

Wage expectations for higher education students in Spain*

César Alonso-Borrego[†]
Universidad Carlos III de Madrid

Antonio Romero-Medina
Universidad Carlos III de Madrid

May 5, 2011

We assess the ability of students to forecast future earnings, by using data on expected wages self-reported by college students with different graduation horizons. We find a significant gender effect, by which wages expectations are systematically lower for women than for men, this effect being different with different degree type in the case of junior students. However, the gender gap in actual wages exceeds the gender gap in wage expectations, suggesting that women do not fully account for the gender gap in their future earnings. Also, junior students with low college performance unrealistically predict higher expected wages. The comparison between students' wage expectations and actual wages for college graduates indicates that college students tend to over predict future wages, and overprediction is higher for women and for fresher students.

Keywords: Wage differentials, College choice, Ordered response.

JEL Nos.: I23, J24, J31, C24, C25.

*We are in debt with Juanjo Dolado and Ricardo Mora for valuable comments. 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. The purpose of this study is to compare earnings expectations of college students with the earnings reported by recent colleges graduates. This exercise will allow us to determine whether specific individual characteristics can be identified as potential sources of error in expectation formation process.

Furthermore, the availability of subjective expectations allows to learn about the process of expectations formation, and to improve our understanding of individual behavior (Dominitz, 1998). Our main findings are that both gender and past academic performance play a key role in the ability to form wage expectations. Furthermore, it is rather reassuring that student predictions become more realistic as they approach graduation.

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, who report the wage they expect to earn after completing college. Comparing data on wage expectations with average actual wages for young college graduates, we can assess to what extent self-reported measures of expected wages are realistic, and how they depend on their graduation horizon. As far as the optimal education level and the degree choice is linked to the market returns to education (Betts, 1996), the results may shed light on how education choices are ruled.

Among the previous contributions, we should mention Dominitz (1998), Das and van Soest (1999), and, more recently, Webbink and Hartog (2004), Jerrim (2009) and Delaney et al. (2010). 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. Along the same path, Carvajal et al. (2000) found that in general students expectations are in line with the salaries of recent graduates.

We have used in our analysis demographic individual data, character trait information as well as life ambition data. The main data for this study was gathered in Madrid from 2001 to 2005 among the students in the different public universities of the region. We also considered, as a benchmark, wages

by gender for young college graduated employees, between 25 and 35 years old, working in the Madrid region, from the 2002 Survey of Wage Structure, carried out by the National Institute of Statistics for comparable individuals. The main data offers two advantages to our purpose. First, we have students following different degrees, where the same degree can be followed in different colleges, which allows for broader comparisons. Second, all the colleges belong to a single region, which is a largely populated and a economically dynamic region which attracts workers from all over the country. Thus, we can assume that they have the regional markets as their implicit target.

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.

Provided that earnings expectations accurately estimate future earnings, we can argue, following Chevalier (2007), that they pose the advantage of being free for some types of endogeneity. In particular, individuals already working and reporting low earnings might claim to be less financially motivated than similar individuals with higher earnings.

We find that expected wages for first-year students are affected not only by the gender and poor academic performance (before college and in college), but also by the variables determining their degree type, degree preferences and 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 poor academic performance plays a relevant role in determining expected wages. However, 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 for most junior students, except those following social science degrees. In any case, 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 answer 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. It also provides specific information about the characteristics of the degree chosen by the student, as well as on the choice motivation and the alternative degrees considered. Finally, we have some information related to individual performance. In Table 2, we provide the statistics of the main variables.

Regarding parents’ education, we find that both parents’ levels of education are highly correlated, what makes us to concentrate on the educational level of the father; in particular, whether the father has a university degree (University father). The percentage of respondents whose father achieved a university degree amounts to 41 percent of the sample.

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. Since all sample individuals are actually college students, their minimum college 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.

We explore the characteristics of the degree chosen by the student, as well as on the choice motivation and the alternative degrees considered. We have broken down the degrees by duration (short or long), and into five categories: Science & Engineering (S&E hereinafter), Health, Social Sciences (which also includes Law), Humanities, and Education. The students must fill a common application form, which consists on a prioritized list of alternative colleges. We consider information on whether the first three choices of the respondent featured a particular degree that could be chosen in several colleges (Same degree). About 15 percent of students considered the same degree offered in different colleges. The attended degree was the first choice for approximately 60 percent of students. Long degrees clearly dominate, amounting to 80 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. Regarding choice motivation, vocation and economic independence feature among the main reasons offered by 70 percent of students. Other reasons such as family influence and the difficulty of the degree were alleged by 40 percent of students.

Two features are particularly relevant for our analysis: performance and gender. Student performance will be determinant in the ability to forecast income after graduation. 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.

We find different patterns in the variables under study by gender. Regarding family characteristics, the percentage of students belonging to high-income households is clearly lower for females than for males. Regarding degree types, S&E are dominated by men, whereas Health and Education are dominated by women. Regarding individual characteristics and performance, men and women differ in terms of academic performance. In particular, a higher percentage of women passed the access examination at their first attempt, and follow college degrees that were their first choice. Women also seem to perform better at college, with a higher proportion of grants awarded,¹ and a lower proportion of individuals who are forced to repeat one academic year. Finally, women report to be satisfied in college to a higher

¹It must be noted that grants are awarded for economic reasons, but certain requirements on academic performance must be accomplished.

extent than men.

2.2 Complementary data

In order to assess the quality of wage forecast by college students, we make use of actual wages for comparable individuals as a benchmark. For this purpose, we consider wages by gender for young college graduated employees, between 25 and 35 years old, working in the Madrid region, from the 2002 Survey of Wage Structure, carried out by the National Institute of Statistics (INE hereafter).

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.

²We 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.

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 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,

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

where $E_k(\bullet)$ represents the mathematical expectation, conditional on her information set, and α_k^e , Φ_k^e , $\theta_{1k}^{e'}$ 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 $\alpha_k^e + \Phi_k^e S + \theta_{1k}^{e'} \mathbf{Z}$.

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. 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.

3.2 Empirical specification

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. Also, since we will consider estimation for students in a particular course, the term will be part of the constant term. Therefore, we can write our empirical model as:

$$\ln W_i^* = \beta' \mathbf{X}_i + u_i. \quad (2)$$

As we have mentioned in Section 2, we do not fully observe W_i^* , but a discretized version of it, W_i ,

³Betts (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.

that can be defined as:

$$W_i = j, \text{ if } \mu_{j-1} < W_i^* < \mu_j \quad (j = 1, \dots, 5), \quad (3)$$

where the values μ_j , $j = 1, \dots, 5$ are known. Our estimation approach follows 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). In particular, our model is a pointwise censored model, since the thresholds determining the different categories are known, so there is no need to estimate them as parameters.⁴The model can be estimated by Maximum likelihood after assuming a distribution for the unobservables. Besides, our knowledge of the thresholds allows to obtain projections for expected wages as in a standard linear model.

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. We have implemented separate maximum likelihood estimates for first-year and junior students. The corresponding results are shown in Table 6.

We have considered two different specifications. The first one corresponds to the more general model, including all the selected variables, and the second 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, on the one hand, and being a repeater, on the other. We have reported the p-values of the corresponding tests. The goodness-of-fit of the estimated specifications is

⁴An 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 a standard 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.

reasonably good. We will concentrate our comments on the results regarding our preferred specifications.

4.1 Determinants of expected wages

We first focus on the effects in wage expectations of two factors: gender and academic performance. We will discuss later the role of further variables in expectation formation.

4.1.1 The role of gender

The gender effect is significantly negative for first year students, and its magnitude does not depend on the type of the degree. Hence, women realistically expect lower wages than men with similar characteristics. The fact that women present higher college attendance and better academic performance, as well as a greater expected probability of college graduation, is consistent with this result. Such characteristics propitiate a greater effort in information gathering, which results in lower overprediction of future wages. Nevertheless, the estimated magnitude does not offset the positive gender gap which is observed between men and women. This result suggests that women do not fully account for the gender gap in their future earnings.

Remarkably, unlike first-year students, the gender effect differs by degree types in the case of junior students. These differential gender effects across degrees are jointly significant. In particular, Science & Engineering students of any gender expect similar wages after completing college studies. However, women studying Health degrees expect higher wages than men. These results contrast with the negative gap against women in actual wages for S&E and Health graduates in Table 4. For the remaining degree types, there is a negative gap in wage expectations against women, very specially for Education degrees. Interestingly, the lowest gender gap in actual wages happens for Education graduates, yet female Education students appear as the most pessimistic about their future wage prospects. It is worth noticing that in Spain the educational sector is highly feminized and mostly publicly owned. This might explain both the fact that salaries are similar between men and women and the fact that are significantly lower than the average.

4.1.2 The role of under-performance

We have also controlled for under-performance at college through the variable Repeater, for which we have allowed for interactions with the degree type. In the case of fresher students, being a repeater does not affect wage expectations after completion. We can thus conclude that under-performing at a early stage of college studies does not alter expectation formation.

Besides, we also find that fresher 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 measures of college performance, like Repeater. Actually, the correlation between Satisfied and Repeater is below 6 percent and clearly non significant. The effect of this variable, though, is negligible in the case of experienced students.

The role of under-performance, as measured by the variable Repeater, is relevant, and differs across degree types. In line with Webbink and Hartog (2004), we find that, in S&E, Health and Education degrees, repeaters expect higher wages than non repeaters. We find, though, the opposite result for Humanities students, and no differences between repeaters and non repeaters in Social Sciences degrees.

4.1.3 Other factors

In addition to gender and performance, we have also controlled for other factors. Most of them appear irrelevant for junior students.

Pre-university background, such as the grade achieved in the exam to access college, barely affects student expected wages. As an exception, having followed a Science curriculum in high school has a negative and significant effect on wage expectations for fresher students. Regarding family background, living in a high income family does not affect expectations. However, the fact that the father has a university degree tends to increase wage expectations for fresher students, but it is irrelevant for more experienced students.

In the case of junior students, neither pre-university nor household variables exhibit any effect on expected wages. These variables might have affected the choice of college degree undertaken, which can be related with the ability to predict expected wages shortly after the students have chosen and started their college degree. However, such variables do not affect expectations formation by junior students, to the extent that their degree completion is close and their (past) degree choice is distant. The only

exception is the variable indicating whether the degree undertaken was the first student's choice, with significant negative effect. This result suggests that students who succeeded in entering their most preferred degree tend to expect lower wages.

With respect to the features behind the choice of college degree, we find different patterns among first-year and junior students. The fact that the student consistently applied to the same degree in different colleges exhibit a negative and significant effect. We interpret this result as that vocational students, who show strong preferences for a specific degree, tend to expect lower wages than those showing a higher taste for degree diversity.

We also find that the degree speciality, controlled by binary variables that are individually and jointly significant, do matter on expectation formation by first year students. We observe that first-year students attending degrees different than S&E (the reference group) expect lower wages, specially in the case of Humanities and, mostly, Educational degrees. This results resembles the evidence reported in Table 4 about actual wages, by which college graduated in these latter degrees exhibit the lowest average actual wages. The results are substantially different, though, for junior students, for which, as mentioned earlier, differences across degrees for female and repeaters appear. Also, the degree duration (as captured by the variable Short degree) affects positively wage expectations, but it is only significant for junior students.

4.2 Wage forecasts

We use our preferred estimates from Table 6 to predict student mean expected wages by gender and degree type in Table 7. We have also produced, in Table 8, the percentage difference between mean wage expectations and the mean actual wages for young college graduated employees in Madrid (reported in Table 4). The comparison with the mean actual wages indicates that first-year students tend to overestimate their potential wages to a great extent.⁵ The rankings of expected and actual wages by degree and gender, though, appear 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.

The level of overestimation is generally lower in the case of junior students, suggesting that students become more realistic as they approach graduation. This pattern has been pointed out by Betts (1996)

⁵The only exception is for male students in short Educational degrees, for which the mean expected wage is in line with the mean actual wage.

and Smith and Powell (1990). They reach the conclusion that, in their case, college seniors may have more accurate information and form more realistic expectations. 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, for which mean expected wages are much lower. There are, however, two degrees that exhibit a different pattern: Education and Social Sciences. In Education degrees, junior students overpredict even more than first-year students. In the case of long Social Sciences degrees, we find a negative gap between expected and actual wages.

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. However, for each degree and duration, the gap between expected and actual wages by gender are narrower when students are closer to graduation.

In fact, the differences in mean actual wages by gender and degree are higher than the corresponding differences in 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. As extreme case, long Educational degrees exhibit the highest gap between mean expected and actual wages.

To understand these results, it must be noted that the actual wages by degree that we use as reference correspond to young college graduates 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.

For first-year and junior students, the percentage difference between mean expected wages by degree type and duration, in accordance with being or not a Repeater is presented in Table 9. It must be noted that some cells have been computed with very few observations. We find that being a repeater is a source of distortion, to the extent that junior repeaters, in comparison with their remaining peers, still heavily overpredict. This result is in accordance with those in Webbink and Hartog (2004), by which repeaters wrongly tend to be more optimistic than non repeaters.

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. Junior students are less affected by pre-university or family background variables when computing their expected wages. Mostly, the type and the duration of the degree are the major variables affecting wage expectations of junior students. But also gender and college performance play a relevant role in their expectation formation. In the case of gender, female students realistically predict lower wages than their male counterparts. In the case of performance at college, repeaters tend to produce wage predictions of lower quality.

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 as graduates. 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. In this sense, 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.

The negative gap between expected and actual wages in Social Sciences deserves special attention. It suggests that very pessimistic expectations by such students in these degrees. 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.

5 Conclusions

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.

We have found that the degree type has a relevant role on expected wages. There were also differences on how students produce their wage expectations depending on their degree year. 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 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 relevant. Women in their first-year expect, other things equal, lower salaries than their male counterparts, and this gender gap is independent of the type of degree. Such gender differences persist for more experienced students, but differ by degree type. This evidence is consistent with the fact that women systematically earn lower mean wages than men. Also, college performance appears relevant in wage predictions for junior students: repeaters tend to overpredict wages to a larger extent than non repeaters. We interpret this result as those student with worse college curriculum also yield wage predictions of lower quality.

We also find that students with gross underperformance have a distinct pattern under and over

forecasting their future earnings in comparison with their peers. The existence of a clearly differentiated pattern might signal both a different expected career path or the presence of different information processing abilities.

Shortly after high school graduation, college students systematically over predict 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, on average, close 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, although still above.

We interpret over prediction, in line with Carvajal et al. (2000), as that students' expectations do not conform to market realities, to the extent that they usually expect to hold better job positions than what they actually will hold. Our paper offers information on the relevant characteristics that describe those students more in need of adequate counseling by the placement offices, in order to focus unrealistic expectations that might hinder the transition of college graduates into the labor force.

References

- [1] Alonso-Borrego, C., R. Mora, 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., 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, 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] Carvajal, M.J., D. Bendana, A. Bozorgmanesh, M.A. Castillo, K. Pourmasiha, P. Rao, J.A. Torres, (2000), Inter-gender differentials between college students' earnings expectations and the experience of recent graduates. *Economics of Education Review* 19: 229–243.
- [7] Chevalier A., (2007) Education, Occupation and Career Expectations: Determinants of the Gender Pay Gap for UK Graduates. *Oxford Bulletin of Economics and Statistics* 69(6):819- 842.
- [8] Das, M., A. van Soest, (1999), A panel data model for subjective information on household income growth. *Journal of Economic Behaviour and Organization* 40: 409-426.
- [9] Delaney, L., C. Harmon, C. Redmond, (2010), Decomposing Gender Differences in College Student Earnings Expectations. University College Dublin Geary Institute Discussion Paper Series. Dublin, Ireland.
- [10] Dominitz, J., (1998), Earnings expectations, revisions and realizations. *The Review of Economics and Statistics* 80 (3): 374-388.
- [11] Dominitz, J., C.F. Manski, (1996), Eliciting student expectations of the returns to schooling. *The Journal of Human Resources* 31(1), 1-26.
- [12] Jerrim, J., (2009), Wage Expectations of UK Students: Are They Realistic? Southampton Statistical Sciences Research Institute. Mimeo.
- [13] Heckman, J. J., (1979), Sample Selection Bias as a Specification Error. *Econometrica* 47(1), 153-161.
- [14] Miranda, A., 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.
- [15] Papke, L.E., (1998), How are participants directing their participant-directed individual account pension plans? *American Economic Review* 88: 212–216.
- [16] Smith, H.L., B. Powell, (1990), Great expectations: variations in income expectations among college seniors. *Sociology of Education*, 63 (3): 194–207.
- [17] Webbink, D., 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

Variable	All		Female		Male	
	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 (1396)	2170 (1172)	2476 (2147)	1938 (1095)	1136 (867)	2228 (1564)
Female	2208 (1074)	1866 (915)	2001 (1482)	1744 (901)	1252 (730)	1787 (1151)
All	2488 (1324)	1983 (1032)	2218 (1832)	1831 (996)	1289 (7944)	2015 (1397)
National average, Short degree						
	S&E	Health	Soc. Sci.		Educational	All
Male	2130 (999)	1829 (660)	1663 (1029)		1379 (658)	1959 (974)
Female	1899 (822)	1701 (553)	1279 (721)		1326 (607)	1551 (682)
All	2088 (973)	1724 (576)	1410 (857)		1342 (623)	1754 (864)
Madrid average, Long degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2870 (1549)	2104 (1126)	3243 (2940)	2172 (1137)	1298 (1128)	2682 (2023)
Female	2381 (1139)	1737 (861)	2440 (1816)	1939 (899)	1329 (812)	2114 (1418)
All	2716 (1449)	1879 (984)	2788 (2399)	2052 (1022)	1315 (963)	2413 (1784)
Madrid average, Short degree						
	S&E	Health	Soc. Sci.		Educational	All
Male	2408 (982)	1578 (583)	1852 (1413)		1320 (766)	2250 (1044)
Female	2160 (757)	1718 (468)	1323 (743)		1210 (598)	1686 (724)
All	2354 (942)	1696 (487)	1545 (1095)		1696 (487)	2005 (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

Determinants of expected wages for college students in Madrid
Pointwise censored model without selection

	First-year students				Junior students			
	(i)		(ii)		(iii)		(iv)	
Gender	-0.207 [†]	(0.091)	-0.117 [†]	(0.059)	-0.004	(0.115)	-0.001	(0.114)
Repeater	0.033	(0.087)	0.062	(0.058)	0.249 [†]	(0.118)	0.251 [†]	(0.115)
Health×Female	0.155	(0.221)			0.316	(0.252)	0.325	(0.259)
Education×Female	0.012	(0.134)			-0.276	(0.169)	-0.310 [†]	(0.155)
Soc. Sci. ×Female	0.207	(0.154)			-0.240	(0.170)	-0.241	(0.175)
Humanities×Female	0.107	(0.132)			-0.169	(0.145)	-0.162	(0.144)
Health×Repeater	-0.196	(0.202)			0.082	(0.251)	0.111	(0.238)
Education×Repeater	-0.237	(0.215)			0.070	(0.201)	0.069	(0.198)
Soc. Sci. ×Repeater	0.246	(0.171)			-0.325	(0.216)	-0.321	(0.212)
Humanities×Repeater	0.012	(0.133)			-0.508 [†]	(0.160)	-0.495 [†]	(0.150)
Health	-0.254	(0.209)	-0.206 [†]	(0.090)	-0.402	(0.252)	-0.436*	(0.255)
Education	-0.398 [†]	(0.127)	-0.475 [†]	(0.120)	-0.063	(0.114)	-0.053	(0.109)
Soc. Sci.	-0.288 [†]	(0.134)	-0.183*	(0.100)	0.052	(0.155)	0.048	(0.145)
Humanities	-0.285 [†]	(0.119)	-0.290 [†]	(0.099)	0.096	(0.121)	0.077	(0.109)
Public secondary	-0.019	(0.054)			0.001	(0.064)		
Access grade	0.001	(0.003)			0.003	(0.004)		
Access at first attempt	0.039	(0.075)			-0.031	(0.068)		
University father	0.113 [†]	(0.056)	0.130 [†]	(0.055)	0.053	(0.064)		
Science secondary	-0.205 [†]	(0.079)	-0.223 [†]	(0.079)	0.006	(0.084)		
Grant	-0.018	(0.068)			-0.063	(0.084)		
First choice	-0.058	(0.056)	-0.052	(0.052)	-0.144*	(0.074)	-0.142 [†]	(0.072)
Same degree	-0.146*	(0.077)	-0.135*	(0.074)	0.010	(0.087)		
Economic indep.	0.047	(0.063)			0.034	(0.080)		
Parental influence	0.039	(0.060)			-0.098	(0.071)	-0.087	(0.068)
Vocation	-0.030	(0.069)			-0.001	(0.100)		
Difficulty	-0.110*	(0.061)	-0.096	(0.059)	0.030	(0.074)		
Satisfied	0.124*	(0.066)	0.101	(0.065)	0.006	(0.092)		
High family income	0.101	(0.067)	0.097	(0.065)	-0.06	(0.076)		
Short degree	-0.079	(0.069)	-0.091	(0.067)	-0.134*	(0.077)	-0.134*	(0.072)
Wald tests of group variables (% p-values)								
Reasons of choice	34.6				72.5			
Access grade	79.4				74.3			
Degree types	2.0		0.1		25.1		21.5	
Female × Degree types	65.8				15.8			
Repeater × Degree types	27.9				0.4			

We have controlled for the survey year included the corresponding binary variables. Standard errors in parentheses.

* and [†] denote significance at 10 and 5 percent, respectively.

Table 7

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	4433 (807)	2878 (634)	3393 (830)	2790 (570)	1861 (520)	3798 (1082)
Female	3821 (629)	3113 (612)	3324 (938)	2430 (722)	1458 (607)	3173 (1037)
All	4201 (800)	3051 (609)	3350 (896)	2625 (666)	1571 (602)	3484 (1103)
1st-year, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	3416 (966)	2839 (464)	2752 (643)		1556 (39)	3083 (936)
Female	3114 (826)	2853 (786)	2776 (791)		2040 (595)	2763 (833)
All	3280 (907)	2851 (742)	2766 (718)		1971 (575)	2878 (881)
Junior, Long degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2971 (476)	2278 (390)	2709 (405)	2567 (448)	2301 (517)	2682 (488)
Female	3124 (461)	2720 (569)	2049 (385)	1919 (555)	1713 (534)	2224 (673)
All	3033 (471)	2584 (553)	2302 (506)	2215 (601)	1868 (580)	2421 (641)
Junior, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	2392 (525)	1834 (265)	2469 (205)		1940 (453)	2333 (493)
Female	2452 (515)	2289 (371)	1923 (182)		1185 (471)	2131 (550)
All	2415 (516)	2247 (383)	2105 (322)		1391 (566)	2207 (536)

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

Table 8

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	54.5	36.8	4.6	28.4	43.4	41.6
Female	60.5	79.3	36.2	25.3	9.7	50.1
All	54.7	62.4	20.2	27.9	19.5	44.4
1st-year, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	41.9	79.9	48.6		17.8	37.0
Female	44.2	66.0	109.9		68.6	63.9
All	39.4	68.1	79.1		58.7	43.5
Junior, Long degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	3.5	8.2	-16.5	18.2	77.3	0.0
Female	31.2	56.6	-16.0	-1.0	28.9	5.2
All	11.7	37.5	-17.4	7.9	42.0	0.3
Junior, Short degree						
	S&E	Health	Soc. Sci.	Humanities	Educational	All
Male	-0.7	16.2	33.3		46.9	3.7
Female	15.5	33.2	45.4		-2.0	26.4
All	2.6	32.5	36.3		12.0	10.1

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

Table 9

Percentage difference in average expected wages for working graduates
in Madrid between repeaters and non repeaters by degree type and duration

	S&E	Health	Soc. Sci,	Humanities	Educational	All
1st-year, Long degree	6.9	<i>-10.5</i>	-0.9	<i>-0.3</i>	<i>-2.2</i>	6.1
1st-year, Short degree	30.8	<i>25.7</i>	-11.5		2.1	22.6
Junior, Long degree	29.0	<i>27.1</i>	9.2	-20.1	<i>64.0</i>	20.5
Junior, Short degree	9.1	<i>35.4</i>	<i>-6.4</i>		<i>67.6</i>	21.9

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*.