
α_3	8.089	0.036

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B6.—ESTIMATED PROPORTIONS OF STUDENTS IN EACH MAJOR, AT EACH LEVEL OF POSTSECONDARY EDUCATION, BY TYPE

	Type 1	Type 2	Type 3
Major			
Sciences	0.00	44.66	50.34
Humanities and social sciences	41.02	55.34	17.49
Law, economics, and management	58.98	0.00	32.16
Length			
Dropout	26.40	47.55	0.00
Associate degree	12.43	16.73	0.86
Bachelor's degree	26.81	24.42	4.58
First year of master's degree	19.77	10.53	29.43
Master's degree and Ph.D.	14.58	0.77	65.14

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B7.—MEANS AND STANDARD DEVIATIONS OF LOG EARNINGS DISTRIBUTIONS, BY TYPE

	Mean	Standard Error
Whole sample	8.65	0.57
Type 1	8.62	0.32
Type 2	8.48	0.31
Type 3	8.84	0.29

Source: Surveys Générations 1992 and 1998 (CEREQ).

TABLE B8.—RELATIVE IMPORTANCE OF TYPE-SPECIFIC ENDOWMENTS VERSUS INDIVIDUAL REGRESSORS IN SCHOOLING DECISIONS

Predicted Major Choice	R^2 Type-Specific Endowments	R^2 Regressors
Sciences	0.41	0.54
Humanities and social sciences	0.45	0.69
Law, economics, and management	0.48	0.56
Predicted length of studies	0.86	0.91

Source: Surveys Générations 1992 and 1998 (CEREQ).