UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper
Economic Series 10-16
June 2010

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Wage expectations for higher education students in Spain*

César Alonso-Borrego† Universidad Carlos III de Madrid

Antonio Romero-Medina Universidad Carlos III de Madrid

Julio 13, 2010

Abstract

We use data on expected wages self-reported by college students to assess the hypothesis that the positive gap between expected and actual wages would decrease as students approach graduation. Our estimation results confirm this hypothesis. The amount and the quality of student information, used to forecast wages, improves with student experience. We find that expected wages for first-year students are affected not only by the degree type and academic performance, but also by the variables determining their degree preferences and their household environment. In the case of junior students, the degree type and length affects expected wages, though neither pre-university performance nor household environment influence their wage forecasts.

Key Words: Wage differentials, College choice, Ordered response.
JEL classification: I23, J24, J31, C24, C25.

* We are in debt with Ricardo Mora for valuable comments. We also thank audiences at the 2nd International Conference on Educational Economics in Athens, FEDEA (Madrid), and the XXXIII Simposium of the Spanish Economic Association in Zaragoza. We acknowledge research funding from the Spanish Ministry of Science and Technology, Grant Nos. ECO2009- 11165 (first author) and ECO2008-02738 (second author).

† Corresponding Author. Department of Economics, Universidad Carlos III de Madrid, 28903 Getafe, Spain. Fax: (+34) 91 624 98 75. e-mail: cesar.alonso@uc3m.es.

1 Introduction

The accuracy of income expectations by students and its relation to educational decisions is at the heart of the "human capital" model. Wage expectations influence individual choices on education, investment, and labor supply. As Dominitz (1998) points out, availability of subjective expectations allows to learn about the process of expectations formation, and to improve our understanding of individual behavior. Despite its importance, the number of studies that assess the accuracy of income expectations is small, and the evidence is mixed. This paper contributes to the debate by using individual data on college students, which provide individual expected wages after completing college, to analyze the informational aspects of the problem. Using our data, we can assess to what extent self-reported measures of expected wages are realistic and changes as graduation approaches, by comparing them with average actual wages for young employees with similar college degrees. The results may shed light on how education choices are ruled. In particular, as far as the optimal education level and degree choice is linked to the market returns to education (Betts, 1996), we are intrigued on how accurately college students perceive their future earnings.

Among the previous contributions, we should mention Dominitz (1998), Das and van Soest (1999), and, more recently, Webbink and Hartog (2004) and Jerrim (2009). The results show how differences may arise because of the particular degrees or particular colleges in which the sample is conditioned, differences in the sample sizes, etc. In particular, Webbink and Hartog (2004), using longitudinal data, find that systematic under or over estimation seldom manifests. Finally, Jerrim (2009) finds that full-time college students in the UK usually overestimate their starting salaries.

In our data set, the wage values are surveyed as discrete ordered categories, whereby respondents are offered a choice among several monetary intervals. Therefore, our baseline econometric model consists of a discrete ordered choice model in which the thresholds correspond to known monetary values. Unlike an ordered response model with unknown thresholds, we can identify the scale of estimated parameters and thus obtain predictions of individual wages.

We estimate the model for expected wages, considering two different subsamples according to the time horizon for degree completion. Namely, we consider first-year college students and penultimate-year or junior college students. Our data set contains information about the degree and academic year for each student, as well as gender, pre-university and college academic performance, and socioeconomic background. We also include individual information by each student before entering university, and additional reasons behind their degree choice.

We find that expected wages for first-year students are affected not only by the degree type and academic performance (before college and in college), but also by the variables determining their degree preferences and their household environment. Furthermore, expected wages predicted by the empirical model imply a high level of overprediction with respect to actual observed wages in occupations requiring their corresponding college degrees. In the case of junior students, the type of college degree plays a relevant role in determining expected wages, but neither pre-university performance nor household environment influence their wage forecast. In general, given the expected wages predicted by the empirical model, there remains a positive gap between expected and actual wages for junior students. Hence, the gap between mean expected wages and actual wages tends to decrease as the students' horizon to graduation approaches.

The remainder of the paper is organized as follows. In Section 2 we outline the data sets, the variables, and alternative model specifications. Sections 3 and 4 present the econometric framework and our estimation results. Section 5 provides some concluding remarks.

2 Data

2.1 The survey

The primary source of data is a survey financed by the Madrid regional authority and carried out in the academic years 2000/2001, 2003/2004 and 2004/2005. The survey explored attitudes and opinions with regard to the higher education system of young students registered in public universities in the Madrid region. The survey design is based on a nationwide data set produced jointly by the Centro de Investigaciones Sociológicas (National Sociological Institute) and the Ministry of Education in 1990, known as "Los jóvenes ante la Universidad" ("Young people facing college education").

Our data set provides information regarding wages expected after graduation. Each student is asked her expected monthly wage after concluding her studies: "What is the monthly wage that you are expecting after graduating?". Among the 1659 students surveyed from all public universities in Madrid, we had 288 who did not answered or answered "Don't know". The answers provided by 1371 students were surveyed into five discrete categories. The categories are: between 450 and 901 euro; between 901 and 1803 euro; between 1803 and 3606 euro; between 3606 and 5409 euro; and more than 5409 euro. In Table 1, we show the marginal relative frequencies of expected wages for each wage category in our sample. Expected wages exhibit a remarkable unimodal profile, whereby 53 percent of students chose

the third category (between 1803 and 3606 euro per month). We also find that a sizeable proportion of respondents reported expected wages in the highest, unbounded category (more than 5409 euro). If we focus on first-year and penultimate-year or junior students, we still observe the unimodal profile, corresponding to the third category, while the proportion of right-censored expected wages is much higher among first-year students.

The data set also contains information on gender, academic and personal status, and socioeconomic background for each student. For the latter, there are data on parents' education, their labor market status, and household income. The academic information provides details on secondary studies completed by the respondent, the ranking of alternative degrees considered, the degree actually followed, and college performance. Information on secondary (pre-university) studies includes details such as whether the secondary academic center was public or private (Public secondary), if the science field of specialization was attended (Science secondary), the score achieved to access university (Access grade), and whether the access examination was passed at the first attempt (Access at first attempt).

We also gathered specific information about the characteristics of the degree chosen by the student, as well as on the choice motivation and the alternative degrees considered. In terms of alternatives considered, respondents had to provide a prioritized list of alternative colleges within the Madrid university district and in Spanish universities outside of Madrid. We consider information on whether the first three choices of the respondent featured a particular degree that could be chosen in several universities (Same degree). Data on university studies included details on whether the attended degree was the student's first choice (First choice), the degree attended and its duration (short or long). For the sake of comparison with complementary data, we have grouped the degrees into five categories: Science & Engineering (S&E hereinafter), Health, Social Sciences (which also includes Law), Humanities, and Education.

Information on college performance includes the degree course, whether the student has repeated an academic year because of failing grades (Repeater), whether she was granted a scholarship (Grant), and whether she is working and/or searching for a job (Work). It must be noted that grants are awarded for economic reasons, but certain requirements on academic performance must be accomplished.

In Table 2, we provide the statistics of the main variables. A 57 percent of the respondents were women, which is close to the 52 percent of women registered in Madrid public universities at such time. Concerning family characteristics, approximately 20 percent reported that they belonged to a high-income household. Information on the educational level of parents is collected in eight categories:

illiterate, below primary, completed primary, professional high school, lower secondary, complete secondary, short university degree, and long university degree. We find that the educational levels of both parents are highly correlated: the t-statistic for linear regression of mother's education vs. father's education is 28.72, with a Kendall statistic for ordinal correlation of 0.46 and a p-value of less than 0.001 percent. We thus concentrate on the educational level of the fathers, in particular, whether the father has a university degree (University father). The percentage of respondents whose father achieved a university degree (long or short) amounts to 41 percent of the sample.[1]

Nearly 60 percent of the students undertook secondary studies in a public high school, and approximately half followed a science field of specialization in secondary education. The minimum requirements to study in any Spanish university are completing secondary education, and passing an access examination with a minimum grade of 50 points out of 100. Since all sample individuals are actually college students, their minimum access grade was 50, with an average value of 68 points. Besides, 84 percent of the respondents passed the access examination at their first attempt. With regard to alternative colleges considered, 15 percent considered the same degree offered in different colleges. Vocation and Economic independence motivate degree choice for about 70 percent of students. Other reasons such as family influence and the difficulty of the degree were alleged by 40 percent of students in the sample.

The attended degree was the first choice for approximately 60 percent of the sample. Long degrees clearly dominate, amounting to80 percent; of these, approximately 35 percent correspond to S&E disciplines, 27 percent to Social Sciences, 18 percent to Humanities, 13 percent to Health, and the rest to Educational oriented studies. The performance of college students in our sample can be summarized as follows. Less than 20 percent were awarded a grant. Approximately 30 percent of the students have failed and repeated at least one academic year. Finally, 20 percent of students reported that they are satisfied with their studies, and nearly 20 percent are simultaneously working (full-time or part-time work) and studying.

Splitting the sample statistics by gender reveals differences in family income; the percentage of students belonging to high-income households is clearly lower for females than for males. The breakdown by degree groups reveals some interesting patterns. S&E are dominated by men, whereas Health and Education are dominated by women.

[1] The remaining parental educational levels correspond to between 10 and 18 percent of respondents, except for the two lower levels, which jointly account for 15 percent of the sample. Compared to the Spanish population as a whole, the educational level of sample fathers is slightly above the average educational level of Spanish parents with children of university age. This same result is observed if we consider maternal education. This bias is coherent with the pervasive intergenerational inertia in educational levels within the same family.

Female students are somewhat different than male students, particularly in terms of academic performance. A higher percentage of women passed the access examination at their first attempt, and a higher proportion of women are following college degrees that were their first choice. Women also seem to perform better at college, with a higher proportion of grants awarded, and a lower proportion of repeaters. Also, a higher proportion of women report to be satisfied in college. This preliminary information in Table 2 only allows comparison of sample averages, but most differences are generally non significant. Besides, a conditional analysis is needed to provide a proper account of these apparent differences.

2.2 Complementary data

We complement the information from our primary data source with the Survey of Wage Structure, carried out by the National Institute of Statistics (INE hereafter, which is the Spanish acronym) to investigate the structure and distribution of wages in Spain for a variety of variables such as age, sex, education level, and region of residence. For comparison with our primary data set, we use 2002 wage data.

In Table 3, we present the marginal distribution of actual wages for young college graduated employees, between 25 and 35 years old. For the sake of comparison, we have split wages into the same discrete groups as for our sample expected wages in Table 1. We can see that the wage distribution in Madrid is shifted to the right with respect to the distribution at the national level. In line with expected wages by college students, the distribution for actual wages in Madrid exhibits a unimodal profile in the third category (between 1803 and 3606 euro). However, there are differences with the distributions of expected wages, specially for first-year students, with less than 16 percent expecting wages below 1803 euro, while the proportion of young graduate wages in Madrid within such wage interval is around 35 percent. The distribution of wages expected by junior students is slightly shifted to the right with respect to the distribution of actual wages for young graduates in Madrid, yet they do not differ very much. Then, the comparison of the marginal distributions points out that college students tend to overpredict wages after graduation, with the extent of overprediction being much higher for first-year than for junior students.

In Table 4, we present the average monthly wage for young college graduated employees, split by gender and by the occupation related with each degree type and length. Since this information is widely publicized and easily accessible, it is reasonable to assume that it is part of the information set that

university students used when computing their expected wages. In this breakdown, we must note there are not short degrees in Humanities.

It must be noted that the average wages in this complementary data set are representative of the population of employees.[2] Therefore, such information is potentially affected by two sources of selection bias. The first one is related to the decision on labor participation, which differs for women and men. In the age range 25-35 years, females exhibit lower participation rates than men. The second source of selection bias arises from the fact that the Survey of Wage Structure reports wage earnings for employees, and therefore is restricted to those who decide to be wage earners and find a job. However, it is not possible to control for these sources of sample selection, since both participation decisions take place after graduation and may thus be conditional on events that take place after the survey. In any case, we use the data in Table 4 as a benchmark to evaluate expected wages of college students in our sample.

Analysis of the data in Table 4 reveals three findings. First, average earnings are greater for men than for women, with a gap ranging between 20 and 35 percent. Second, average wages are generally higher for occupations requiring long degrees with respect to short degree occupations. The only exception appears for occupations related with Educational degrees, for which short degrees exhibit, at the national level, slightly higher average earnings than those with long degrees, while, for the Madrid region, the wage gap remains positive for women. The wage gap between long and short degrees is remarkably high in Social Sciences, being even higher in the case of Madrid. For Science & Engineering and Health, men exhibit higher wage gaps between short and long degrees than women. Third, employees in the Madrid region with degrees in S&E, Social Sciences and Humanities, enjoy earnings that lie substantially above the national average. Differences in the cost of living and in industrial composition account for these differentials. However, in the case of Health and Educational degrees, the national average wages are similar to those in Madrid. We also find that long S&E and Social Sciences degrees are the ones with the highest average wages, whereas Education degrees show the lowest average wages in any duration. S&E also exhibit the highest average wages among the short degrees.

There is a positive gender wage gap between men and women for most degree types, with the exceptions of long Educational degrees (both in Spain and in Madrid), and short Health degrees in Madrid, for which the gender wage gap is negative. Among the potential reasons for the gender gap, we should mention three: pure gender discrimination; the possibility that, with all other things equal, firm-specific accumulated human capital tends to be lower for women because they are more likely to experience

[2]We have disregarded self-employment status, for which nevertheless there are not reliable data sources about earnings. The proportion of young college graduated who were self-employed in Madrid, in accordance with the data from the 2002 Spanish Labor Force Survey, was smaller than 10 percent.

discontinuities in their professional career; and occupational segregation. In the latter case, women are more likely to face restrictions that force them to choose occupations with lower wages in exchange for non-wage compensations such as greater time flexibility. Regarding this, in those occupations for which the wage gap is reversed (related with long Educational and short Health degrees), women predominate.

We concentrate our analysis on two different subsamples of students, which correspond to extreme cases of the time to graduation, first-year and penultimate-year or junior students. We expect both their characteristics and wage expectations to differ very much for these two particular groups. In particular, we expect students closer to completion to have much lower uncertainty about their academic prospects, as well as a better informed assessment of their job market prospects after graduation.

In Table 5 we present the sample distribution of college students in our sample, for these two particular groups. In line with our earlier comments, our sample exhibits a low proportion of men in Educational and Health degrees of any duration, whereas S&E degrees, specially of long duration, are dominated by men.

3 Empirical analysis

3.1 Basic model

We use a stylized model of human capital accumulation and investment in education that suits the needs of our empirical analysis as in Webbink & Hartog (2004). For any individual, we assume that her individual wage, W^*, is proportional to her amount of human capital, H. Assuming, without loss of generality, that unobserved individual factors are on average equal to zero, the expected log wage for a level of education S and a given set of observed individual factors equal to Z is $\alpha + \Phi S + \theta_1' \mathbf{Z}$.

Moreover, for a university student in the k-th academic year of her college degree, her expected wage after graduation will depend on the information set determining her expectation. In particular,

$$E_k(\ln W^*) = \alpha_k^e + \Phi_k^e S + \theta_{1k}^{e\prime} \mathbf{Z} + E_k(v), \tag{1}$$

where $E_k(\bullet)$ represents the mathematical expectation, conditional on her information set, and α_k^e, Φ_k^e, θ_{1k}^e represent the expected returns in the wage equation of the corresponding variables in that information set. Assuming that $E_k(v)$ is equal to zero, then the expected average wage becomes $\alpha_k^e + \Phi_k^e S + \theta_{1k}^{e\prime} \mathbf{Z}$.

Therefore, the differential between average expected wages and average actual wages arises from the differences between the expected and actual returns of each variable,

$$[E_k(\ln W^*) - E(\ln W^*)] = (\alpha_k^e - \alpha) + (\Phi_k^e - \Phi)S + (\theta_{1k}^e - \theta_1)'\mathbf{Z}. \quad (2)$$

Note that this differential ultimately depends on the distribution of information across students. Student information sets are related to the amount and quality of a student's knowledge about the economic value of her college degree, and to the time until receiving a wage as a graduate, i.e., her prediction horizon. We thus expect that the gap between expected and actual wages would be greatest at the beginning of a university course and would decrease as the student approaches graduation.

Fresher students face much more uncertainty about their career prospect than junior students. First, their graduation probability is much lower for the first-year students, so that the effort that they put on computing an accurate forecast of their future wages is much lower. Second, in general, the time horizon until they get into the labor market is much longer for first-year students. Third, first-year and junior students differ in their attitudes and their maturity. The first ones have finished secondary school very recently, while the last ones are very prone to graduating and getting into the labor market. Since gathering information is costly, first-year students are less willing to pay effort on learning about wages after graduation than junior students, who expect to be searching for a job in a much shorter time.[3] Fourth, the weights and the characteristics that the different aspects of the personal environment have differ among first-year and junior students. In both cases, their personal environment is determined by their household and their college peers. In the case of first-year students, information is dominated by the one provided by the household, so that when forecasting wages they rely much more on the earnings opportunities of their parents. But, since parents are in a very different stage of their life cycle, the information that provides about expected earnings may be less informative than the information that graduates in their first job can provide. Junior students are also influenced by their corresponding household. But the characteristics and attitudes of their peers are very different: they are more mature and much more concerned with their labor market prospects. Also, they are much closer to fresh graduates already working or searching for a job.

[3]Betts (1996) poses the existence of countervailing forces which make uncertain when information acquisition occurs more intensively. On the one hand, the marginal value of information may be greater in the early degree years, before high sunk costs make it costly to the student to change her career path. On the other hand, as far as information about the labor market acquired by the student does mainly come out from informal exchanges with peers, faculty, and others, then more experienced students might show an informational advantage over freshmen.

3.2 Empirical specification

The sample students differ in their academic and personal information and in their degree year, so that there is individual heterogeneity in their levels of human capital accumulation and other individual characteristics. Such heterogeneity affects how students forecast their wages after college completion. In particular, with all other things equal, differences in the degree year, which reflect the time to completion, affect the student's opportunity cost of education, as well as the amount and quality of her information. These differences may thus lead to differences in subjective valuation of the same college studies. We therefore consider the aforementioned breakdown between first-year and junior college students. Using this breakdown, the years of education for each group can be taken as constant, and therefore ΦS_i will be part of the constant term for each group.

In addition to the variables that characterize socioeconomic background and may be associated with human capital accumulated before higher education, it is also important to account for further individual characteristics. In particular, gender and the academic curriculum during secondary education may have a systematic effect on the subjective valuation of wages. Thus, we extend the vector of covariates, denoting it as $\mathbf{X}_i$. In addition to unobservables affecting human capital obtained before higher education, there are individual characteristics that are unobserved in the data that affect subjective valuations. Therefore, we can write our empirical model as:

$$\ln W_i^* = \beta' \mathbf{X}_i + u_i. \tag{3}$$

If W_i^* were observed, appropriate estimates of β could be obtained through OLS under certain conditions. However, we have emphasized in Section 2 that we do not fully observe W_i^*, but a discretized version of it, W_i, that can be defined as:

$$W_i = j, \text{ if } \mu_{j-1} < W_i^* < \mu_j \quad (j = 1, \ldots, 5), \tag{4}$$

where the values μ_j, $j = 1, \ldots, 5$ are known. We can also define indicator variables for each category as:

$$d_{ij} = 1(W_i = j) = 1(\mu_{j-1} < W_i^* < \mu_j) \quad (j = 1, \ldots, 5). \tag{5}$$

The censored nature of the observed dependent variable W_i invalidates OLS as an estimation method. We address this problem using the strategy developed for models with multiple ordered responses that

has been applied when using contingent-type data as, for example, in Cameron and Quiggin (1994), Cai, Deilami and Train (1998), and Papke (1998). Our empirical model is thus an ordered response model, yet in our case the thresholds determining the different categories are known, so there is no need to estimate them as parameters.

Even though the observed variable W_i is ordinal, knowing the cutoff points implies that no normalization is required to identify the vector β and the likelihood function will generally depend on both β and $Var(u_i|\,\mathbf{X}_i) = \sigma^2$. Maximum likelihood estimation can be carried out after assuming a distribution for u_i, $F(.)$. The probability that respondent i chooses wage category j is:

$$\Pr(W_i = j|X_i) = \Pr(\mu_{j-1} < W_i^* < \mu_j) \tag{6}$$

$$= \Pr(\ln \mu_{j-1} < \ln W_i^* < \ln \mu_j) \tag{7}$$

$$= F(\ln \mu_j - \beta' \mathbf{X}_i) - F(\ln \mu_{j-1} - \beta' \mathbf{X}_i). \tag{8}$$

Then, the log-likelihood takes the form:

$$\ln L(\beta, \sigma) = \sum_{i=1} \sum_{j=1} d_{ij} \ln \Pr(W_i = j|\,\mathbf{X}_i).$$

Given our knowledge of thresholds, we can obtain projections for expected wages as in a standard linear model.

Note, in contrast, that if the cutoff points were not known, the parameter vector would only be identified up to a normalization. In such a case, it is usually assumed that $\sigma = 1$ and, therefore, the scale of β conveys no information. An important practical advantage of exploiting wage thresholds by means of the pointwise censored model is that we do not need further assumptions about the distribution of the right tail to compute individual expected wages. More precisely, in an ordered probit in which the information on threshold values is not exploited, we must introduce an additional assumption for the right tail of the wage distribution (for declared expected monthly wages above 5409 euro). Using results from the standard ordered probit estimates, we have found that predicted individual expected wages are very sensitive to this additional assumption.

4 Results

To assess the quality of wage forecasts by college students in Madrid, we exploit the information about expected wages reported by the students in our sample. Our estimates can be subsequently used to compute individual predictions of expected wages and compare them with average actual wages for working graduates. The values reported for expected wages represent subjective valuations. This means that the estimated effects of the conditioning variables will combine the influence of these variables on the potential wage, on the one hand, and the quality of the information used in computing wage expectations, on the other.

Expected wages are censored into five wage categories, with the highest category being unbounded to the right. Given that we observe wage thresholds, the scale of the parameters is identified. Thus, the variance of the error term can be estimated, together with the remaining parameters of interest, by maximum likelihood. Moreover, although both the ordered probit and pointwise censored models are consistently estimated by maximum likelihood, the latter is more efficient as it exploits the information available on monetary thresholds in the questionnaire.

4.1 Determinants of expected wages

The maximum likelihood estimates for first-year and junior students are reported in Tables 6 and 7, respectively. The conditioning variables, which have been described in Table 2, contain individual and household characteristics, curricular variables before and during the college studies, characteristics of the degree, and reasons why the degree was chosen. Among the degree characteristics, we have considered binary variables for the degree type: Science & Engineering (S&E), Health, Educational, Social Sciences and Humanities. These binary variables have also been interacted with gender, the degree duration (short or long), and whether the individual has repeated at least one year because of failing grades. For both first-year and junior students, we have considered three different specifications. The first one corresponds to the more general model, including all the selected variables, and the last one is the model of our choice, being the one that provides the best fit to the data. To achieve such specification, we have tested for the significance of several variables, both at the individual and at the joint level, removing those which were clearly insignificant. In particular, we have tested for the joint significance of the set of variables corresponding to degree types interacted with gender, degree duration, and being a repeater. The p-values of the corresponding tests are reported in Tables 6 and 7. The model adjustment is reasonably good for the two student groups. We will concentrate our comments on the results regarding

our preferred specifications.

We first discuss the estimates for first-year students. Among the pre-university variables, those related with access grade, as a measure of academic performance shortly before university entrance, are irrelevant in determining the expected wages of first-year college students. However, we find that a Science curriculum in high school has a negative and significant effect on wage expectations. Among the characteristics behind the choice of degree, both the difficulty of the studies and the fact that the student consistently applied to the same degree in different colleges exhibit a negative and significant effect. We infer from this result that students who show strong preferences for a specific degree tend to expect lower wages than those showing a higher taste for degree diversity.

Among household variables, having a father who has a university degree, as well as living in a high-income household, has a positive effect, yet only the first variable is clearly significant. This implies that the higher the educational level and/or the income within the household, college beginners tend to expect higher wages after completing their studies.

The effect of gender is significantly negative, and its magnitude does not depend on the type or the duration of the degree. Hence, women realistically expect lower wages than men with similar characteristics. The sign of the gender effect is consistent with the fact that women present higher college attendance and better academic performance, as well as a greater expected probability of college graduation. This propitiates a greater effort in information gathering that results in a lower wage forecasts. However, the estimated magnitude does not offset the positive gender gap which is observed between men and women.

With regard to college characteristics, we find a negative, though marginally significant, effect of short degree. Such effect is similar for any degree, because the interactions with the degree type were non significant. Hence, students attending short college degrees expect, in general, lower wages. We also find that the set of binary variables that control for degree types are individually and jointly significant. Given that the reference group (for which we have omitted its binary variable) is S&E, and given that all of them exhibit negative coefficients, we can conclude that students attending degrees different than S&E expect lower wages. This occurs mostly for Humanities and, very specially, Educational degrees. This results resembles the evidence reported in Table 4, by which these two degree groups exhibit the lowest average actual wages among college graduated employees.

We have also controlled for college under-performance through the variable Repeater, for which we have allowed for interactions with the degree type. Students in Social Sciences degrees who have repeated

tend to compute higher expected wages. However, for the remaining degrees, we do not find a significant effect of having been a repeater.

Besides, we also find that college students declaring to be satisfied in their studies tend to expect higher wages after completion. This variable is a subjective indicator of college satisfaction, which captures a different effect than the measures of college performance, like Repeater. Actually, the correlation between Satisfied and Repeater is below 6 percent and clearly non significant.

In the case of junior students (whose results are shown in Table 7), the number of variables affecting the magnitude of wage expectations is smaller than for first-year students. In particular, neither pre-university nor household variables exhibit any effect on expected wages. These variables had some importance with regard to the college degree chosen, so that they are related with the ability to predict expected wages shortly after the students have chosen and started their college degree. However, after a long time since their degree choice, and when the completion horizon is much shorter, such variables do not longer play a role in the information set by which students generate their expectations. The only exception is the variable indicating whether the degree undertaken was the first student's choice, with a marginally significant negative effect. This result suggests that students who succeeded in entering their most preferred degree tend to expect lower wages.

Unlike first-year students, neither gender nor satisfaction appear to play any role on expected wages. Therefore, experienced students of any gender with similar characteristics expect similar wages after completing college studies. This evidence contrasts with the fact that, other things equal, women exhibit lower observed wages than men.

With regard to the degree characteristics, we find that following a short degree has a negative and significant effect on wage expectations for those following Education, Humanities, and to a lesser extent, Social Sciences degrees. However, we do not find significant differences by degree duration in the case of S&E degrees.

Regarding college performance, being a repeater affects differently depending on the degree type. In particular, we find a positive and significant effect for S&E and Educational degrees (which are the reference groups), null effect on Social Sciences (given that the corresponding interaction fully offsets the effect), and a negative and significant effect for Humanities degrees. Our findings are in line with the results in Webbink and Hartog (2004) that repeaters expect higher wages.

4.2 Wage forecasts

In Table 8, we use our preferred estimates from Tables 6 and 7 to predict student mean expected wages by gender and degree type. We have also produced, in Table 9, the percentage difference between mean wage expectations and the mean actual wages for young college graduated employees in Madrid (reported in Table 4).

In comparison with the mean actual wages reported in Table 4 (either in Spain or in Madrid), mean expected wages for first-year students are, in general, much higher than actual wages for employees in occupations related with such degree.[4] Interestingly, although mean expected wages are systematically higher than mean actual wages, the rankings of wages expected by first-year students and actual wages by degree and gender are very similar. Namely, mean wages for long Social Sciences, and very specially, S&E degrees are generally among the highest, while the lowest mean wages usually correspond to Educational degrees. We find that, for most degree types, expected wages are greater than average actual wages for young college graduated employees in Spain, and even in the Madrid region, where wages are higher. Hence, first-year college students tend to overestimate their potential wages to a great extent. With the exceptions of Health and short Education degrees, expected wages for first-year female students are, on average, lower than those for men in the same group. However, considering that actual wages for graduates are systematically lower for women than for men, the relative level of overestimation is still higher for female than for male first-year students.

For each degree and gender, in general, first-year students expect much higher wages than junior students. The fact that mean expected wages, in general, move closer to mean actual wages for junior than for fresher students is in accordance with the hypothesis that the wage expectations formation improves as students approach graduation. For each degree duration, the mean expected wages for junior students are very much alike among degrees, with the only exception of short Educational degrees, which exhibit much lower mean expected wages than the remaining short degrees. In addition, although gender had no significant effect on expected wages for junior students, gender differences in mean expected wages remain. We attribute this result, as we have already discussed in the descriptive evidence, to the fact that women and men differ each other in terms of pre-university and college academic performance, and other characteristics that affect wage forecasts. However, for each degree and duration, the differences in expected wages by gender are narrower when students are closer to graduation.

[4] The only exception is for male students in short Educational degrees, for which the mean expected wage is in line with the mean actual wage.

In fact, there are substantial differences in the mean actual wages by gender and degree, to a much greater extent than the differences found for mean expected wages. Consequently, we observe substantial differences in the gap between mean expected wages and mean actual wages for junior students by degree and by gender. We find two extreme cases. Long Educational degrees exhibit the highest gap between mean expected and actual wages. On average, Education junior students overpredict even more than first-year students. On the other hand, long Social Sciences degrees exhibit a negative gap for any gender. In the remaining degrees, junior students keep overpredicting wages after graduation, though to a much lesser extent than first-year students.

To understand these results, it must be noted that the actual wages by degree that we use as reference correspond to young college graduated who are employed in occupations that typically require the corresponding degree. Hence, when making the comparison between expected and actual wages we are considering observed wages in specific occupations that are directly linked to the corresponding degree.

In the case of Educational degrees, we deduct from Table 4 that the occupations that are directly related to such degrees are worse paid than occupations associated to other college degrees. The higher gap between expected and actual wages can be partly explained by the fact that some students will consider aiming at other occupations that are not specific of Educational degrees, which can provide them better salaries.

Concerning junior students in Social Sciences degrees, the negative gap between expected and actual wages suggests that they exert very pessimistic expectations as they approach graduation. There are several explanations to these results. First, in the case of many Social Sciences graduates, their first job after college usually takes the form of a training contract. The remuneration of this contract is quite below the one of a standard contract. If students expect such situation, it is clear that their wage forecast is associated with their short-run earnings after graduation, and, therefore, their expectations would be much lower than the average wage of young employee graduated in Social Sciences. Second, and more realistically, the career path for a Social Science graduate is more uncertain than for other graduates. Whereas there are occupations that require being a graduate in degrees like Health or S&E, most occupations associated with Social Sciences degrees are not exclusive of graduates in such degrees. This makes that any long degree graduate can compete for Social degree occupations. As a consequence, the prospects of Social Sciences graduates may appear more uncertain than in the aforementioned degrees. In particular, a large fraction of them may end up underemployed, which is, indeed, the case. Very often, long degree graduates end up working in Social Science occupations that only require short college

degrees, or even lower education levels. Unlike first-year students, junior students in long Social Sciences degrees perceive this potential situation, and weight this possibility when computing their expected wages.

In order to ascertain what characteristics make students to incur in systematic prediction errors, in Table 10 we report, for first-year and junior students, the percentage difference between mean expected wages by degree type and duration, in accordance with enjoying or not three relevant characteristics. These are Repeater, High income, and University father. It must be noted that some cells have been computed with very few observations. We find that being a repeater is the main source of distortion, to the extent that junior repeaters, in comparison with their remaining peers, still heavily overpredict. Having a university father is another source of distortion among first-year students. However, this source of distortion vanishes for juniors. Interestingly enough, living in a high income household is fairly neutral, although for most degree types the number of observations within cells is very small.

4.3 Accuracy of expectations

A conclusion from our results is that as the time horizon towards graduation shortens, students' ability to realistically compute their expected wages improves. For first-year students, variables related with academic performance and with the features behind the degree choice affected wage expectations. In particular, those with better pre-university performance, those with preferences for a specific degree, those to whom a higher difficulty of college studies motivated degree choice, tend to expect lower wages. Also, for certain degrees, college repeaters (which indicates under-performance) tend to expect higher wages. We can then conclude that positive curricular characteristics make students to be more realistic in forecasting future wages. Regarding household variables, a higher father's educational level, and a high household income, make first-year students more optimistic about future wages after graduation. We thus observe that past and current academic performance, as well as family background variables, affect how first-year students compute wage expectations.

On the contrary, junior students are not affected by pre-university curricular variables nor by family background when computing their expected wages. After some years in college, these groups of variables, which affected degree choice, are no longer relevant in computing expected wages. In addition, gender is not relevant either. Mostly, the type and the duration of the degree are the major variables affecting wage expectations of junior students. We interpret such differences between fresh and junior college students as follows. Individual characteristics capture both individual quality effects and individual

ability to compute expected wages. Such ability is much more limited for first-year students, for which information uncertainty on both career prospects and future wages is greater. Among these students, those with lower quality (i.e., with worse academic performance), those less motivated when choosing college degree, and those with a better economic situation within the household, tend to fail more in forecasting future wages.

It must be noted that the overestimation of expected wages with respect to actual wages for young working graduates is actually greater than reflected in Table 8. This is because the individuals in our sample are not strictly comparable with the sample for which average actual wages were computed. This later sample is restricted to young graduates who have decided to work and have indeed found a job. In contrast, our sample comprises students who have not yet graduated. For those who graduate, some will eventually not work, either because they decide not to enter the labor market or because they will not find a job. Moreover, a proportion of them will drop out of college before graduation. Therefore, it is possible that part of the apparent improvement in the formulation of expectations with increasing degree years merely reflects sample selection of students who are much more likely to work in jobs that require a university education. Our results are consistent with Betts (1996), who find that student in higher years prove to be much better informed with respect to the labor market than fresher students.

It is worth mentioning that a sizable proportion of respondents failed to declare either their expected wages or the degree that were attending, what leads to a potential sample selection problem (cf. Heckman, 1979). The proportion of non respondents amounted to 22 percent among first-year and 17 percent for junior students. Namely, if unobserved factors in respondents' wages were correlated with factors affecting the answer probability, then estimates based on the subsample of respondents would be inconsistent. A proper treatment of this problem would require a sample selection model with an additional auxiliary equation that would account for the respondent's decision to answer their expected wages and their college degree. The fact that expected wages are observed in a discrete fashion complicates the treatment of selection bias, which would require estimating a sample selection pointwise censored model. Miranda and Rabe-Hesketh (2006) have proposed maximum likelihood procedures and developed subroutines in Stata to estimate nonlinear models like ours subject to sample selection. However, there are two major problems when estimating sample selection nonlinear models. First, maximum likelihood estimation is computationally very demanding. Second, and more importantly, the concavity of the likelihood function is not ensured, so that it may occur that convergence is not attained in the presence of a moderate number of covariates. Indeed, we find this problem when we approach the estimation of our specifications of interest controlling for sample selection.[5]

[5]In an earlier version of this paper, we estimated simplified versions of our expected wages model, in which we consider

5 Conclusions

It is well known that expected earnings and expected returns to education are major determinants in deciding both the type and the amount of education attended. However, empirical knowledge about expectations and their formation is scarce. Furthermore, most of the existing studies exploit data collected from European and US students, mainly in business and economics degrees, to study the determinants of wage expectations, but not students' capacity to predict future earnings (see Brunello et al, 2004, for references).

This paper models the wage forecasting of college students in Madrid universities. We use a microeconomic data set previously exploited by Alonso-Borrego et al. (2007) that includes academic, personal and household characteristics, as well as reported expected wages. This data set includes students from all universities and most degrees available in Madrid. This rich data set avoids concentrating on students in a particular type of college degree. Differences in time to completion may affect students' subjective valuation of college degrees. Such differences may affect individual processing of relevant information. For this reason, we considered two different subsamples, first-year and junior students.

Since expected wages are surveyed into five discrete categories, OLS estimation was inappropriate. Instead, we must consider an ordered response model to account for the nature of the observed dependent variable. We also exploit information on wage thresholds to obtain more efficient estimates than those provided by a standard ordered probit model with unknown thresholds. We cannot asseverate that our sample is representative of the population, but we are confident that our data and our empirical approach allows to circumvent most of the drawbacks because of the data sets exploited in the related literature.

We have found that the degree type has a relevant role on expected wages. There were also differences depending on the student degree year, so that expected wages depend on pre-university academic performance for first-year students and on college performance for later-year students. Comparison of mean predicted expected wages with mean actual wages for young working graduates reveals a positive gap, which reflects that college students tend to overpredict their wages after graduation. This gap tends to narrow for junior students. This reflects that expectations became more realistic as students approach graduation. The information set is strongly influenced by the student's personal environment

a smaller number of covariates, ignoring most interactions, and estimated a simpler ordered probit model, in which we ignored the known values of the wage thresholds, subject to sample selection. Under such estimates, we could not reject the hypothesis that sample selection was exogenous. Unfortunately, when we attempted to estimate our model of interest subject to sample selection, we were not able to achieve convergence.

at the beginning of college studies. Such influence dilutes as students approach graduation and their uncertainty about their career prospects is reduced.

The role of gender in the change on wages expectations is particularly intriguing. While women in their first-year expect, other things equal, lower salaries, such differences among men and women vanishes for junior college students. Again, new information sources dilute the previous perception. It is important to remark that, despite the lack of significance of gender, junior females expect lower salaries than males, what is due to differences among men and women in other characteristics.

Shortly after high school graduation, college students systematically overpredict their future wages. Even though first-year female students realistically predict lower wages than comparable men, their level of overprediction with respect to actual wages is higher. As college students approach graduation, their wage expectations, with some exception, become more realistic, and tend to be closer, on average, to the corresponding actual wages. It is important to remark that the breakdown by degree types shows differential patterns in accordance with the specific career prospects of each degree. We also find that students with a better defined career path exhibit wage expectations that keep closer to the observed wages. We also find that systematic mistakes are due to the different sources of information and the different weights that students give to such information sources along their college cycle.

References

[1] Alonso-Borrego, C., R. Mora and A. Romero-Medina (2007), Una valoración de la enseñanza superior a partir de los salarios esperados y los salarios sombra de los estudiantes universitarios. Cuadernos Económicos de ICE 74: 11–38.

[2] Betts, J.R. (1996), What do students know about wages? Evidence from a survey of undergraduates. The Journal of Human Resources 31(1), 27-56.

[3] Brunello, G., C. Lucifora, R. Winter-Ebmer (2004), The wage expectations of European Business and Economics students. The Journal of Human Resources 39(4), 1116-1142.

[4] Cameron, T.A. and J. Quiggin (1994), Estimation using contingent valuation data from a dichotomous choice with follow-up questionnaire. Journal of Environmental Economics and Management 27(3): 218–234.

[5] Cai, Y., I. Deilami and K. Train (1998), Customer retention in a competitive power market: Analysis of a 'double-bounded plus follow-ups' questionnaire. The Energy Journal 19(2): 191–215.

[6] Das, M., and van Soest, A. (1999), A panel data model for subjective information on household income growth. Journal of Economic Behaviour and Organization 40: 409-426.

[7] Dominitz, J. (1998), Earnings expectations, revisions and realizations. The Review of Economics and Statistics 80 (3): 374-388.

[8] Dominitz, J., and C.F. Manski (1996), Eliciting student expectations of the returns to schooling. The Journal of Human Resources 31(1), 1-26.

[9] Jerrim, J. (2009), Wage Expectations of UK Students: Are They Realistic? Southampton Statistical Sciences Research Institute. Mimeo.

[10] Heckman, J. J. (1979), Sample Selection Bias as a Specification Error. Econometrica 47(1), 153-161.

[11] Miranda, A. and S. Rabe-Hesketh (2006), Maximum likelihood estimation of endogenous switching and sample selection models for binary, count, and ordinal variables. The Stata Journal 6(3): 285–308.

[12] Papke, L.E. (1998), How are participants directing their participant-directed individual account pension plans? American Economic Review 88: 212–216.

[13] Webbink, D. and J. Hartog (2004), Can students predict starting salaries? Yes! Economics of Education Review 23: 103-113.

Table 1

Monthly expected wages of Madrid college students

Relative frequency (%)	All	1st-year	Junior
Between 450 and 901 euro	4.9	2.6	6.3
Between 901 and 1803 euro	17.9	12.7	22.2
Between 1803 and 3606 euro	52.7	47.5	55.6
Between 3606 and 5409 euro	13.8	18.7	9.9
More than 5409 euro	10.8	18.4	6.0
Number of non -missing observations	1371	385	284
Number of missing observations	288	111	58

Source: Young people facing college education, 2001, 2004 and 2005.

Table 2
Main variables and descriptive statistics

	All		Female		Male	
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.
Female	0.57	0.50				
Family						
High family income	0.19	0.39	0.14	0.35	0.26	0.44
University father	0.41	0.49	0.40	0.49	0.42	0.49
Pre-university						
Public secondary	0.58	0.49	0.59	0.49	0.57	0.49
Science secondary	0.52	0.50	0.50	0.50	0.54	0.50
Access grade	67.78	9.32	67.69	9.50	67.90	9.07
Examination passed at first attempt	0.84	0.37	0.87	0.34	0.81	0.40
Choice reasons:						
Economic independence	0.66	0.47	0.63	0.48	0.70	0.46
Vocation	0.76	0.43	0.74	0.44	0.78	0.42
Parental influence	0.40	0.49	0.38	0.48	0.42	0.49
Difficulty	0.39	0.49	0.39	0.49	0.40	0.49
Choice set						
Same degree	0.15	0.35	0.15	0.36	0.14	0.35
University degree chosen						
First choice	0.61	0.49	0.66	0.47	0.54	0.50
Long degree	0.80	0.40	0.79	0.41	0.81	0.39
Science & Engineering (S&E)	0.35	0.48	0.28	0.45	0.44	0.50
Health	0.13	0.34	0.18	0.39	0.07	0.25
Educational	0.07	0.25	0.09	0.28	0.04	0.24
Soc. Sciences	0.27	0.44	0.29	0.45	0.25	0.43
Humanities	0.18	0.38	0.16	0.37	0.20	0.40
College performance						
Grant	0.17	0.37	0.18	0.39	0.15	0.35
Repeater	0.30	0.46	0.27	0.45	0.35	0.48
Satisfied	0.21	0.41	0.25	0.43	0.16	0.37
Working	0.18	0.39	0.18	0.39	0.18	0.39
Survey year						
2004	0.31	0.46	0.25	0.43	0.40	0.49
2005	0.56	0.50	0.61	0.49	0.50	0.50

Source: Young people facing college education, 2001 2004 and 2005.

All the variables are binary except for Access grade, which ranges between 50 and 100.

Table 3

Monthly earnings for college graduated employees aged 25-35 years

Relative frequency (%)	National	Madrid
Between 450 and 901 euro	9.4	7.4
Between 901 and 1803 euro	36.2	28.3
Between 1803 and 3606 euro	46.8	50.5
Between 3606 and 5409 euro	6.3	11.3
More than 5409 euro	1.3	2.5

Source: Calculated from "National Survey of Wage Structure", 2002.

Table 4

Monthly average earnings (in euro) for college graduated employees aged 25-35 years, by degree type and gender

	National average, Long degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2599	2170	2476	1938	1136	2228
	(1396)	(1172)	(2147)	(1095)	(867)	(1564)
Female	2208	1866	2001	1744	1252	1787
	(1074)	(915)	(1482)	(901)	(730)	(1151)
All	2488	1983	2218	1831	1289	2015
	(1324)	(1032)	(1832)	(996)	(7944)	(1397)

	National average, Short degree				
	S&E	Health	Soc. Sci.	Educational	All
Male	2130	1829	1663	1379	1959
	(999)	(660)	(1029)	(658)	(974)
Female	1899	1701	1279	1326	1551
	(822)	(553)	(721)	(607)	(682)
All	2088	1724	1410	1342	1754
	(973)	(576)	(857)	(623)	(864)

	Madrid average, Long degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2870	2104	3243	2172	1298	2682
	(1549)	(1126)	(2940)	(1137)	(1128)	(2023)
Female	2381	1737	2440	1939	1329	2114
	(1139)	(861)	(1816)	(899)	(812)	(1418)
All	2716	1879	2788	2052	1315	2413
	(1449)	(984)	(2399)	(1022)	(963)	(1784)

	Madrid average, Short degree				
	S&E	Health	Soc. Sci.	Educational	All
Male	2408	1578	1852	1320	2250
	(982)	(583)	(1413)	(766)	(1044)
Female	2160	1718	1323	1210	1686
	(757)	(468)	(743)	(598)	(724)
All	2354	1696	1545	1696	2005
	(942)	(487)	(1095)	(487)	(960)

Source: Calculated from "National Survey of Wage Structure", 2002.

Standard deviations in parentheses.

Table 5
Sample distribution of college students in Madrid
by degree type, year and duration of degree and gender

1st-year, Long degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	108	5	40	34	7	194
Female	66	14	69	29	18	196
All	174	19	109	63	25	390

1st-year, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	22	4	10		2	38
Female	18	24	14		12	68
All	40	28	24		14	106

Junior, Long degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	28	8	26	32	5	99
Female	19	18	42	38	14	131
All	47	26	68	70	19	230

Junior, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	29	3	7		3	42
Female	18	30	14		8	70
All	47	33	21		11	112

Source: "Young people facing college education", 2001, 2004 and 2005.

Table 6
Expected wage for first-year college students
Pointwise censored model without selection

Public secondary	−0.0206	−0.0162	
Access grade	−0.0083	−0.0083	
Access at first attempt	−0.7000	−0.6774*	
First attempt × Access grade	0.0110	0.0108*	
University father	0.1039†	0.1055†	0.1230‡
Science secondary	−0.2115‡	−0.2143‡	−0.2195‡
Grant	−0.0118	−0.0136	
First choice	−0.0626	−0.0689	
Ideal choice	−0.0139	−0.0144	
Same degree	−0.1411†	−0.1471†	−0.1638‡
Reason: Economic independence	0.0514		
Reason: Vocation	0.0353		
Reason: Parental influence	−0.0211		
Reason: Difficulty	−0.1100†	−0.1023†	−0.1002†
Female	−0.1579*	−0.1499*	−0.1181‡
Repeater	−0.2063*	−0.2070*	−0.1515
Satisfied	0.1328‡	0.1320‡	0.1098†
Working	−0.0262	−0.0270	
High family income	0.0981*	0.102*	0.1046*
Short degree	0.0199	0.0208	−0.0914*
Health	−0.2424*	−0.2526*	−0.1862†
Educational	−0.4450‡	−0.4544‡	−0.4520‡
Soc. Sci.	−0.2835‡	−0.2929‡	−0.2162‡
Humanities	−0.2972‡	−0.3107‡	−0.3025‡
S&E×Short degree	−0.0469	−0.0498	
Soc. Sci. ×Short degree	−0.2106	−0.2179*	
S&E×Female	−0.0592	−0.0642	
Soc. Sci. ×Female	0.1651	0.1544	
Humanities×Female	0.0617	0.0382	
S&E×Repeater	0.2228*	0.2264*	0.1910
Soc. Sci. ×Repeater	0.5063‡	0.5194‡	0.4231‡
Humanities×Repeater	0.2396*	0.2327*	0.1907
Wald tests of group variables (% p-values)			
Reasons of choice	33.3	8.8	8.6
Access grade	44.1	45.9	
Degree types	3.5	2.4	0.3
Short degree × Degree types	43.5	38.5	
Female × Degree types	51.1	55.6	
Repeater × Degree types	9.4	6.9	13.6

We have controlled for the survey year included the corresponding binary variables.
*, † and ‡ denote significance at 20, 10 and 5 percent, respectively.

Table 7
Expected wage for junior college students
Pointwise censored model without selection

Public secondary	−0.0149		
Access grade	−0.0018		
Access at first attempt	−0.4920		
First attempt × Access grade	0.0072		
University father	0.0590		
Science secondary	−0.0145		
Grant	−0.0440		
First choice	−0.1104*	−0.1222*	−0.1058*
Ideal choice	−0.0570		
Same degree	0.0071		
Reason: Economic independence	0.0430		
Reason: Vocation	−0.0711		
Reason: Parental influence	0.0066	−0.0747	
Reason: Difficulty	0.0067		
Female	0.1526	0.1531	−0.0652
Repeater	0.2398*	0.2722†	0.2595†
Satisfied	−0.0144	−0.0075	
Working	0.0699	0.0690	
High family income	0.0260	0.0533	
Short degree	−0.3904‡	−0.3973‡	−0.3284‡
Health	0.0045	0.0011	
Educational	−0.1623	−0.1667	
Social Sciences	0.1251	0.1338	
Humanities	0.1432	0.1387*	
S&E×Short degree	0.4455‡	0.4500‡	0.3688‡
Soc. Sci. ×Short degree	0.2711*	0.2852*	0.1632
S&E×Female	−0.1288	−0.1292	
Soc. Sci. ×Female	−0.3683†	−0.3817†	
Humanities×Female	−0.3039†	−0.2954†	
S&E×Repeater	−0.0088	−0.0227	−0.0441
Soc. Sci. ×Repeater	−0.3088*	−0.3438*	−0.3153*
Humanities×Repeater	−0.4964‡	−0.4998‡	−0.4152‡
Wald tests of group variables (% p-values)			
Reasons of choice	90.1		
Access grade	53.7		
Degree types	19.7	17.3	
Short degree × Degree types	2.4	2.6	2.2
Female × Degree types	18.3	18.5	
Repeater × Degree types	1.1	0.5	0.01

We have controlled for the survey year included the corresponding binary variables.
*, † and ‡ denote significance at 20, 10 and 5 percent, respectively.

Table 8

Monthly average expected wages (in euro) for college students in Madrid by degree type, year and duration of degree and gender

	1st-year, Long degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	4426	2742	3424	2810	1906	3802
	(780)	(715)	(877)	(598)	(602)	(1075)
Female	3815	3200	3326	2416	1487	3178
	(606)	(621)	(983)	(722)	(632)	(1044)
All	4194	3079	3362	2629	1605	3488
	(777)	(660)	(942)	(681)	(641)	(1103)

	1st-year, Short degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	3415	2841	2901		1266	3106
	(924)	(709)	(794)		(12)	(968)
Female	3122	2862	2714		2042	2756
	(825)	(669)	(757)		(545)	(786)
All	3283	2859	2792		1931	2881
	(882)	(661)	(762)		(575)	(868)

	Junior, Long degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2872	3141	2488	2390	2345	2610
	(443)	(365)	(356)	(401)	(549)	(475)
Female	2772	2662	2226	1905	2134	2262
	(342)	(379)	(426)	(578)	(664)	(575)
All	2832	2809	2326	2127	2189	2412
	(404)	(431)	(418)	(558)	(628)	(561)

	Junior, Short degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2683	2161	2214		1626	2492
	(651)	(304)	(161)		(452)	(640)
Female	2549	1844	2065		1246	2001
	(566)	(219)	(112)		(536)	(535)
All	2632	1872	2115		1350	2185
	(617)	(240)	(145)		(523)	(622)

Source: Calculated from "Young people facing college education", 2001, 2004 and 2005.
Standard deviations in parentheses.

Table 9

Percentage difference between average expected wages and average wages observed for working graduates in Madrid

by degree type and duration, degree year and gender

1st-year, Long degree						
	S&E	Health	Soc. Sci,	Humanities	Educational	All
Male	54.2	30.3	5.6	29.4	46.9	41.7
Female	60.2	84.3	36.3	24.6	11.9	50.3
All	54.4	63.8	20.6	28.1	22.0	44.6

1st-year, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	41.8	80.0	56.6		-4.1	38.0
Female	44.6	66.5	105.2		68.8	63.4
All	39.5	68.5	80.8		55.5	43.7

Junior, Long degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	0.1	49.3	-23.3	10.0	80.7	-2.7
Female	16.4	53.3	-8.8	-1.8	60.6	7.0
All	4.3	49.5	-16.6	3.7	66.5	0.0

Junior, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	11.4	36.9	19.5		23.1	10.8
Female	18.0	7.3	56.1		3.0	18.7
All	11.8	10.4	36.9		8.7	9.0

Source: Own calculations from "Young people facing college education", 2001, 2004 and 2005 and "National Survey of Wage Structure", 2002.

Table 10
Percentage difference between average expected wages
observed for working graduates in Madrid by characteristics

	1st-year, Long degree					
	S&E	Health	Soc. Sci,	Humanities	Educational	All
Repeater / Non repeater	4.0	*-29.6*	19.9	-2.3	*-23.1*	7.8
High income / Non high income	16.8	n.a.	-2.2	8.0	*53.5*	2.2
Univ. father / Non univ. father	24.2	36.5	11.9	24.4	27.7	24.2

	1st-year, Short degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Repeater / Non repeater	28.1	*2.8*	9.6		*-18.5*	20.5
High income / Non high income	0.3	*41.6*	*-7.2*		*28.6*	10.9
Univ. father / Non univ. father	12.4	30.9	19.5		0.3	17.3

	Junior, Long degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Repeater / Non repeater	26.9	*33.8*	6.6	-16.3	*62.5*	18.1
High income / Non high income	-1.3	*18.5*	8.2	28.4	*22.9*	6.9
Univ. father / Non univ. father	1.4	-1.8	1.1	-8.8	2.6	0.8

	Junior, Short degree					
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Repeater / Non repeater	6.6	30.8	*-6.7*		*87.9*	27.9
High income / Non high income	1.1	*-8.3*	*4.8*		*49.8*	11.0
Univ. father / Non univ. father	1.2	-1.6	2.4		*12.5*	9.3

Source: Own calculations from "Young people facing college education", 2001, 2004 and 2005.
Those ratios with less than 5 observations in either group are in *italics*.