# The Effect of Labor Market Information on Community College Students' Major Choice

**RACHEL BAKER**

UNIVERSITY OF CALIFORNIA IRVINE

**ERIC BETTINGER**

STANFORD UNIVERSITY & NBER

**BRIAN JACOB**

UNIVERSITY OF MICHIGAN & NBER

**IOANA MARINESCU**

UNIVERSITY OF CHICAGO & NBER

**Abstract**

An important goal of community colleges is to prepare students for the labor market. But are students aware of the labor market outcomes in different majors? And how much do students weigh labor market outcomes when choosing a major? In this study we find that less than 40% of community college students in California rank broad categories of majors accurately in terms of labor market outcomes. However, students believe that salaries are 13 percent higher than they actually are, on average, and students underestimate the probability of being employed by almost 25 percent. We find that the main determinants of major choice are beliefs about course enjoyment and grades, but expected labor market outcomes also matter. Experimental estimates of the impact of expected labor market outcomes are larger than OLS estimates and show that a 1% increase in salary is associated with a 1.4 to 1.8% increase in the probability of choosing a given broad category of majors.

## Introduction

Community colleges students represent over 40% of total undergraduate enrollment in the United States (Ma and Baum, 2015). These students are more likely than their four-year peers to be underrepresented minorities and to come from disadvantaged backgrounds (Bailey, Jenkins & Leinbach, 2005). An explicit mission of community colleges is to improve the labor market outcomes of these disadvantaged student populations (American Association of Community Colleges; California Community Colleges). Much attention, at the federal, state, and institutional level, has recently been paid to the labor market outcomes of community college graduates from these career-oriented programs (see, for example, Jepsen, Troske & Coomes, 2014; Dadgar & Trimble, 2014; Xu & Trimble, 2015).

Labor market outcomes associated with various degrees and certificates vary dramatically. Many awards are associated with significant positive increases in earnings and probability of employment, while others have been shown to have insignificant effects on earnings (Stevens et al, 2015; Bahr, 2015). This means that, for community colleges to best fulfill their mission of improving the labor market outcomes of their students, students need to be aware of differences in outcomes across awards. The degree to which this is true is not clear.

The existing research on students' knowledge of labor market outcomes has focused on students attending selective four-year colleges, while very little is known about the behavior of community college students. Students in selective four-year colleges have broadly accurate information about the relative rankings of labor market returns associated with various degrees, but many students make large errors in specific estimations (Arcidiacono, Hotz & Kang, 2010; Huntington-Klein, 2015). In particular, lower income and lower GPA students are less well informed (Attanasio and Kaufmann, 2012; Betts, 1996). Given that community college students are disproportionally disadvantaged and have weaker access to informed social networks, it is unclear to what extent previous findings apply to community college students.

Furthermore, just having the information is not enough; it should also used as a basis for decision making. This premise, that expected labor market outcomes should be an important part of the major choice decision for students in community colleges, is foundational in a number of initiatives being undertaken by high-profile national players. The Aspen Institute, Jobs for the Future, the American Association of Community Colleges, various state community college systems, and most notably the U.S. Department of Education's College Scorecard have recently focused efforts on helping students, and schools, use labor market information in making decisions. But these efforts can only succeed if students take labor market outcomes into account when choosing a major.

In this paper, we provide empirical evidence on the potential efficacy of such initiatives. We examine to what extent community college students factor in labor market outcomes when choosing a degree. Studies of students in selective four-year colleges indicate that degrees are mostly chosen for non-pecuniary reasons, with expected salary playing a significant but minor role (Beffy et al, 2012; Wiswall and Zafar, 2015a; Stinebrickner & Stinebrickner, 2014). In this paper we ask if that is also the case for community college students. If community college students face stronger incentives to make ends meet, and if degrees in community colleges have more heterogeneous returns, one may expect labor market outcomes to play a larger role in determining major choice.

First, leveraging unique administrative data on the labor market outcomes of community college students, we examine the accuracy of students' beliefs about the labor market outcomes of different majors for alumni of their college. An important innovation of our study is that we examine beliefs about both wages and employment, while previous literature has focused mostly on wages. Furthermore, we do not ask students the outcomes for a national sample of students, but about labor market outcomes for graduates from their own college, which should be the most directly relevant and perhaps the best known. We find that only 13% of students correctly rank the four broad categories of majors in terms of salary. Even if we ask about the two highest paying vs. the two lowest paying majors, only 40% of students rank the majors correctly.. We also find systematic biases: students overestimate salaries by 13% and underestimate the probability of finding employment by almost 25% in almost all fields. Finally, we find that low income students are more likely than their higher income peers to have large errors in estimating the probability of employment.

Second, we examine the determinants of major choice. An important innovation of our study is to compare the estimates from observational data with the estimates from an experiment. Indeed, with the exception of Hastings et al (2015) and Wiswall and Zafar (2015a,b) who focus respectively on Chilean students and NYU students, most of the previous literature is non-experimental. In our experiment, we provide students with hypothetical information about the labor market outcomes for various majors and elicit their probability of choosing a given major given this new information. In our observational data, expected course enjoyment is the most important predictor of major choice, followed by expected grades. Salary has a positive relationship with major choice, while the effect of the probability of employment is not statistically significant. In our experimental study, we find that salary has a significant effect on major choice, and this effect is quantitatively larger than in the observational data. Based on our experimental estimates, a 1% increase in salary is associated with about a 1.4% (for business) to 1.8% (for humanities) increase in the probability of choosing a major. On the other hand, a 1% increase in the probability of employment is associated with a 0.2 to 0.6% increase in the probability of choosing a major, but this effect is not always statistically significant. Comparing the experimental and non-experimental estimates, we find evidence consistent with students' lack of information dampening the impact of labor market outcomes on major choice in observational data. Overall, these results show that estimates based on observational data are a poor predictor of the impact of labor market information on major choice.

Surprisingly, community college students are not very different from their peers in four-year schools: they also put the most weight on factors other than labor market outcomes, such as course enjoyment, when choosing a major. The preeminence of factors other than labor market outcomes in major choice is partly explained by students' lack of information about these labor market outcomes. Our results suggest that better information about labor market outcomes, as well as increasing the salience of information about labor market outcomes, could serve community colleges' goal of improving their students' labor market prospects.

## Background and literature review

College students choose majors in a complicated choice environment; there many options (sometimes upward of 150) that vary along a number of dimensions, including required courses, expected rigor, and perceived labor market outcomes. Examining how students make this decision has important consequences, as the choice of a major can have implications for future earnings and employment (Wolniak et al, 2008), students often receive little direct guidance on how to make such a decision, and most schools offer relatively little structure to direct this decision making process (Scott-Clayton, 2011; Grubb, 2002). Many students, particularly those in broad-access two- and four-year schools, report feeling anxious or confused about how to best choose a major (Rosenbaum et al, 2006; Scott-Clayton, 2011).

A number of studies have attempted to estimate the effects of various factors that students use to choose a major. Much prior work is based on traditional human capital theory, which posits that students will make educational decisions (how long to persist, what degree to pursue) based on a calculation of the perceived costs relative to the perceived discounted benefits (Becker, 1962). Costs and benefits can be both financial and psychological.

Prior research has found that non-pecuniary factors, such as perceived enjoyment and ability, are the most important factors in choosing a major. Expected enjoyment is the biggest driver of choice (Beffy et al, 2012; Wiswall and Zafar, 2015a; Stinebrickner & Stinebrickner, 2014). Enjoyment can mean many things: a sense of satisfaction both from courses in the program and from the expected career, a belief that a certain field will allow them to fulfill altruistic goals (Grubb, 2002), or hope that a given field will help students gain the approval of their parents (Zafar, 2009). A student's perception of his/her ability in a given field is also an important consideration (Arcidiacono et al, 2010; Stinebrickner & Stinebrickner, 2011). Zafar (2009) finds that such non-pecuniary outcomes explain about 45% of choice behavior for males and 75% for females.

Expected labor market outcomes (expected salary, expected probability of finding a job, etc.) play a role in the choice of a major, though these factors are generally less important than expected enjoyment and ability. The effect of expected earnings on major choice is small but significant. Wiswall and Zafar (2015a) show that a 1% increase in beliefs about self earnings in a major increases the log odds of graduating in that major by 2%. However, beliefs about self earnings are correlated with unobserved reasons why students prefer a certain major: when using experimental variation in beliefs about self earnings, a 1%

increase in beliefs about earnings in a given major only increases the log odds ratio of graduating in that major by 0.28%.

However, using estimated labor market returns to make choices about majors is a good strategy only so far as students have access to accurate information. Previous research has examined students' expectations about labor market outcomes, particularly students' beliefs about wages. Recent empirical work indicates that while college students can generally accurately rank the returns to various major categories and levels of awards (Arcidiacono, Hotz & Kang, 2010; Huntington-Klein, 2015), many students make significant errors in estimating returns to various degrees (Betts, 1996; Arcidiacono, Hotz & Kang, 2010). In addition, there is important heterogeneity; students tend to get more accurate over time (students in their fourth year offer more accurate predications than students in their first year) (Zafar, 2011; Betts, 1996), students from lower income families tend to underestimate the returns to schooling (Attanasio and Kaufmann, 2012; Betts, 1996), students with higher GPAs make smaller errors (Betts, 1996), and there is significant variation in accuracy across majors (Wiswall and Zafar, 2015b). However, for certain awards, especially certificates and associate degrees, the labor market benefits might be on the margin of employment or the number of hours worked, rather than on hourly rate. To date, no study has examined the accuracy of students' knowledge labor market returns in terms of both probability and stability of employment.

While there is a literature on major choice and labor market outcomes for four-year college students, this paper is the first to look at major choices for students in community colleges. We argue that it is important to examine this population for four main reasons. First, as compared to students in four-year schools, students in community colleges are more likely to be first generation college students, more likely to be from low-income backgrounds, and more likely to be underrepresented minorities (Bailey, Jenkins & Leinbach, 2005). In addition to consequences for equity and social mobility, this has important implications for the potential effects of information about labor market outcomes. Namely, we hypothesize that students in community colleges may be more sensitive to information gaps, as they have thinner social networks, less access to information, and are more likely to use in-school sources to get information about earnings (Scott-Clayton, 2011; Hastings et al 2015).

Second, on average, community colleges offer many more award options than four-year schools, and these awards are more heterogeneous in both focus and length (purely vocational, academic, short- and long-term certificates, associate degrees). This creates a complicated decision making environment in which students must weigh a number of potential options using a range of criteria (Scott-Clayton, 2011; Rosenbaum, Deil-Amen & Person, 2006 ). In order to improve the labor market outcomes of community college students, it is important to understand how they navigate this complex decision.

Third, community college students have traditionally persisted and graduated at relatively low rates. While the vast majority of students in community colleges have the intention of earning a bachelors degree when they start, very few actually do (Rosenbaum & Rosenbaum, 2013). For this reason, it may be more efficient for many students to aim for a relatively fast transition to the labor market with a one or two year degree or certificate that would enhance their earnings and job opportunities. This high

probability of not completing a bachelors degree creates a different decision making environment than in four-year schools and necessitates conceptualizing the effect of labor market information anew.

Fourth, the training that occurs in community colleges is tied to a wider range of labor market outcomes than four-year training, and returns are more heterogeneous across two-year degrees and certificates than they are across four-year degrees. The average returns to an associate degree or community college certificate are positive (Bailey & Belfield, 2011). However, a number of recent studies have shown that these effects vary greatly by field, with some degrees yielding insignificant earnings returns and others, often degrees in health-related fields, yielding large positive returns (Stevens et al, 2015; Bahr, 2015; Dadgar & Weiss, 2012; Jepsen, Troske & Coomes, 2014; Jacobson, LaLonde and Sullivan, 2005). Additionally, the labor market returns to a community college award go beyond salary; the effects on probability of employment and on hours worked is at least as significant as the effect on wages (Dadgar & Weiss, 2012; Jacobson, LaLonde and Sullivan, 2005; Marcotte et al, 2005).

In this study, we surveyed 376 students at two community colleges in California's Bay Area. We asked students about their knowledge of labor market outcomes, their curricular plans, and how they would respond to specific information about labor market outcomes in different fields of study. We examine how community college students differ from students at selective four-year schools. Importantly, we are the first study to examine a range of labor market outcomes (beyond salary), so we can examine which specific labor market outcomes are most important to students.

## Methodology and Results

Our data come from an original survey administered at two community colleges in California's Bay Area in the winter of 2014. To increase response rates, we administered the survey in person in courses (using paper and pencil). We sent a request to administer the survey to course instructors via the departmental representatives to the faculty senate. Due to the nature of the survey questions, we targeted academic (not strictly pre-vocational) classes with no pre-requisites that enrolled mostly students in their first year.

Our sample of 12 classes consisted of courses that fit our sampling criteria and had instructors who were willing to use 30-40 minutes of class time for the survey. The classes were in a number of disciplines, including psychology, environmental science, language arts and biology. A member of the research team (Rachel Baker) administered the survey in each class and was present to read instructions and answer questions. The 12 classes had a total enrollment of about 470 and daily attendance of about 80%, so we obtained responses from 376 students. In each class, no students who were present opted out of the survey. Students took an average of 25 minutes to complete the survey. We provided an incentive of two $25 Amazon.com gift cards to each class and raffled the gift cards off when all students were finished with the survey. We also gave each participating instructor a $25 Amazon.com gift card.

Due to our sampling strategy, the students in our sample are more transfer- and academically-oriented than the average community college student. Sixty-four percent of the students in our sample stated that their most likely outcome was transfer (as opposed to earning an associates degree or certificate). The students in our sample were also younger than the mean community college student (mean age 23), and

almost a quarter of the students in our sample were in their first term of college. We will return to the implications of these sample characteristics for the interpretation of our results in the discussion section.

Table 1: Student Descriptives presents descriptive information about the students who are included in the sample.

**Table 1: Survey Summary Statistics**

| | mean | sd | N |
|---|---|---|---|
| Age (years) | 23.472 | 7.901 | 369 |
| Self-reported family income is < $50,000 | 0.430 | 0.496 | 349 |
| No parent with college degree | 0.438 | 0.497 | 347 |
| Student is in first term | 0.241 | 0.428 | 357 |
| Self-reported HS GPA is ≥ B | 0.737 | 0.441 | 376 |
| Latino | 0.243 | 0.429 | 342 |
| White | 0.293 | 0.456 | 376 |
| Asian | 0.508 | 0.501 | 376 |
| Black | 0.037 | 0.19 | 376 |
| Pacific Islander | 0.035 | 0.183 | 376 |
| *Student lists as most likely award:* | | | |
| Transfer | 0.489 | 0.501 | 376 |
| AA | 0.074 | 0.263 | 376 |
| Cert | 0.088 | 0.283 | 376 |
| No Degree | 0.101 | 0.302 | 376 |
| Tie or no choice | 0.248 | 0.432 | 376 |
| *Student lists as most likely broad major field:*[†] | | | |
| Science | 0.521 | 0.500 | 376 |
| Business | 0.267 | 0.444 | 376 |
| Language Arts | 0.311 | 0.464 | 376 |
| Humanities | 0.247 | 0.432 | 376 |
| *Expectations about awards:*[#] | | | |
| Expected Salary | $44,689 | $30,361 | 1,131 |
| Exp. Prob. of having a job | 0.435 | 0.24 | 1,223 |
| Exp. Stab. of employment | 0.539 | 0.370 | 1,292 |
| Prob. of enjoyment | 0.539 | 0.430 | 1,316 |
| Predicted grades (1=F, 5=A) | 4.08 | 0.920 | 1,221 |

Notes: [†] Many students stated that they were equally likely of majoring in two or more broad major categories. When students were "most likely" of majoring in two or more broad major fields, we included both. [‡] Students were asked each of these questions for each broad major category, so there are a maximum of three observations per student.

The survey, available in Appendix A, had four main sections: (1) students' self-reported probabilities of choosing each of four broad major categories and their expectations about enjoyment, grades, and labor market outcomes in each of these fields[1]; (2) the same information about the three specific majors that students state they are most likely to choose; (3) experimental questions in which we randomly assign hypothetical labor market information about two major categories to students and ask them how this changes their probability of choosing a given major; and (4) multiple choice questions in which we ask students where they get information about expected salaries for majors and which factors they value when choosing a major. Additionally, we collected background demographic information from each student and linked survey responses to administrative transcript and application data. In this paper, we will be focusing on sections (1) and (3).

To assess students' accuracy in estimating the outcomes associated with various degrees and certificates, we collected data from the California Employment Development Department (EDD) Labor Market Information Division. For 2006-2013 graduates who earned each award offered at the two community colleges in our sample (e.g. AA in Music Technology in 2007, certificate in Network Administration in 2010), we collected data on the probability of employment six months after graduation (proportion of graduates who had positive wages in either of the two quarters after graduation), the mean salary three years after graduation (all adjusted to the 2013 consumer price index), and the stability of employment for graduates in the three years after graduation (number of quarters with positive wages/12 quarters). One important factor that sets this study apart from most previous studies is that we ask students about their expectations for students from their own college and compare their responses to this data (rather than comparing their responses to data on all graduates in the country or to data on students from different areas of study). This is an important improvement on past work, particularly in the context of community colleges, where the local labor market is especially important.

## Descriptive summary results

Seventy-seven percent of students in our sample report that they are considering majors in more than one broad major category. On average, students are considering 2.63 groups with non-zero probability (1/3 of students are considering all four major categories). Many students report 0 or very low probabilities of majoring in Language Arts, and Business & Technology[2] (for each of these two broad

[1] One methodological note is important to include here. On the survey (available in Appendix A), we asked for students' expectations about the labor market outcomes associated with earning an *associates degree* in each field. We did not ask for the labor market outcomes associated with transferring and then earning a *bachelors* degree in each field (in fact, we explicitly told students NOT to answer the question with students who successfully transfer an earn a BA/BS in mind). This, of course, affects the interpretation of our results, but we felt this was the most appropriate course of action for a few reasons. First, estimating labor market outcomes for students who transfer from a two- to four-year college is a function of a number of factors: probability of admission to a four-year school, probability of finishing a four-year school, selectivity and curriculum of both the two- and four-year schools, etc. Second, the time horizon necessary for estimating such returns is necessarily protracted—students will not be entering the labor force for at least four (and likely five or six) years. Finally, while the majority of our students stated that they intended to transfer and earn a bachelors degree, evidence from these schools, and community colleges more generally, indicates that most of these students will not transfer and those who do might not earn a four-year degree. The labor market outcomes associated with earning an associates degree is both simpler to estimate and more plausibly relevant for the students we sampled.

[2] This could be a function of the sample of classes in which we administered the survey.

majors categories, almost half of students state that there is less than a 10% probability that they will select it).

Table 2 : Students' beliefs of labor market outcomes

| | Student Beliefs Across Sample | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Average Salary[a] | | Belief of Salary[b] | | Prob. of Employment[c] | | Belief of Prob. of Employ.[d] | | Percent of Time with a Job[e] | | Belief of Stability of Employ.[f] | |
| | M | Rank | M | Rank | M | Rank | M | Rank | M | Rank | M | Rank |
| Science, Math & Engineering | 46,300 | 2 | 51,919 | 2 | 0.615 | 2 | 0.466 | 2 | 0.611 | 2 | 0.599 | 1 |
| | (27,395) | | (35,501) | | (0.211) | | (0.256) | | (0.196) | | (0.587) | |
| Business, Comp. Sci. & Tech. | 51,473 | 1 | 52,149 | 1 | 0.652 | 1 | 0.498 | 1 | 0.641 | 1 | 0.59 | 2 |
| | (17,885) | | (33,326) | | (0.186) | | (0.235) | | (0.15) | | (0.248) | |
| Arts & Language | 28,820 | 4 | 34,850 | 4 | 0.485 | 4 | 0.369 | 4 | 0.473 | 4 | 0.462 | 4 |
| | (13,635) | | (19,480) | | (0.135) | | (0.221) | | (0.131) | | (0.246) | |
| Humanities & Social Sciences | 33,180 | 3 | 39,702 | 3 | 0.537 | 3 | 0.406 | 3 | 0.529 | 3 | 0.503 | 3 |
| | (12,005) | | (26,629) | | (0.148) | | (0.228) | | (0.143) | | (0.251) | |
| % of students who ranked correctly | | | 15% | | 12% | | | | 9% | | | |
| % of students who correctly ranked top two and bottom two | | | 43% | | 35% | | | | 32% | | | |

Notes: All labor market information comes from the California Employment Development Department, Labor Market Information Division and represents 2006-2010 graduates from two community colleges in Northern California. [ace] Mean annual salary of graduates three years after graduation, in 2013 dollars; percent of graduates who are employed six months after graduation; fraction of quarters that graduates have non-zero earnings in the three years after graduation. [bdf] Students were asked: Average annual salary a recent graduate from CCD with an AA/AS in this field (who doesn't continue to a BA/BS) will make 3 years after graduation; probability a graduate will have a job six months after graduation; percent of time a graduate will be with a job in the first three years after graduation.

In Table 2, we see that students in our sample have broadly accurate views of the average labor market outcomes of graduates in various fields. In the aggregate, they generally correctly rank fields in all outcomes (salary, probability of employment, and stability of employment), which is similar to Zafar's (2012) finding for students knowledge of relative salaries. However, we find that individual students rarely rank all four fields correctly; only 13% of students in our sample were able to correctly rank the expected salaries.[3]

We find that students generally overestimate expected salaries, which matches what Hastings et al (2015) find with Chilean students, and is different from what Zafar (2011) found at a selective U.S. college. The only field in which they don't overestimate is Business, Computer Science & Technology. Students overestimate the most (in % terms) the salaries in Arts & Languages.[4] In all fields, students underestimate the probability of being employed. The degree of underestimation is remarkably consistent across fields—on average, students think the probability of being employed is about 75% of what it actually is. Students also underestimate the average stability of employment (percentage of the time with an income), but not nearly as much as they underestimate the probability of being employed. Overall, students overestimate salaries and underestimate the employment outcomes of awards.

There is a relatively wide distribution of students' beliefs about their own outcomes as compared to others' outcomes. Students' expected outcomes for themselves are higher than the population on average (10% higher for probability of having and stability of a job, 4% higher for salary), but there is a wide range, as shown by Figure 1, which plots the distribution of the difference between own outcomes beliefs and other students outcomes beliefs.

[3] However, as shown in Table 2, the average salaries of students who graduated with Science, Math & Engineering degrees is quite close to the average salaries of students who graduated with Business, Computer Science & Technology degrees. Likewise, the average salaries of the other two broad major fields are quite close. If we relax our definition of correct ranking, and measure what percent of students correctly state that the expected salaries from Science, Math & Engineering and Business, Computer Science & Technology degrees are higher than from Arts & Language and Humanities & Social Science degrees, we see that 40% of students correctly rank the four expected salaries.

[4] One methodological note is important here. Our measure of the "true" salary is restricted to those graduates who have positive earnings in a given quarter, and thus included graduates who had positive earnings but did not work full-time for the whole quarter. Students in our survey may have been thinking of someone who is employed full-time for the full year, which could explain why they overestimate salary.




Figure 1: Expectations for Self vs. Others

## Errors in beliefs about labor market outcomes

To examine students' errors in beliefs about labor market outcomes we used a similar measure as Zafar (2011) and Betts (1996):

$$error_{if} = \frac{s_{if} - s_{obs,f}}{s_{obs,f}} * 100$$

Where $s_{if}$ is student *i*'s estimate of the labor market outcome of field *f* and $s_{obs,f}$ is the observed labor market outcome for field *f*. This measure represents the percent error in the student's belief relative to the actual outcome: for example, a value of 10 for salary error means that the student believes the salary to be 10% higher than it actually is. We constructed two measures for the error in beliefs: (1) error measure defined above, (2) absolute value of the error measure defined above. Since errors can be both positive and negative, a mean error close to zero does not necessarily indicate uniformly low errors. Reporting the absolute value of errors in addition to the mean error allows us to examine a measure of the dispersion of errors.

Table 3: Student Errors in Predicting Labor Market Outcomes

| | | | Median by Major | | | |
|---|---|---|---|---|---|---|
| | Mean | Median | Business | Humanities | Language & Arts | Science |
| | | | *Error* | | | |
| Salary | 13.5 | 4.1 | -2.9 | 5.5 | 4.1 | 8.0 |
| Probability of Employment | -24.0 | -23.3 | -23.3 | -25.5 | -38.1 | -18.8 |
| Stability of Employment | -4.3 | -5.5 | -6.4 | -5.5 | 5.6 | -18.2 |
| | | | *Absolute Error* | | | |
| Salary | 47.0 | 35.2 | 32 | 35.6 | 38.8 | 35.2 |
| Probability of Employment | 40.0 | 38.1 | 23.3 | 44.1 | 38.1 | 35 |
| Stability of Employment | 38.3 | 32.3 | 24.8 | 32.3 | 36.6 | 30.9 |

In Table 3 we see that on average, students overestimate salary by 13.5% and underestimate both probability and stability of employment, by 24% and 4.3% respectively. The absolute errors indicate that students generally make large errors; 50% of students make errors larger than 30% for each of the labor market outcomes. Table 3 also shows that students' errors differ by major field. Students' errors are the largest, on average, for humanities and the smallest for business. Students are more likely to underestimate the labor market outcomes associated with business than they are for other fields.

### *Regression Outcomes*

We examine student errors in two ways: asking first if certain groups of majors are associated with larger or smaller errors, and asking second if certain groups of students have particularly large errors.

We first estimated a model that includes student fixed effects. Thus, the results from this model express the relative error for a given labor market outcome for a student. This student fixed effects model allows us to ask: do students systematically make larger or smaller errors for certain subjects or for certain labor market outcomes? We predicted two outcomes: (1) error measure as defined above, (2) absolute value of the error measure defined above. We use the model below:

$$\boldsymbol{Y_{i,j,bm} = \Gamma_{bm}\beta_1 + \Theta_j\beta_2 + \phi_i\beta_3 + \beta_4 own_{bm,i} + \varepsilon_i} \quad (2)$$

$Y_{i,j,bm}$ is student *i*'s error in predicting labor market outcome *j* in broad major category *bm*. $\boldsymbol{\Gamma_{bm}}$ is a vector of broad major category fixed effects, $\boldsymbol{\Theta_j}$ is a vector of dummy variables for each labor market outcome (expected salary, probability of employment, stability of employment), $\boldsymbol{\phi_i}$ is a vector of student

fixed effects and $own_{bm,i}$, a dummy variable equal to one if major category *bm* is the intended major of student *i*. Standard errors are clustered at the student level.

The results from model (2), exploring the heterogeneity in errors by major and labor market outcomes are presented in Table **4** (for reference, unconditional errors for each broad major category are shown in Table **3**).

Table 4: Predicting Errors in Labor Market Outcome Estimates, Student Fixed Effects

| | % Error | \|% Error\| |
|---|---|---|
| Business/CS | -2.7851 | -7.2918*** |
| | (2.5226) | (2.0593) |
| Arts & Lang. Arts | 4.6116* | -0.0710 |
| | (2.5340) | (2.0686) |
| Hum & Soc. Sci. | 2.7257 | -0.4588 |
| | (2.4971) | (2.0385) |
| Prob. Of Employment | -36.0296*** | -5.7012*** |
| | (2.2236) | (1.8152) |
| Stability of Employment | -16.6109*** | -8.9726*** |
| | (2.1849) | (1.7836) |
| Intended Major | 6.5463*** | -2.6477 |
| | (2.0774) | (1.6958) |
| Constant | 9.7618*** | 49.2940*** |
| | (2.3434) | (1.9130) |
| Student Fixed Effects | X | X |
| Observations | 3,646 | 3,646 |
| R-squared | 0.3474 | 0.2222 |

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1
Omitted major category is Sci, Math & Eng, Omitted LMO is salary. Unit of observation is at the student-by- major category-by-outcome level.

Individual students have smaller errors in estimating the average labor market returns to Business/CS than they do estimating the labor market returns to science, math and engineering (the omitted category). On average, students overestimate the labor market outcomes for Arts & Language Arts degrees. Both of these results broadly match the results shown in Table 3, which confirms that the patterns we see across the sample hold when we look within students. Individual students maker smaller errors in estimating the probability and stability of employment than they do in estimating expected salary, and these errors are more likely to be underestimates than the errors they make when estimating salary (both the percent error and the absolute value of these errors are negative). As we saw in Table 3, variance in errors is much larger across outcome statistics than it is across disciplines. Students are generally no more accurate in estimating the labor market outcomes of their intended major as compared to other majors (the coefficient on intended major when predicting the absolute value of the error is not significant), but errors in intended majors tend to be positive (the coefficient on intended

major when predicting error is significant and positive), implying that students tend to underestimate outcomes in their non-intended majors.

We next estimate a model that examines differences between groups of students. We use the following model:

$$\boldsymbol{Y_{i,j,bm} = \Gamma_{bm}\beta_1 + X_i\beta_2 + \beta_4 own_{bm,i} + \varepsilon_i} \quad (3)$$

All terms are as defined above, and $\mathbf{X_i}$ is a vector of student-level covariates which includes a dummy variable indicating if a student is from a low income family, a dummy variable for gender, and a dummy variable indicating if the student is in his/her first term. Standard errors are clustered at the student level. We estimate this model six times: error and absolute value of error for each of the three labor market outcomes (stability and probability of employment and expected salary). The coefficient of interest is $\boldsymbol{\beta_2}$. That is, for a particular labor market outcome, within a broad major category, do certain groups of students have larger or smaller errors?

The results from model 3 are presented Table 5 below.

Table 5: Relationship between Student Characteristics and Errors in Predicted Labor Market Outcomes

| | Salary | | Prob. of Employ. | | Stab. of Employ. | |
|---|---|---|---|---|---|---|
| | % Error | \|% Error\| | % Error | \|% Error\| | % Error | \|% Error\| |
| Low Income | 5.7057 | 6.2176 | -6.2072 | 4.5211* | -2.1294 | -0.9153 |
| | (7.8733) | (5.8909) | (4.0932) | (2.3615) | (5.1506) | (3.4223) |
| Female | -8.0773 | 4.2466 | 2.0368 | -0.5228 | -1.8945 | 4.9655 |
| | (7.2837) | (5.3630) | (4.0802) | (2.3492) | (4.9109) | (3.2043) |
| First Term | 9.5481 | 17.4074** | -3.5417 | 3.0486 | -6.6470 | -4.4768 |
| | (10.8579) | (8.1647) | (4.7938) | (2.8071) | (5.0778) | (3.1791) |
| Intended Major | 9.7176** | 3.5495 | 6.6093*** | -3.7389*** | 4.1830 | -6.5145** |
| | (4.0045) | (3.6694) | (1.9760) | (1.3850) | (3.4821) | (3.0853) |
| Constant | 8.6512 | 37.5259*** | -26.1322*** | 39.9803*** | -1.1887 | 43.2329*** |
| | (6.2830) | (4.5982) | (3.9879) | (2.3647) | (7.3769) | (6.4344) |
| Field Fixed Effects | X | X | X | X | X | X |
| Observations | 1,064 | 1,064 | 1,142 | 1,142 | 1,210 | 1,210 |
| R-squared | 0.0214 | 0.0213 | 0.0116 | 0.0275 | 0.0043 | 0.0125 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Low income students are more likely than their higher income peers to have large errors in estimating the probability of employment (fourth column of Table 6), though these errors are not systematically over- or under-estimates. Low- and high-income students do not have significantly different expectations about salary, which is different from what Hastings et al (2015) found. The difference with Hastings et al. (2015) may be explained by the fact that we consider a sample of relatively disadvantaged students while

Hastings et al study the whole population of college-bound Chilean students. Students in their first term make significantly larger errors in estimating salary than students who have been enrolled longer (Table 6, second column). Their errors are not systematically over- or under-estimates. Students are more accurate in estimating labor market outcomes (stability and probability of employment) in their intended major than all other majors, similar to what Arcidiacono et al (2012) found. Students are more likely to overestimate the salary in their intended major as compared to other majors[5] (first column).

Why do students overestimate salary in their intended major? Students may correctly think that they will make more in their intended major than students who do not choose this same major. However, we do not compare students' beliefs about their *own* salary with their beliefs about the salary for other students, nor do we ask them to compare their beliefs about their own salary in their intended major as compared to other majors. Instead, we are asking students about the salary of *past graduates* from their college in a range of majors. Thus, comparative advantage cannot explain intended major overestimation. Instead, it is more likely that being overoptimistic about salary makes students more likely to prefer a certain major. We will explore the determinants of major choice – salary among them – below.

Overall, we find that individual students are typically unable to correctly rank majors, and students as a whole have biased beliefs, overestimating salaries (by 13%) and underestimating the probability of employment (by 25%). Furthermore, low income students make larger errors in estimating the probability of employment. We now move on to examining the impact of students' beliefs about labor market outcomes on their major choice.

## Determinants of Major Choice

### *Non-experimental results*

First, we use observational data to examine the relationships between choice and students' estimates of labor market outcomes, ability and enjoyment. Summary statistics for the variables shown in this correlation table can be found in Table 6 .

Table 6: Correlations between Probability of Choice and Estimates of Outcomes

| | Prob. of choosing | Expected Salary | Exp. Prob. of having a job | Exp. Stab. of employment | Predicted enjoyment | Predicted grades |
|---|---|---|---|---|---|---|
| Prob. of choosing | 1 | | | | | |
| Expected Salary | .0921 * | 1 | | | | |
| Exp. Prob. of having a job | .1042 ** | .1792 *** | 1 | | | |
| Exp. Stab. of employment | .0966 ** | .1449 *** | .6613 *** | 1 | | |
| Prob. of enjoyment | .6326 *** | 0.0398 | .0726 * | 0.0617 | 1 | |
| Predicted grades | .4806 *** | -0.0118 | 0.0539 | 0.0575 | .5657 *** | 1 |

[5] Differences in errors in estimation are presented visually in the Appendix B in kernel density plots.

Table 6 shows that the probability of choice is significantly correlated with all expected outcomes. Labor market outcomes (salary, probability and stability of having a job) are significantly positively correlated with each other. The belief about salary is not correlated with beliefs about enjoyment or ability; this is in contrast to the finding of Wiswall and Zafar (2015a) that taste for majors is correlated with salary expectations. On the other hand, the belief about having a job is correlated with the belief about enjoyment.

To examine how these beliefs predict the probability of choosing a given field, we estimate the following model based on Blass, Lach, and Manski (2010), and also used by Wiswall and Zafar (2015a):

$$\left(ln\pi_{k,i} - ln\pi_{b,i}\right) = \beta_0 + \beta_1\left(lnw_{k,i} - lnw_{b,i}\right) + \beta_2 X_{k,i} + \varepsilon_{k,i} \quad (4)$$

where $\pi_{k,i}$ is the stated probability for student i to major in field k. The subscript b stands for Business, Computer Science & Technology, as everything is relative to this major category. The key explanatory variable is $lnw_{k,i} - lnw_{b,i}$, i.e. the difference in belief about salary of past graduates in this field relative to humanities. Thus, the variation we use to identify $\beta_1$ is a combination of the true difference in the labor market outcomes between field *k* and Business, Computer Science & Technology, and students' errors in estimating this difference.[6]

In some specifications, we also add differences in beliefs about having a job, job stability (both shown in Table 7), and predicted enjoyment and average grade (shown in Table 8). We also extend our baseline specification in two ways: adding a vector of individual controls ($X_{k,i}$, including gender, race, measures of SES, and high school GPA) and adding a vector of student fixed effects.

Individual controls account for the propensity of certain types of students to choose a major other than Business, Computer Science & Technology. For example, a positive coefficient on female would mean that females are more likely to choose something other than Business, Computer Science & Technology. To the extent that certain types of students make systematic errors about the labor market outcomes of different majors, this could lead to bias in the coefficients on the labor market outcomes if we do not control for student type. For example, if we assume that females systematically underestimate the labor market outcomes for Business, Computer Science & Technology, this could explain their choices. In this case, controlling for gender will weaken the relationship between labor market outcomes and major choice. (Appendix C shows the estimated relationship between various individual characteristics and major choice).

[6] The reason why everything must be expressed relative to a baseline choice is that the econometric model we use here is an extension of the multinomial logit (Blass, Lach, and Manski 2010). In the standard logit, there are only two choices, so the reference choice is easily determined as "the other choice." However, in a multinomial logit, there are three or more unordered choices. Therefore, a reference choice must be used as the default choice to which other choices are being compared.

The results from the student fixed effects model express the relative relationship between a labor market outcome and probability of choosing a major *for a student*. This student fixed effects model allows us to control for heterogeneity in individual students' preference for other majors relative to business. A similar logic can be applied as the one discussed above for female students. That is, if some individual students have a greater preference for business vs. other majors, this could lead them to overestimate the labor market outcomes of business relative to other majors. In such a scenario, we may overestimate the importance of labor market outcomes in major choice. Introducing student fixed effects allows us to control for student heterogeneity in the preference for other majors relative to business. The fixed effects estimate informs us about the choice between humanities, language and arts, and science after taking into account a student's general preference for business relative to these other majors.

Table 7: Relationship between Major Choice and Predicted Labor Market Outcomes

| | Stated Probability of Choosing a Given Major Category | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Predicted Salary | 0.7227** | | | | 0.7139*** | 1.0954** |
| | (0.3168) | | | | (0.2606) | (0.4608) |
| Pred. Prob. of Employment | | 0.3152 | | | 0.1703 | 0.1232 |
| | | (0.2118) | | | (0.2132) | (0.2885) |
| Pred. Enjoyment | | | 0.9869*** | | 0.7831*** | 0.8675*** |
| | | | (0.0418) | | (0.0509) | (0.0800) |
| Predicted Grades | | | | 1.8033*** | 0.8245*** | 0.6566*** |
| | | | | (0.1110) | (0.1338) | (0.2376) |
| Constant | 1.0418*** | 0.9557*** | 0.0756 | 0.0718 | 0.0958 | -0.3063* |
| | (0.1851) | (0.1818) | (0.1357) | (0.1561) | (0.1644) | (0.1731) |
| Student Fixed Effects | | | | | | X |
| Observations | 594 | 594 | 594 | 594 | 594 | 594 |
| R-squared | 0.0088 | 0.0050 | 0.3792 | 0.2462 | 0.4249 | 0.7905 |

Robust standard errors clustered at the student level are shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Unit of observation is at the student-by-broad major level. Business is the reference major.

In Table 7, columns 1 and 2, we can see that, when we predict probability of choosing a major as a function of either salary or the probability of employment, salary is predictive but probability of employment is not. The estimate on salary implies that a 10% change in salary is associated with a 7% change in the probability of majoring in a given broad major field. In the next two columns we repeat the same specifications controlling individually for each of two key determinants of major choice: beliefs about enjoyment and average grades. Predicted enjoyment and estimated grades are significant determinants of major choice; a 10% increase in either enjoyment (grades) is associated with about an 10% (18%) increase in the probability of majoring in a broad major category relative to business. When we control for all four beliefs together, these three predictors remain significant and positive.. Finally, when we include individual student fixed effects (column 6), we see that the relationships remain largely the same. In sum, these results confirm that when the choice is between majors other than business, students favor the higher paying major and the majors that they think they will enjoy more and do better in relative to business. This is true both across students and for any given student. The estimated relationship between various individual characteristics and major choice is shown in Appendix C.[7]

Overall, we find that enjoyment, grades, and predicted salary are the most important determinants of major choice (enjoyment is even more important than grades, as can be assessed by an R-squared that is 1.5 times as large when we include only enjoyment vs. only grades as explanatory variables). Even though the average relationship between salary and major choice is of similar magnitude to the relationships between enjoyment and grades and major choice, labor market outcomes do not explain much of the

[7] We estimated these same models using each of the other three broad major categories as the reference as well. The results are qualitatively similar. The results are shown in Appendix D.

variance in major choice across students, as can be inferred from the very small R-squared for labor market outcomes.

### *Experimental*

In this section, we explore the effect of labor market outcomes on major choice by randomly assigning information about salary and the probability of employment in business and humanities to students. In our experiment, we present hypothetical labor market outcomes to students. There are two versions of the hypothetical labor market outcomes that were randomly assigned to students (described below, and presented graphically in Appendix E). We instruct students to act as if these outcomes were the real outcomes of past graduates.

Specifically, in the survey (see Appendix A), students are asked to recall and write down their beliefs about the probability of being employed, their expected salary, and their probability of choosing Business, Comp. Sci. & Technology. They are then given, on the same page, two different hypothetical scenarios with respect to the employment opportunities and salary in Business, Comp. Sci. & Technology. For each scenario, corresponding to a specific employment rate and salary in the field of Business, Comp. Sci. & Technology, students report their probability of choosing this major. The same procedure is repeated on the next page for Humanities, Education and Behavioral/Social Sciences.

In all, students saw four hypothetical scenarios: two for Business, Computer Science & Technology and two for Humanities, Education and Behavioral/Social Sciences. In order to separately identify the effects of expected salary and the effects of expected probability of employment, we use a factorial design. That is, each student saw all possible combinations of high and low estimates of the probability of employment (85-90% and 35-40%, respectively) and high and low estimates of salary ($75-80,000 and $30-$35,000, respectively). Each student was exposed to all four conditions: high employment, high salary; high employment, low salary; low employment, low salary; low employment, high salary. However, each student did not see all four conditions for each of the two majors. This design is depicted graphically in Appendix E.

In the experiment we randomly assign hypothetical information to students, and we therefore eliminate the source of bias due to correlation between students' beliefs about labor market outcomes and other determinants of major choice.

We first run this specification[8]:

$$ln\pi'_{k,i} = \beta_0 + \beta_1 lnw'_{k,i} + \epsilon'_{k,i} \quad (5)$$

[8] In the specification using non-experimental data, we had used a generalization of the ordered logit because students had to choose between several different majors with different labor market outcomes. Identification came from the variation across majors in expected labor market outcomes. Here, for the experimental results, the students report probability of choice given variation in labor market outcomes for the *same* major. In Appendix G we present results where we re-analyze the experimental data using methods more similar to what we used for the non-experimental data. That is, we predict the difference in probabilities of choosing majors using the difference in stated labor market outcomes. The results are not qualitatively different from what we present here.

Where $\pi_{k,i}$ is the probability for student *i* to major in field *k*. The key explanatory variable is $lnw'_{k,i}$, the expected labor market outcome for major *k* (salary (column (1), probability of employment (column 2) or both (column 3) in Table 10 and Table 11). We run this specification separately for each of the two major groups: Business, Computer Science & Technology, and Humanities, Education and Behavioral/Social Sciences. We ask each student about two hypothetical scenarios for each of the two broad major categories, so we have two observations per student per broad major field. Because the labor market information is randomly assigned, it should not be correlated with any omitted factors. We also add two additional specifications where we control for students' previously stated beliefs about salary and employment, as well as previously stated probability of choosing each major category (from the observational analyses described above, which they were prompted to re-enter in the experimental tables).[9]

The impact of salary on major choice is positive and highly significant based on our experiment: column 1 in Tables 9 and 10 shows that a 1% increase in salary leads respectively to a 1.4% and 1.8% increase in the probability of choosing a major. The effect of salary on major choice is larger for humanities than for business, but the difference is not significant (results from a joint regression including an interaction between business and salary not shown). The impact of the probability of employment when included alone is positive but insignificant (Tables 10 and 11, col. 2). When salary and the probability of employment are included together, the impact of the probability of employment becomes larger and significant for humanities majors (Table 11, col. 3) implying that a 1% increase in the probability of employment leads to a 0.46% increase in the probability of choosing a major.[10]

Labor market outcomes thus have a large impact on the probability of choosing a major according to our experimental estimates. While the effects seem large, their magnitude is within the range of what was found in the prior literature. For example, Wiswall and Zafar (2015a) show found effects of a 1% increase in earnings that vary between 0.3% and 2%. Our experimental estimates suggest that the true impact of labor market outcomes is towards the higher end of this range.

[9] We also tried a specification that includes an interaction between the "true" salary and "true" probability of employment. We find that this interaction is very small and not statistically significant; students do not seem to disproportionately prefer majors with high salary and probability of employment.

[10] In separate regressions we examine if "true" salary has differential effects for students based on their previous probability of choosing this major. In Appendix F we present results from these models, which show that students who previously stated there was zero chance they would select this major seem to be most affected by this new salary information.

Table 8: Experimental Estimates of Effect of Labor Market Outcomes on Major Choice: Business, CS, and Technology

| | Stated Probability of Choosing Business/CS/Technology | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ln(hypothetical salary) | 1.4244*** | | 1.3903*** | 1.4540*** | 1.5003*** |
| | (0.2735) | | (0.2767) | (0.2888) | (0.2473) |
| ln(hypo. prob. of employ.) | | 0.5569 | 0.3647 | 0.3880 | 0.4016 |
| | | (0.3417) | (0.3362) | (0.3511) | (0.3006) |
| Prev. Belief of Salary | | | | -0.2329 | 0.0040 |
| | | | | (0.2702) | (0.2371) |
| Prev. Belief of Prob. Employ. | | | | 0.1065 | -0.1634 |
| | | | | (0.2547) | (0.2235) |
| Prev. Prob of Choosing | | | | | 0.4639*** |
| | | | | | (0.0367) |
| Constant | -17.9108*** | -2.2720*** | -17.3465*** | -15.3494*** | -16.7223*** |
| | (2.9632) | (0.2197) | (3.0301) | (4.3134) | (3.7039) |
| Observations | 431 | 431 | 431 | 376 | 362 |
| R-squared | 0.0595 | 0.0062 | 0.0621 | 0.0721 | 0.3434 |

Robust standard errors, clustered at the student level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Observations are at the student-by-scenario level.

Table 9: Experimental Estimates of Effect of Labor Market Outcomes on Major Choice: Humanities, Education and Behavioral/Social Science

| Stated Probability of Choosing Humanities/Education/Behavioral & Social Science | | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| ln(hypothetical salary) | 1.8166*** | | 1.9043*** | 1.9834*** | 2.0498*** |
| | (0.2850) | | (0.2863) | (0.2933) | (0.2562) |
| ln(hypo. prob. of employ.) | | 0.1861 | 0.4625* | 0.5442** | 0.5557** |
| | | (0.2614) | (0.2506) | (0.2565) | (0.2241) |
| Prev. Belief of Salary | | | | -0.0552 | -0.0651 |
| | | | | (0.2411) | (0.2169) |
| Prev. Belief of Prob. Employ. | | | | -0.2947** | -0.4622*** |
| | | | | (0.1492) | (0.1331) |
| Prev. Prob of Choosing | | | | | 0.4341*** |
| | | | | | (0.0430) |
| Constant | -21.7554*** | -1.8889*** | -22.4405*** | -22.9385*** | -22.3808*** |
| | (3.1365) | (0.1941) | (3.1355) | (4.0630) | (3.5470) |
| Observations | 410 | 410 | 410 | 361 | 345 |
| R-squared | 0.0906 | 0.0012 | 0.0980 | 0.1212 | 0.3755 |

Robust standard errors, clustered at the student level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Observations are at the student-by-scenario level.

Adding controls to a true experiment should not affect the size of the estimates on the variable that is randomly assigned. This is indeed what happens in Table 10 & Table 11 for true salary (compare coefficients in column 1 and columns 4-5). It is also remarkable that the initial belief about salary no longer matters conditional on true belief (column 4), which is a way of validating that we succeeded in fully changing students' beliefs. Indeed, if we had only updated the beliefs to some limited amount, we would expect both the true salary and the past belief to matter (positively) for student choice.

The initial probability of choosing a major has a positive and highly significant impact on hypothetical major choice after the experiment, and also triples the R-squared (Table 10 & Table 11, column 5). This is consistent with the fact that there are other important determinants of major choice beyond salary, so people who are inclined to choose a certain major are still inclined to choose it even after accounting for new beliefs about salary.

### *Experimental vs. non-experimental*

In this section, we attempt to explain why the experimental impact of salary on major choice is greater than the nonexperimental impact. We do not find evidence to support four possible explanations. First, it is not the case that the difference is due to the fact that the experiment only deals with humanities and business majors while the observational data considers all majors. Second, the difference is not due to differences in modelling strategies between the experimental and non-experimental analyses (relative vs. absolute change). We provide evidence concerning these two explanations in Appendix G.

Third, of the variables we measured on the survey, we did not find a variable that is negatively correlated with earnings beliefs and is an important determinant of major choice. The variables we measured were exhaustive based on prior literature, and include all that we think are important determinants of major choice. The explanation therefore does not seem to be an omitted variable bias.

A fourth explanation that we do not find support for is more psychological: by only varying labor market outcomes, we make those a lot more salient in our experimental intervention, and we are thus perhaps overestimating their impact. However, two factors make this explanation unlikely. First, the fact that the probability of employment is not such a strong determinant of major choice in the experiment, both in absolute terms and relative to the non-experimental data, even though it was listed before salary in our scenarios, makes this psychological explanation less likely. Second, we explicitly remind people of their original stated probability of choosing this major. This implicit anchoring should make respondents more likely to stick to what they first said[11], and makes it less likely that we are overestimating the impact of labor market outcomes on major choice in the experiment relative to the non-experimental estimates.

Instead, we find support for an explanation based on students' uncertainty. From a Bayesian perspective, if students are more uncertain about some of the determinants of major choice, they should put less weight on these determinants. For example, if students are more uncertain about labor market outcomes than about how much they would enjoy the major, they should put less weight on labor market outcomes than on enjoyment when choosing a major. In the experiment, we provide the "true" labor market outcomes, thus removing uncertainty and increasing the estimated impact of labor market outcomes on major choice. This could therefore explain why the experimental estimates of the impact of labor market outcomes on major choice are larger than the non-experimental estimates.

To test this uncertainty hypothesis, we first calculate for each student the percent error (as described above). We take the absolute value of the error and add up the errors for each student. This sum of errors tells us which students make the largest mistakes; these students are likely to be the least certain about their salary predictions, and thus their belief about salary should matter less to their major choice. In Table 11 below, we use the observational data and run specifications similar to those in Table 7, adding an interaction between the salary belief and the sum of errors (columns 2 and 5). The interaction term in the model that includes student fixed effects is significant and negative (column 5); the relationship between perceived salary and choice is indeed larger for those who make smaller errors. In columns 3 and 6, we divide the students in below the median sum error and above the median sum error, thus creating the "large error" dummy. This shows that the relationship between salary and choice is indeed larger for those who make smaller errors because the point estimate of salary is larger in columns 3 and 6 than in columns 1 and 4, and the interaction term is significant and negative in the model with student fixed effects. These results can thus provide evidence of why experimental results yield a higher impact of salary than the non-experimental results: in the non-experimental case, students who are uncertain about labor market outcomes are likely to put less weight on them when choosing their majors.

[11] Consistent with this interpretation, the original probability of choosing a major has a strong and positive effect on the probability of choosing this major in the experiment in Table 9 and Table **10**.

Table 10: Relationship between Student Errors and Role of Salary in Major Choice

| | Stated Probability of Choosing a Given Major | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ln(Belief of Salary) | 0.7404*** | 0.6199 | 1.3185** | 1.4340*** | 2.5023*** | 3.1833*** |
| | (0.2804) | (0.5032) | (0.5691) | (0.5154) | (0.7251) | (0.9755) |
| Sum of Salary Errors | | -0.0003 | | | 0.0029*** | |
| | | (0.0009) | | | (0.0004) | |
| Sum(error)*ln(Sal. Belief) | | 0.0004 | | | -0.0036** | |
| | | (0.0015) | | | (0.0016) | |
| Error>p50(Error) | | | -0.2156 | | | 3.3620*** |
| | | | (0.3455) | | | (0.3112) |
| Large Error*ln(Sal. Belief) | | | -0.7337 | | | -2.3241** |
| | | | (0.6533) | | | (1.1212) |
| Constant | 1.0332*** | 1.0720*** | 1.1983*** | -0.4667*** | -0.3886* | 0.0914 |
| | (0.1632) | (0.2507) | (0.2660) | (0.1644) | (0.2258) | (0.3112) |
| Student Fixed Effects | | | | X | X | X |
| Observations | 775 | 775 | 775 | 775 | 775 | 775 |
| R-squared | 0.0088 | 0.0093 | 0.0103 | 0.6619 | 0.6669 | 0.6681 |

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1, reference category is business.

## Conclusion

In this paper, we provided new evidence about the role of labor market outcomes in determining major choice among community college students. We find that students have some information about labor market outcomes but this information is limited. Less than 40% of students correctly rank salary across broad major categories. Futhermore, students systematically overestimate salaries (by 13%) and underestimate employment outcomes (by almost 25%). Furthermore, lower income students are more likely to make mistakes in estimating the probability of employment.

We then examine the determinants of major choice among community college students. We find that course enjoyment and grades are the most important determinants. Labor market outcomes also play a role, but this role is underestimated in non-experimental data. Using experimental data, we find that a 1% increase in salary is associated with an almost 2% increase in the probability of choosing a major, while the probability of employment plays a much smaller role. We find evidence consistent with the greater impact of labor market outcomes in our experimental data being driven by students' lack of information at baseline. Indeed, if students are uncertain about labor market outcomes, it can be rational to put less weight on these labor market outcomes when choosing a major.

Our results suggest that providing better information about labor market outcomes has the potential to significantly affect major choice among community college students. We also find that students put the

most weight on course enjoyment when choosing a major. Thus, increasing the salience of labor market outcomes may also help students improve their labor market outcomes.

## References

Arcidiacono, P., Hotz, V. J., & Kang, S. (2010). *Modeling college major choices using elicited measures of expectations and counterfactuals*. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w15729

Attanasio, O., & Kaufmann, K. (2012). *Education Choices and Returns on the Labour and Marriage Markets: Evidence from Data on Subjective Expectations.* Working paper. Retrieved from http://www.homepages.ucl.ac.uk/~uctpjrt/Files/Educational%20Choices_Marriage_fin.pdf

Bahr, P. R. (2015). The labor market return in earnings to community college credits and credentials in California. Ann Arbor, Michigan: Center for the Study of Higher and Postsecondary Education, School of Education, University of Michigan. Retrieved from http://www.soe.umich.edu/people/profile/peter_riley_bahr/

Bailey, T. R., Jenkins, D., & Leinbach, T. (2005). What we know about community college low-income and minority student outcomes: Descriptive statistics from national surveys." New York, NY: Teachers College Community College Research Center.

Becker, G.S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70 (5): 9-49.

Belfield, C. R., & Bailey, T. (2011). The benefits of attending community college: A review of the evidence. *Community College Review*, *39*(1), 46–68.

Beffy, M., Fougère, D., & Maurel, A. (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? Review of Economics and Statistics, 94(1), 334–347.

Betts, J. R. (1996). What do students know about wages? Evidence from a survey of undergraduates. *Journal of Human Resources*, 27–56.

Blass, A., Lach, S., & Manski, C. (2010). Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability. *International Economic Review*, 51(2), 421-440.

Dadgar, M., & Weiss, M. J. (2012). Labor market returns to sub-baccalaureate credentials: how much does a community college degree or certificate pay? Retrieved from http://academiccommons.columbia.edu/catalog/ac:150856

Hastings, Justine, Christopher A. Neilson, and Seth D. Zimmerman. 2015. "The Effects of Earnings Disclosure on College Enrollment Decisions." Working Paper 21300. National Bureau of Economic Research. http://www.nber.org/papers/w21300.

Hastings, Justine S., Christopher A. Neilson, Anely Ramirez, and Seth D. Zimmerman. 2015. "(Un)Informed College and Major Choice: Evidence from Linked Survey and Administrative Data." Working Paper 21330. National Bureau of Economic Research. http://www.nber.org/papers/w21330.

Jacobson, L., LaLonde, R., & G Sullivan, D. (2005). Estimating the returns to community college schooling for displaced workers. *Journal of Econometrics*, *125*(1), 271–304.

Jepsen, C., Troske, K., & Coomes, P. (2014). The labor-market returns to community college degrees, diplomas, and certificates. *Journal of Labor Economics,* 32(1), 95-121.

Ma, Jennifer, and Sandy Baum. 2015. Trends in Community Colleges: Enrollment, Prices, Student Debt, and Completion. New York: The College Board. http://trends.collegeboard.org/sites/default/files/trends-in-community-colleges-research-brief.pdf.

Marcotte, D. E., Bailey, T., Borkoski, C., & Kienzl, G. S. (2005). The returns of a community college education: Evidence from the National Education Longitudinal Survey. *Educational Evaluation and Policy Analysis*, *27*(2), 157–175.

Rosenbaum, J. E., Deil-Amen, R., & Person, A. E. (2006). After admission: From college access to college success. Russell Sage Foundation.

Rosenbaum, James E., and Janet Rosenbaum. 2013. "Beyond BA Blinders: Lessons from Occupational Colleges and Certificate Programs for Nontraditional Students." Journal of Economic Perspectives, 27(2): 153-72.

Scott-Clayton, J. E. (2011). The shapeless river: Does a lack of structure inhibit students' progress at community colleges? Community College Research Center Working Paper #25.

Stevens, A. H., Kurlaender, M., & Grosz, M. (2015). Career technical education and labor market outcomes: Evidence from California Community Colleges. National Bureau of Economic Research Working Paper #21137.

Stinebrickner, T. R., & Stinebrickner, R. (2014). Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model. Journal of Labor Economics, 32(3), 601 – 644

Wiswall, M., & Zafar, B. (2015a). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2), 117-169.

Wiswall, M., & Zafar, B. (2015b). Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies*, 82(2), 791-824.

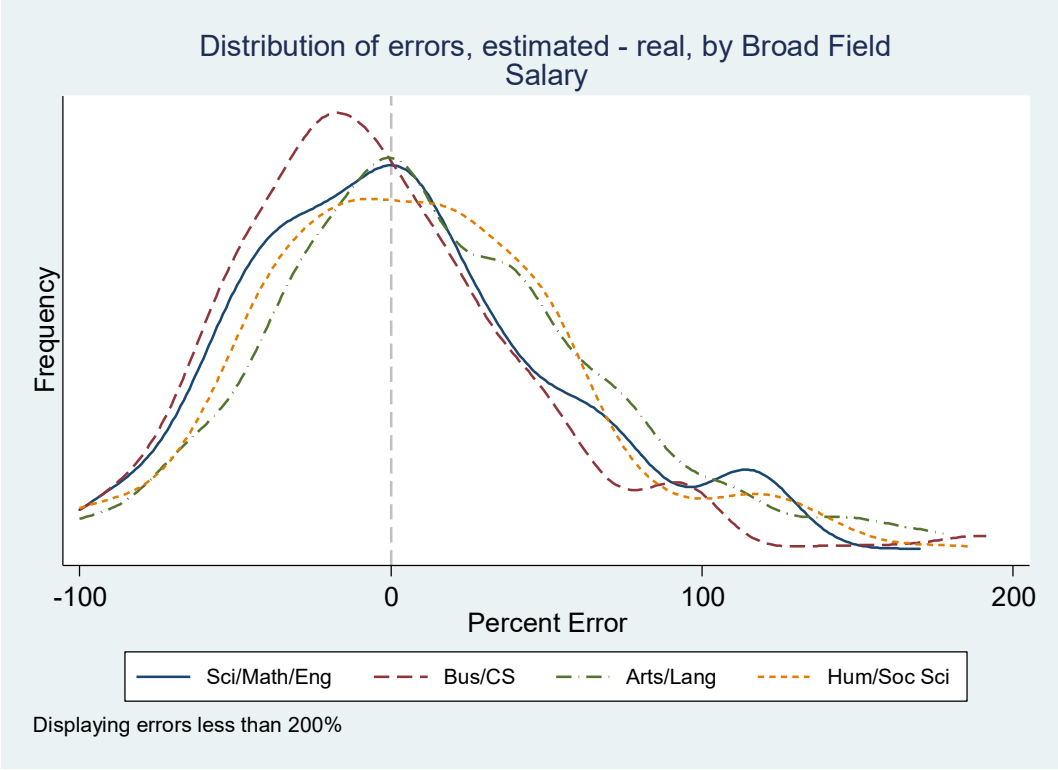
Zafar, B. (2009). College major choice and the gender gap. *FRB of New York Staff Report*, (364). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1348219

Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics*, *29*(2), 301–348.

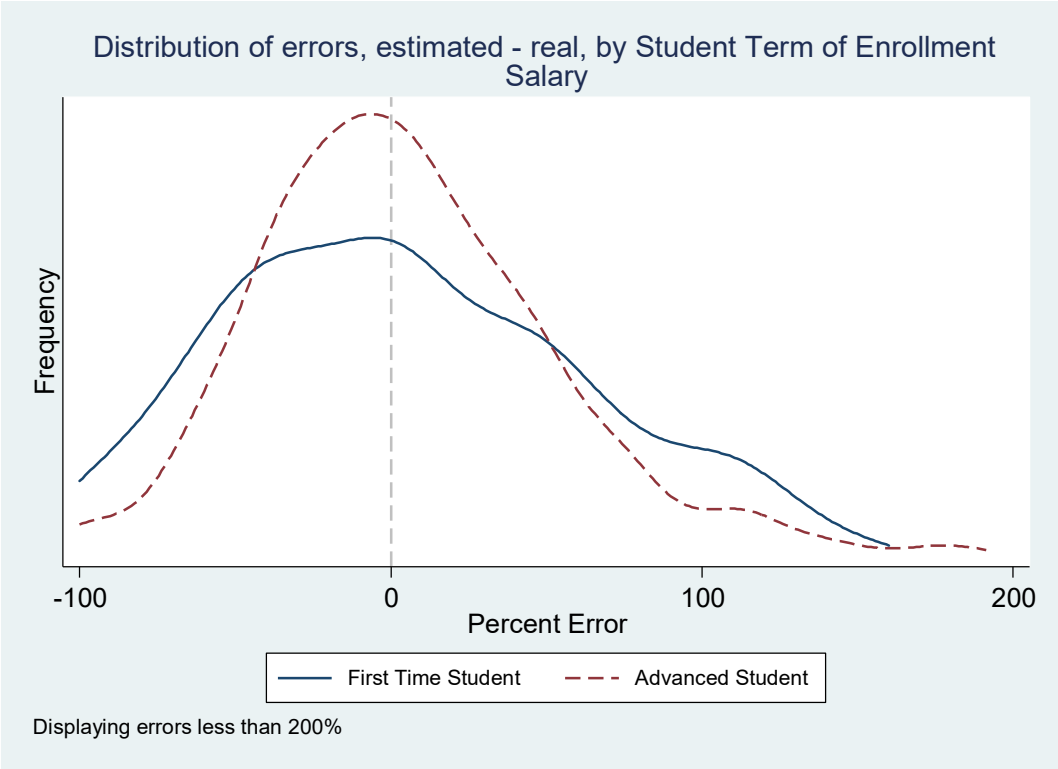
## Appendix A: Student Survey

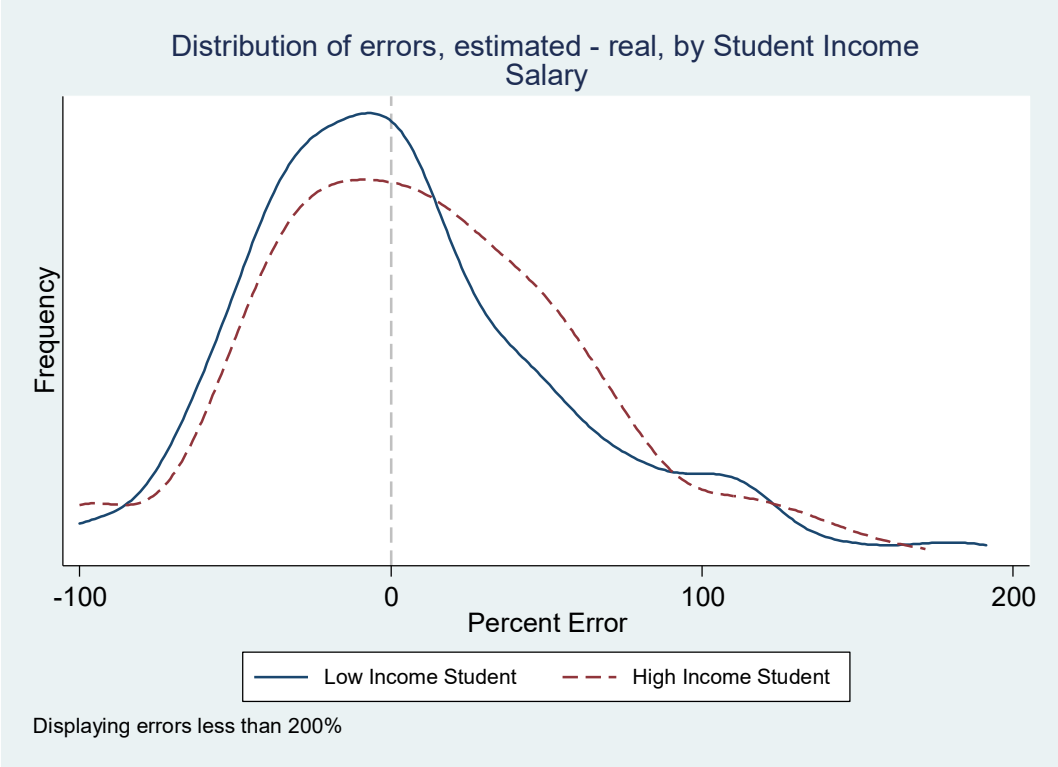
(in attached document)

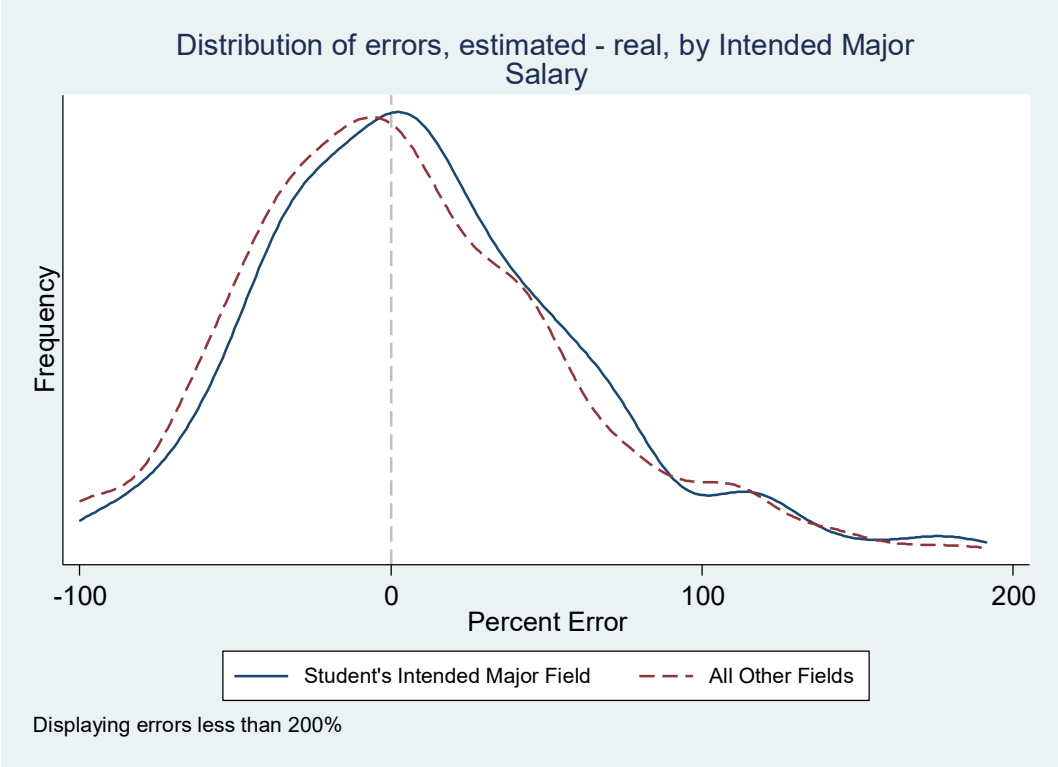
## Appendix B: Students' Errors in labor market outcomes expectations

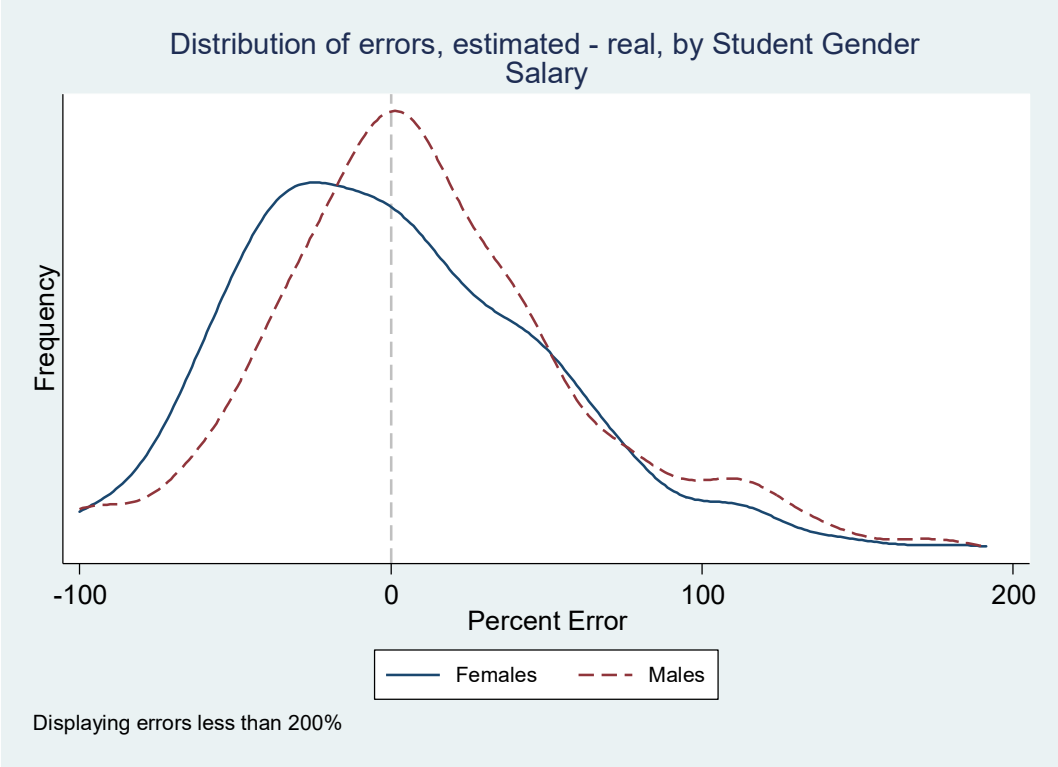


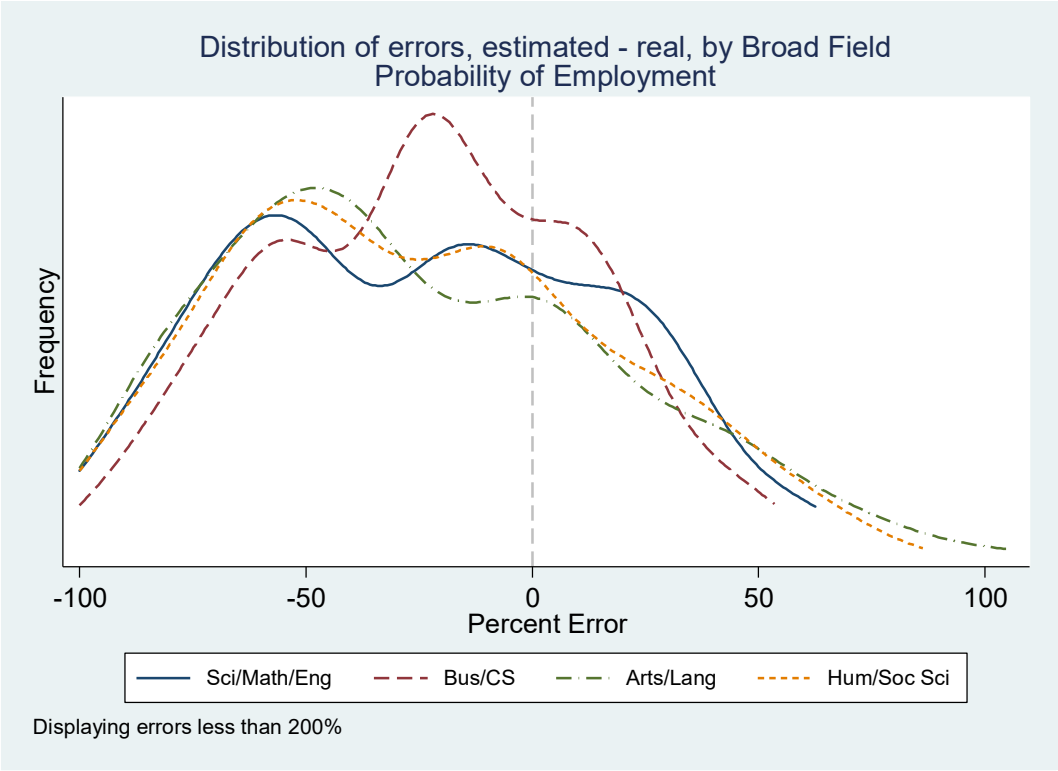

Notes: Interestingly, the mode of the distribution is around zero for all fields but Business/CS where it is below 0. At the same time, there is a substantial fraction of students who over-predict in each field. For business/CS the mode being lower and the over-prediction together results in a correct estimate on average.

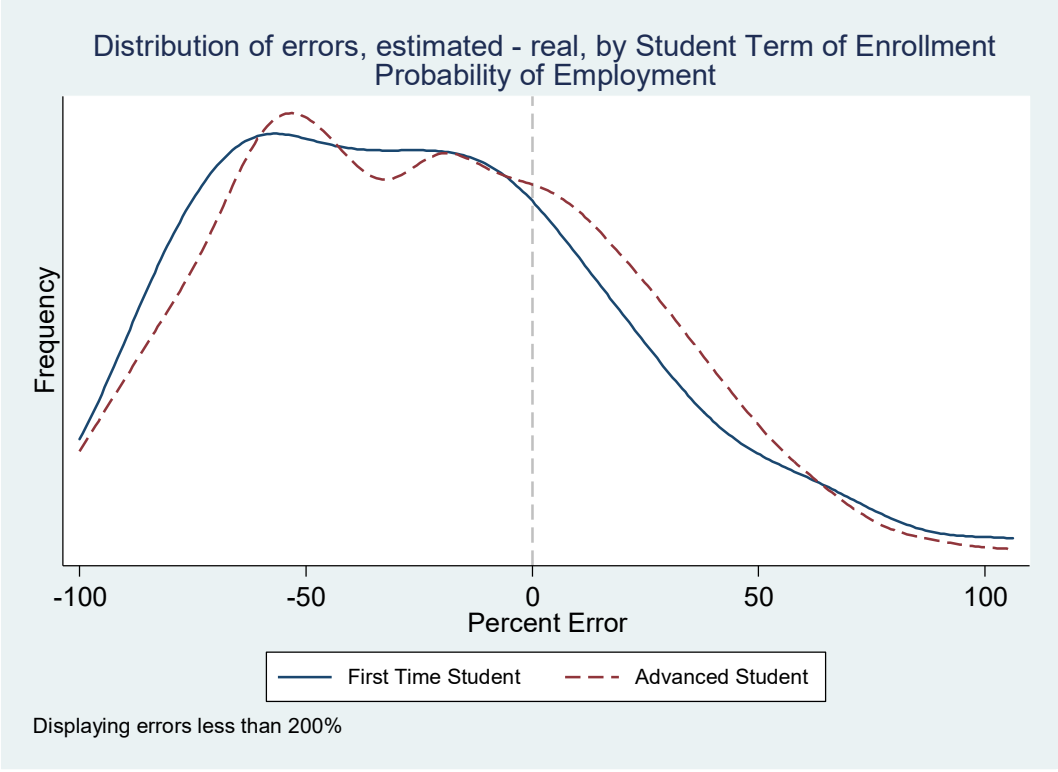

Distribution of errors, estimated - real, by Student Term of Enrollment
Salary
Frequency
-100
0
100
200
Percent Error
First Time Student
Advanced Student
Displaying errors less than 200%

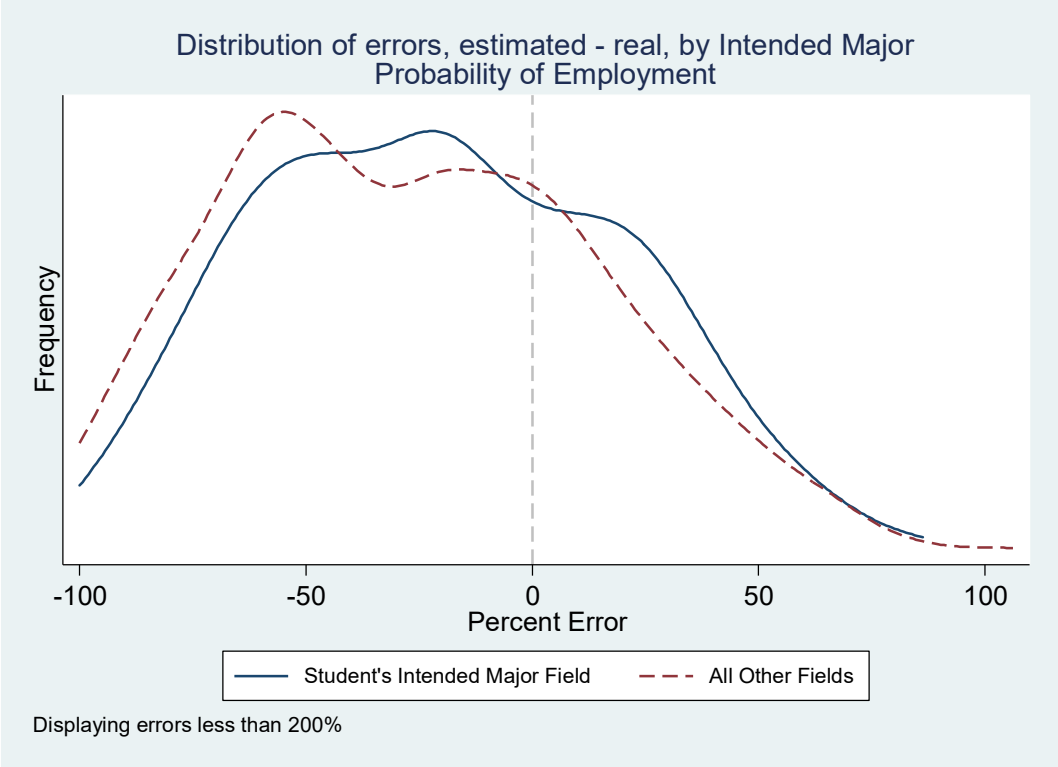

Distribution of errors, estimated - real, by Student Income
Salary
Frequency
-100
0
100
200
Percent Error
Low Income Student
High Income Student
Displaying errors less than 200%

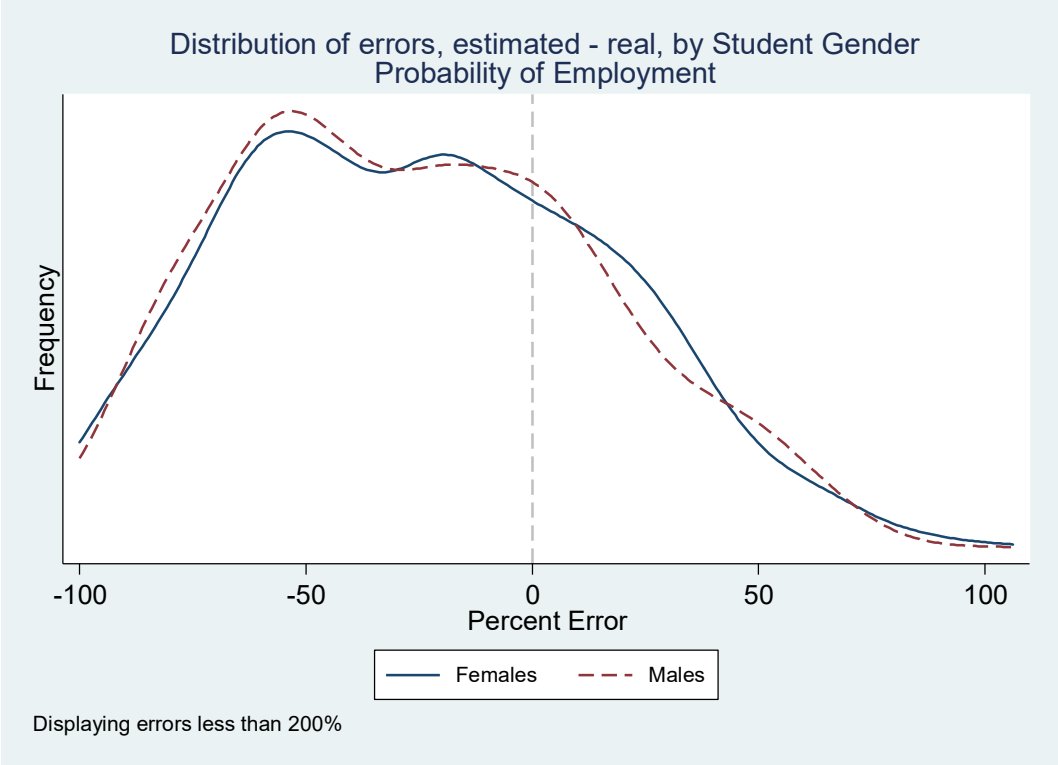

Distribution of errors, estimated - real, by Intended Major Salary
Frequency
-100
0
100
200
Percent Error
Student's Intended Major Field
All Other Fields
Displaying errors less than 200%

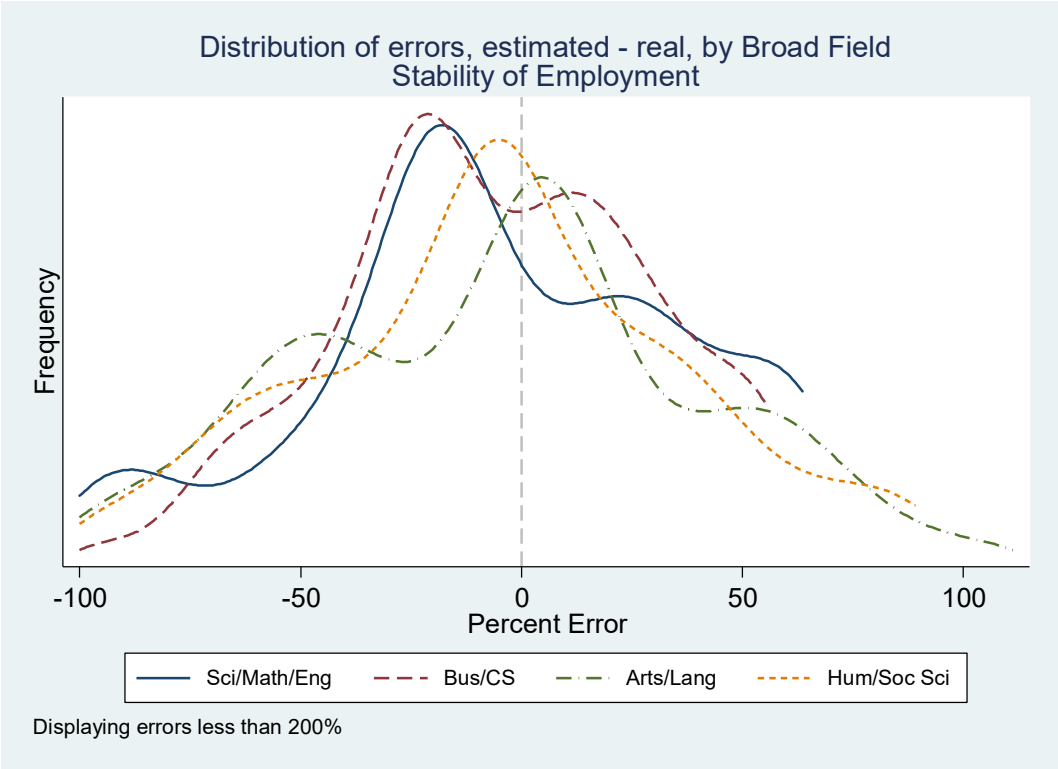

Distribution of errors, estimated - real, by Student Gender Salary
Frequency
-100
0
100
200
Percent Error
Females
Males
Displaying errors less than 200%

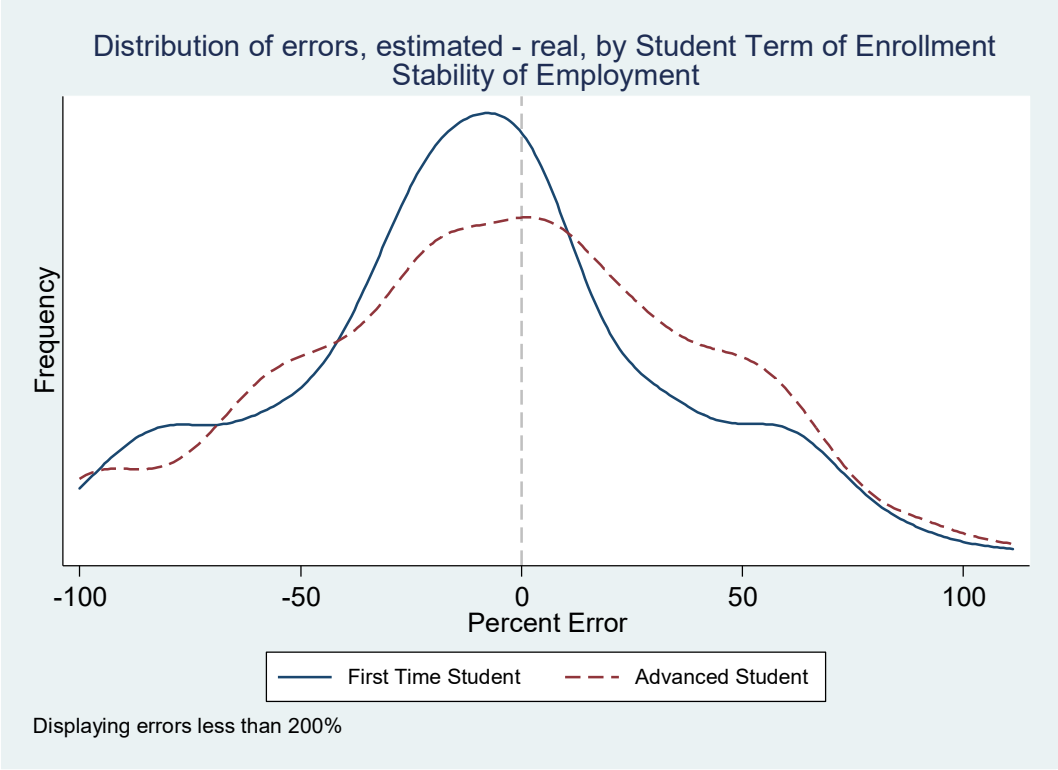

Distribution of errors, estimated - real, by Broad Field
Probability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Sci/Math/Eng
Bus/CS
Arts/Lang
Hum/Soc Sci
Displaying errors less than 200%

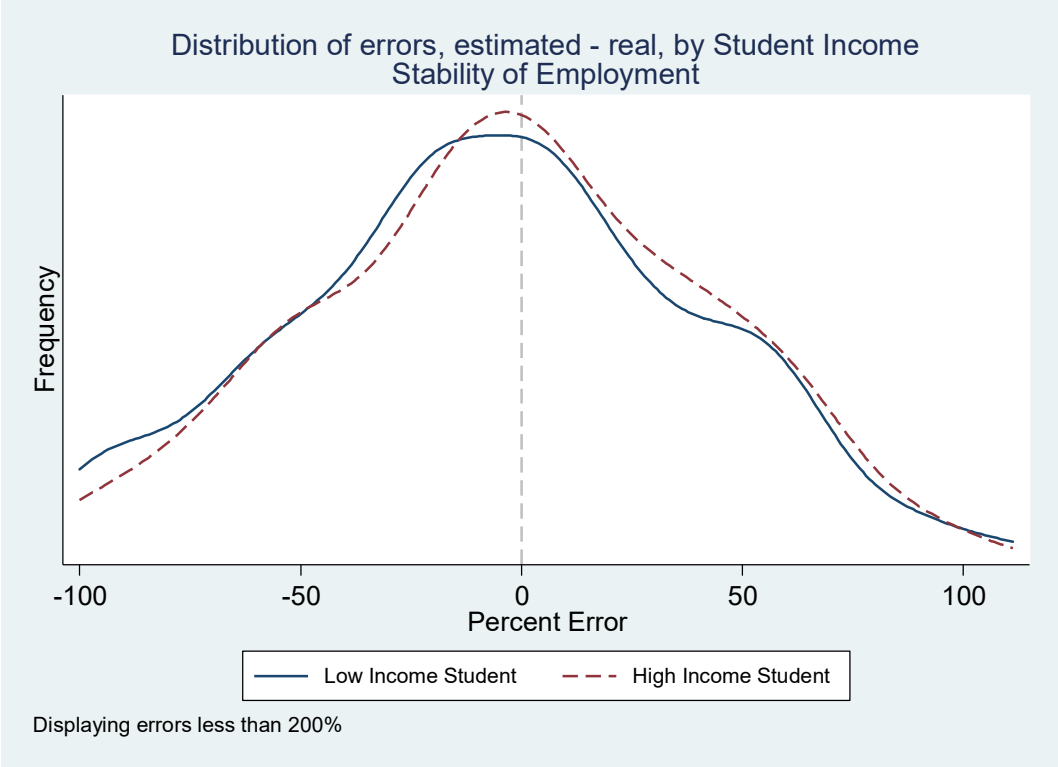

Distribution of errors, estimated - real, by Student Term of Enrollment
Probability of Employment
Frequency
-100
-50
0
50
100
Percent Error
First Time Student
Advanced Student
Displaying errors less than 200%

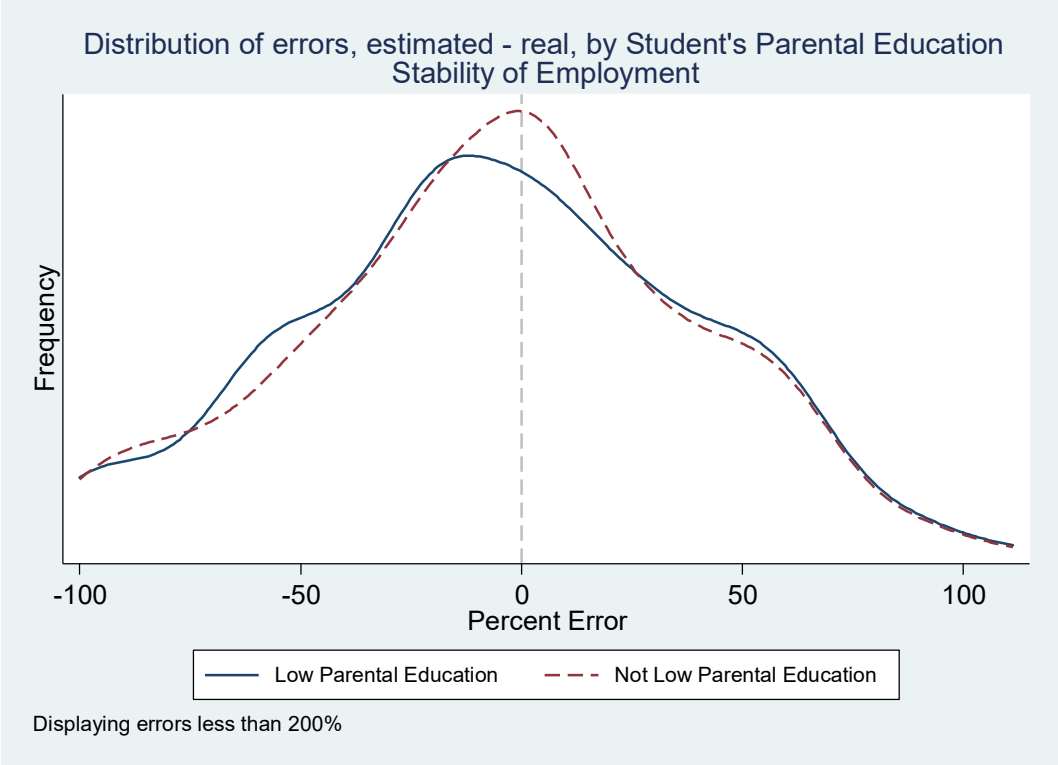

Distribution of errors, estimated - real, by Intended Major
Probability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Student's Intended Major Field
All Other Fields
Displaying errors less than 200%

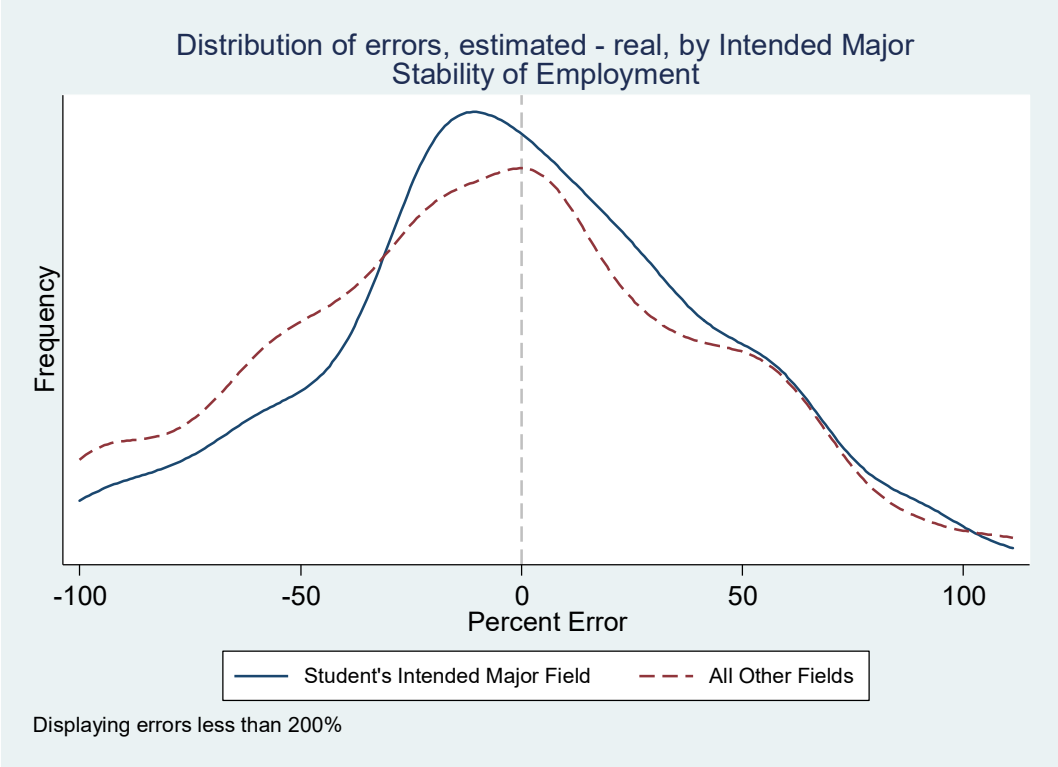

Distribution of errors, estimated - real, by Student Gender
Probability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Females
Males
Displaying errors less than 200%

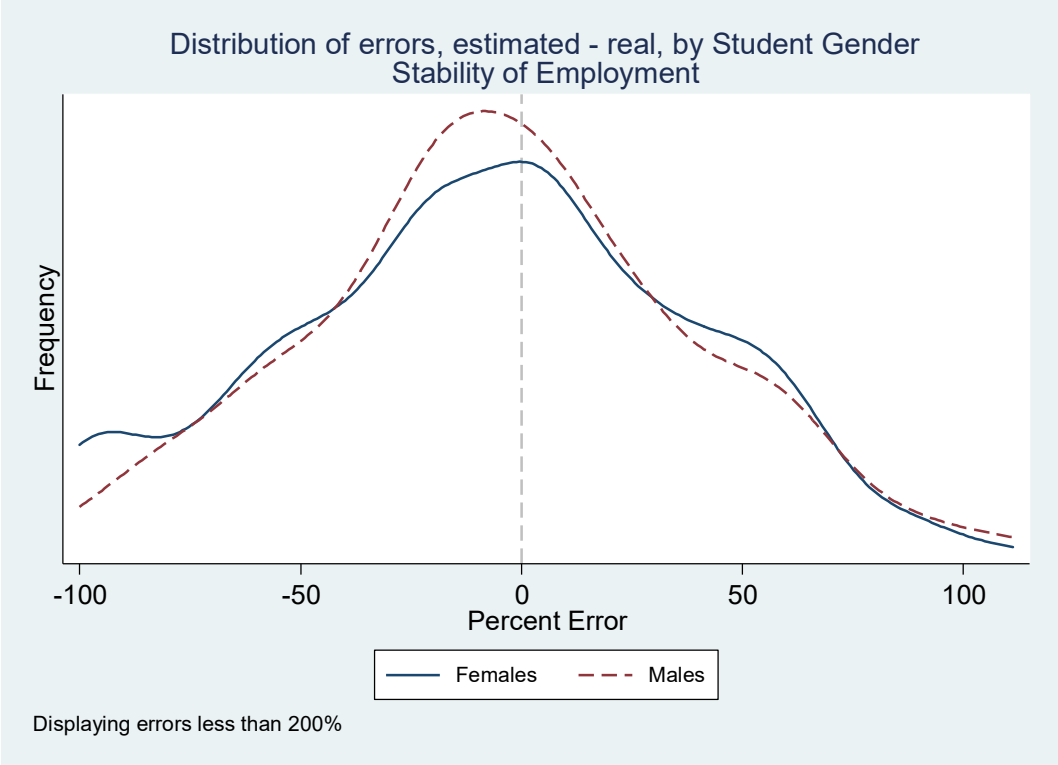

Distribution of errors, estimated - real, by Broad Field
Stability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Sci/Math/Eng
Bus/CS
Arts/Lang
Hum/Soc Sci
Displaying errors less than 200%


Distribution of errors, estimated - real, by Student Term of Enrollment
Stability of Employment
Frequency
-100
-50
0
50
100
Percent Error
First Time Student
Advanced Student
Displaying errors less than 200%


Distribution of errors, estimated - real, by Student Income Stability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Low Income Student
High Income Student
Displaying errors less than 200%


Distribution of errors, estimated - real, by Student's Parental Education Stability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Low Parental Education
Not Low Parental Education
Displaying errors less than 200%


Distribution of errors, estimated - real, by Intended Major
Stability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Student's Intended Major Field
All Other Fields
Displaying errors less than 200%


Distribution of errors, estimated - real, by Student Gender
Stability of Employment
Frequency
-100
-50
0
50
100
Percent Error
Females
Males
Displaying errors less than 200%

## Appendix C: Relationship Between Major Choice and Labor Market Outcomes with Covariates

Table C1: Relationship between various individual characteristics and major choice

| | Stated Probability of Choosing a Given Major Category | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Predicted Salary | 0.7404*** | | 0.6075** | 0.5820* | 0.0658 | 1.1924** |
| | (0.2804) | | (0.2986) | (0.2971) | (0.3668) | (0.5479) |
| Pred. Prob. of Employment | | 0.3901** | 0.3649* | 0.2586 | 0.1445 | 0.5090 |
| | | (0.1725) | (0.2126) | (0.2358) | (0.3070) | (0.3478) |
| Pred. Stab. of Employment | | | | 0.2115 | | |
| | | | | (0.1749) | | |
| Female | | | | | 1.3368*** | |
| | | | | | (0.4551) | |
| Hispanic | | | | | 0.9482* | |
| | | | | | (0.5122) | |
| Asian | | | | | -1.1103** | |
| | | | | | (0.5031) | |
| Black | | | | | 1.0652* | |
| | | | | | (0.6324) | |
| Pacific Islander | | | | | 1.4569** | |
| | | | | | (0.5789) | |
| Other Race | | | | | 0.1563 | |
| | | | | | (0.5655) | |
| Low Income | | | | | -0.1499 | |
| | | | | | (0.4569) | |
| No Parent with College Degree | | | | | 0.6783 | |
| | | | | | (0.4321) | |
| HS GPA = D | | | | | -5.1177*** | |
| | | | | | (0.8716) | |
| HS GPA = C | | | | | -4.9720*** | |
| | | | | | (0.6869) | |
| HS GPA = B | | | | | -5.5604*** | |
| | | | | | (0.6383) | |
| HS GPA = A | | | | | -5.8681*** | |
| | | | | | (0.6962) | |
| Constant | -0.1334 | -0.0216 | -0.1332 | -0.1301 | 5.6236*** | -0.4990*** |
| | (0.1461) | (0.1376) | (0.1505) | (0.1500) | (0.9928) | (0.1672) |
| Observations | 772 | 854 | 738 | 738 | 620 | 741 |
| R-squared | 0.0085 | 0.0109 | 0.0152 | 0.0169 | 0.1136 | 0.6650 |

Robust standard errors clustered at the student level are shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Unit of observation is at the student-by- major category level. Reference category is business.

Table C2: Relationship between various individual characteristics and major choice, including predicted enjoyment and grades

| | Stated Probability of Choosing a Given Major Category | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Predicted Salary | 0.7122*** | | 0.7139*** | 0.7353* | 0.4003 | 1.0954** |
| | (0.2426) | | (0.2606) | (0.3797) | (0.2993) | (0.4608) |
| Pred. Prob. of Employment | | 0.2226 | 0.1703 | -0.0308 | -0.0499 | 0.1232 |
| | | (0.1707) | (0.2132) | (0.3641) | (0.2860) | (0.2885) |
| Pred. Stab. of Employment | | | | 0.0982 | | |
| | | | | (0.2516) | | |
| # Fam. & Friends w/ this Major | | | | -0.2298 | | |
| | | | | (0.1876) | | |
| Predicted Enjoyment | 0.7661*** | 0.8100*** | 0.7831*** | 0.7631*** | 0.8163*** | 0.8675*** |
| | (0.0516) | (0.0493) | (0.0509) | (0.0757) | (0.0620) | (0.0800) |
| Predicted Grades | 0.8307*** | 0.6629*** | 0.8245*** | 0.7460*** | 0.6055*** | 0.6566*** |
| | (0.1321) | (0.1289) | (0.1338) | (0.1938) | (0.1753) | (0.2376) |
| Female | | | | | 0.1011 | |
| | | | | | (0.4074) | |
| Hispanic | | | | | 0.9901** | |
| | | | | | (0.3946) | |
| Asian | | | | | -0.4599 | |
| | | | | | (0.4332) | |
| Black | | | | | 2.8516** | |
| | | | | | (1.1299) | |
| Pacific Islander | | | | | 1.1580 | |
| | | | | | (0.8071) | |
| Other Race | | | | | 0.9362** | |
| | | | | | (0.4282) | |
| Low Income | | | | | 0.2493 | |
| | | | | | (0.3929) | |
| No Parent with College Degree | | | | | 0.2576 | |
| | | | | | (0.3579) | |
| HS GPA = D | | | | | -1.0852 | |
| | | | | | (0.8086) | |
| HS GPA = C | | | | | 0.2415 | |
| | | | | | (0.6451) | |
| HS GPA = B | | | | | 0.1356 | |
| | | | | | (0.5925) | |
| HS GPA = A | | | | | 0.0782 | |
| | | | | | (0.6982) | |
| Constant | 0.1048 | -0.0512 | 0.0958 | 0.0789 | -0.4429 | -0.3063* |
| | (0.1511) | (0.1460) | (0.1644) | (0.2844) | (0.9195) | (0.1731) |
| Observations | 618 | 694 | 594 | 267 | 506 | 594 |
| R-squared | 0.4168 | 0.3925 | 0.4249 | 0.4210 | 0.4464 | 0.7905 |

Robust standard errors clustered at the student level are shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Unit of observation is at the student-by-major category level. Business is the reference category.

Appendix D: Non-experimental results with other reference categories

In this appendix we report results from the non-experimental estimates of the relationship between predicted labor market outcomes and major choice using each of the other three broad major categories as the reference. The results (below) show that when the reference category is science or humanities, controls appear to matter more than they do when the reference category is Language/Arts or business. In the case of science, the effect of earnings is not robust to the inclusion of individual controls or student fixed effects. However, these point estimates are not statistically significantly different from estimates with other controls. When using humanities as the reference category, including estimates for expected grades and enjoyment affects the relationship between salary and choice, and including student fixed effects restores a significant relationship between earnings and choice.

On average, earnings are a significant determinant of choice. However, the significance of this effect changes with different reference majors. For some, the relationship is more significant across students (science), and for some the relationship is more significant within students (humanities and language/arts).

Table D1: Relationship between Major Choice and Predicted Labor Market Outcomes, Science as reference category

| | Stated Probability of Choosing a Given Major Category | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Predicted Salary | 0.6775*** | | 0.6753*** | 0.8398*** | 0.5754 | 0.5629 |
| | (0.2127) | | (0.2279) | (0.3112) | (0.3516) | (0.4653) |
| Pred. Prob. of Employment | | 0.2200* | 0.0912 | 0.0588 | -0.0333 | 0.2629 |
| | | (0.1223) | (0.1697) | (0.2189) | (0.2293) | (0.3380) |
| Pred. Stab. of Employment | | | | -0.1812 | | |
| | | | | (0.1341) | | |
| Pred. Enjoyment | 0.8257*** | 0.8273*** | 0.8318*** | 0.8136*** | 0.8289*** | 0.6960*** |
| | (0.0414) | (0.0413) | (0.0410) | (0.0632) | (0.0604) | (0.0933) |
| Predicted Grades | 0.8292*** | 0.7106*** | 0.8422*** | 0.7884*** | 0.7405*** | 0.8928*** |
| | (0.1404) | (0.1315) | (0.1441) | (0.2179) | (0.2130) | (0.2053) |
| # Fam. & Friends w/ this Major | | | | 0.1020 | | |
| | | | | (0.1979) | | |
| Constant | -0.9547*** | -1.1235*** | -0.9658*** | -0.7727*** | -0.1530 | -0.5245*** |
| | (0.1373) | (0.1294) | (0.1401) | (0.1939) | (1.0128) | (0.1049) |
| Individual Controls | | | | | X | |
| Student Fixed Efects | | | | | | X |
| Observations | 620 | 700 | 595 | 284 | 505 | 595 |
| R-squared | 0.4096 | 0.3875 | 0.4134 | 0.4202 | 0.4291 | 0.8683 |

Robust standard errors clustered at the student level are shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Unit of observation is at the student-by-major category level. Reference category is science.

Table D2: Relationship between Major Choice and Predicted Labor Market Outcomes, Humanities as reference category

| | Stated Probability of Choosing a Given Major Category | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Predicted Salary | 0.4749* | | 0.4223 | 0.6112* | 0.1869 | 1.3204*** |
| | (0.2514) | | (0.2597) | (0.3156) | (0.3941) | (0.3908) |
| Pred. Prob. of Employment | | 0.1051 | 0.1742 | -0.0941 | 0.0405 | 0.0561 |
| | | (0.1322) | (0.1777) | (0.2729) | (0.2312) | (0.3155) |
| Pred. Stab. of Employment | | | | 0.1876 | | |
| | | | | (0.1957) | | |
| Pred. Enjoyment | 0.7521*** | 0.7869*** | 0.7708*** | 0.8335*** | 0.7923*** | 0.8683*** |
| | (0.0560) | (0.0538) | (0.0556) | (0.0846) | (0.0729) | (0.0738) |
| Predicted Grades | 0.8577*** | 0.7911*** | 0.8806*** | 0.7329*** | 0.9004*** | 0.7424*** |
| | (0.1410) | (0.1331) | (0.1433) | (0.1959) | (0.2009) | (0.2105) |
| # Fam. & Friends w/ this Major | | | | 0.1125 | | |
| | | | | (0.1808) | | |
| Constant | 0.2651** | 0.3741*** | 0.2776** | 0.2004 | -0.5263 | 0.8251*** |
| | (0.1300) | (0.1228) | (0.1348) | (0.1704) | (0.7075) | (0.1846) |
| Individual Controls | | | | | X | |
| Student Fixed Efects | | | | | | X |
| Observations | 617 | 691 | 593 | 325 | 503 | 593 |
| R-squared | 0.4060 | 0.3882 | 0.4179 | 0.4157 | 0.4489 | 0.7678 |

Robust standard errors clustered at the student level are shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Unit of observation is at the student-by-major category level. Reference category is humanities

Table D3: Relationship between Major Choice and Predicted Labor Market Outcomes, Language/Arts as reference category

| | Stated Probability of Choosing a Given Major Category | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Predicted Salary | 0.6875** | | 0.7331** | 0.6406** | 0.5009 | 0.6910* |
| | (0.2714) | | (0.2931) | (0.3144) | (0.3620) | (0.3948) |
| Pred. Prob. of Employment | | 0.0259 | 0.1071 | 0.0217 | 0.0905 | 0.0902 |
| | | (0.1312) | (0.1786) | (0.2733) | (0.2395) | (0.3225) |
| Pred. Stab. of Employment | | | | 0.0456 | | |
| | | | | (0.1434) | | |
| Pred. Enjoyment | 0.7998*** | 0.8011*** | 0.8090*** | 0.8559*** | 0.8117*** | 0.8301*** |
| | (0.0556) | (0.0558) | (0.0559) | (0.0736) | (0.0705) | (0.0729) |
| Predicted Grades | 0.8361*** | 0.7088*** | 0.8387*** | 0.6982*** | 0.7442*** | 0.8506*** |
| | (0.1319) | (0.1231) | (0.1328) | (0.1773) | (0.1788) | (0.2199) |
| # Fam. & Friends w/ this Major | | | | 0.1041 | | |
| | | | | (0.1832) | | |
| Constant | 0.6659*** | 0.9200*** | 0.6851*** | 0.4373** | 1.3005* | -0.2532 |
| | (0.1389) | (0.1238) | (0.1434) | (0.1786) | (0.7782) | (0.1996) |
| Individual Controls | | | | | X | |
| Student Fixed Efects | | | | | | X |
| Observations | 615 | 685 | 590 | 326 | 500 | 590 |
| R-squared | 0.4045 | 0.3729 | 0.4122 | 0.4134 | 0.4399 | 0.7722 |

Robust standard errors clustered at the student level are shown in parentheses; *** p<0.01, ** p<0.05, * p<0.1.
Unit of observation is at the student-by-major category level. Reference category is Language/Arts.

# Appendix E: Experimental Design

Form A

| | Business, Comp. Sci. & Tech. | | Humanities, Education, Social Sci. | |
|---|---|---|---|---|
| *Prob. Of Employment* | Low | High | Low | High |
| *Salary* | High | Low | Low | High |

Form B

| | Business, Comp. Sci. & Tech. | | Humanities, Education, Social Sci. | |
|---|---|---|---|---|
| *Prob. Of Employment* | Low | High | Low | High |
| *Salary* | Low | High | High | Low |

# Appendix F: Which students react most to information?

In the main text we examined the effects of the provision of experimentally manipulated "true" labor market outcomes on students' stated probability of choosing a given major. In this appendix we examine of the effects of information vary by students' initial probability of choosing a given major. We accomplish this with the following model:

$$\boldsymbol{ln\pi'_{k,i}} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 lnw'_{k,i}} + \theta_{prev.prob.}\beta_2 + \delta_{prev.prob} * lnw'_{k,i}\,\beta_3 + \boldsymbol{\epsilon'_{k,i}} \quad \text{(C1)}$$

Where $\pi_{k,i}$ is the probability for student i to major in field k. $lnw'_{k,i}$, is the expected labor market outcome for major *k* (salary (column (1), probability of employment (column (2) or both (column (3) in Table  and F2). We also include a vector of dummy variables that indicate no, low (<30%), medium (>= 30% and <= 50%) and high (>50%) previous probabilities ($\boldsymbol{\theta_{prev.prob.}}$) and a vector of these dummies interacted with the experimentally manipulated labor market outcomes ($\boldsymbol{\delta_{prev.prob} * lnw'_{k,i}\,\beta_3}$). Thus, the estimated coefficients $\beta_3$ gives an indication of if the effect of the "true" salary differs between students who previously stated no, low, medium, or high probability of choosing a given major. In some models (columns 4-6), we add students' prior beliefs about expected salary and probability of employment. We run these analyses separately for the business (Table F1) and humanities (Table F2) experiments.

In both tables we see that the effects of labor market information seem to be greatest for students who previously stated that there was zero chance they would choose this field; the interaction between a dummy for membership in this group and information about "true" salary are large and significant for humanities (Table F2) and large and sometimes significant for business (Table F1).  In the case of the humanities major, the effects of labor market information are significantly greater for students who previously had a low probability of selecting this major. This is not true for the business major. The results do not qualitatively change with the addition of students' previous beliefs about salary and probability of employment.

Table F1—Experimental Estimates of Effects of Salary on Choice, by Initial Probability, Business

| | Stated Probability of Choosing Business/CS/Technology | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| "True" salary | 0.9664** | | 0.9638* | 0.8330** | | 0.8061** |
| | (0.4817) | | (0.4969) | (0.3605) | | (0.3297) |
| Prev. Prob.: Zero | -13.2880* | -1.7277*** | -12.1439* | -16.7882*** | -1.9482*** | -15.9942*** |
| | (6.8784) | (0.5266) | (7.0481) | (6.0795) | (0.5358) | (6.1284) |
| Prev. Prob.: Low | 1.1052 | -0.5118 | 1.2403 | -1.0462 | -0.4215 | -1.5462 |
| | (7.6568) | (0.5824) | (7.8048) | (6.6106) | (0.5743) | (6.1468) |
| Prev. Prob.: Med. | 2.5178 | 0.6385 | 3.2394 | 0.7683 | 0.8156** | 2.0351 |
| | (6.4984) | (0.4635) | (6.1875) | (4.8423) | (0.4083) | (4.4828) |
| Prev. Prob.: Zero * "True" Salary | 1.0191 | | 0.9433 | 1.3115** | | 1.2676** |
| | (0.6355) | | (0.6540) | (0.5616) | | (0.5542) |
| Prev. Prob.: Low * "True" Salary | -0.1671 | | -0.1689 | 0.0179 | | 0.0949 |
| | (0.7065) | | (0.7208) | (0.6083) | | (0.5646) |
| Prev. Prob.: Med * "True" Salary | -0.2051 | | -0.2573 | -0.0552 | | -0.1437 |
| | (0.5975) | | (0.5824) | (0.4412) | | (0.4003) |
| "True" Prob. Employment | | -0.1396 | -0.0360 | | -0.0954 | -0.2281 |
| | | (0.5991) | (0.6031) | | (0.6053) | (0.4008) |
| Prev. Prob.: Zero * "True" Prob. Employ | | 1.0979 | 0.6079 | | 0.9895 | 0.8732 |
| | | (0.7975) | (0.7944) | | (0.8169) | (0.6737) |
| Prev. Prob.: Low * "True" Prob. Employ | | 0.3469 | 0.2140 | | 0.2770 | 0.4227 |
| | | (0.8761) | (0.8756) | | (0.8745) | (0.6863) |
| Prev. Prob.: Med * "True" Prob. Employ | | 0.6197 | 0.2952 | | 0.5030 | 0.5054 |
| | | (0.7435) | (0.7072) | | (0.6739) | (0.4866) |
| Prev. Belief of Salary | | | | -0.1171 | | 0.0014 |
| | | | | (0.2177) | | (0.2337) |
| Prev. Belief of Prob. Employ. | | | | | -0.0749 | -0.2178 |
| | | | | | (0.2424) | (0.2299) |
| Constant | -11.7627** | -1.4280*** | -11.7541** | -8.7484* | -1.4703*** | -9.9923** |
| | (5.2307) | (0.3952) | (5.2938) | (4.4564) | (0.4403) | (4.4061) |
| Observations | 431 | 431 | 431 | 392 | 402 | 376 |
| R-squared | 0.2499 | 0.1914 | 0.2534 | 0.3164 | 0.2370 | 0.3504 |

Robust standard errors, clustered at the student level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Low previous probability of choosing is defined as less than 30%, medium is defined as between 30% and 50%. High (>50%) previous probability of choosing is the reference category. Observations are at the student-by-scenario level.

Table F2—Experimental Estimates of Effects of Salary on Choice, by Initial Probability, Humanities

| | Stated Probability of Choosing Humanities/Education/Behavioral & Social Science | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| "True" salary | 0.8373** | | 0.8159** | 0.7536** | | 0.6608* |
| | (0.3316) | | (0.4007) | (0.3381) | | (0.3860) |
| Prev. Prob.: Zero | -23.8803*** | -2.2292*** | -25.4056*** | -28.5206*** | -2.2879*** | -31.0687*** |
| | (7.2236) | (0.4744) | (7.4341) | (7.4372) | (0.4777) | (7.4515) |
| Prev. Prob.: Low | -13.7479** | -0.2870 | -13.3575** | -13.9665** | -0.3505 | -14.2078** |
| | (5.8467) | (0.3750) | (6.2287) | (5.7750) | (0.3748) | (6.0703) |
| Prev. Prob.: Med. | -4.6911 | -0.6153 | -4.9753 | -3.4125 | -0.2183 | -5.1433 |
| | (7.2062) | (0.4664) | (7.8548) | (6.9006) | (0.3926) | (6.8994) |
| Prev. Prob.: Zero * "True" Salary | 1.9470*** | | 2.1407*** | 2.3752*** | | 2.6597*** |
| | (0.6598) | | (0.6844) | (0.6795) | | (0.6834) |
| Prev. Prob.: Low * "True" Salary | 1.1907** | | 1.1854** | 1.2184** | | 1.2739** |
| | (0.5228) | | (0.5669) | (0.5177) | | (0.5526) |
| Prev. Prob.: Med * "True" Salary | 0.3691 | | 0.4007 | 0.2707 | | 0.4536 |
| | (0.6548) | | (0.7322) | (0.6267) | | (0.6429) |
| "True" Prob. Employment | | -0.2844 | -0.0595 | | -0.2488 | -0.2026 |
| | | (0.2986) | (0.3519) | | (0.2988) | (0.3380) |
| Prev. Prob.: Zero * "True" Prob. Employ | | 0.8178 | 1.0054* | | 0.8788 | 1.2781** |
| | | (0.6166) | (0.5998) | | (0.6279) | (0.5982) |
| Prev. Prob.: Low * "True" Prob. Employ | | 0.8789* | 0.5744 | | 0.8813* | 0.8789* |
| | | (0.4968) | (0.4970) | | (0.5020) | (0.4837) |
| Prev. Prob.: Med * "True" Prob. Employ | | 0.1026 | 0.1019 | | 0.3202 | 0.4279 |
| | | (0.5899) | (0.6415) | | (0.5064) | (0.5628) |
| Prev. Belief of Salary | | | | -0.2381 | | -0.0933 |
| | | | | (0.2177) | | (0.2217) |
| Prev. Belief of Prob. Employ. | | | | | -0.4896*** | -0.4922*** |
| | | | | | (0.1322) | (0.1379) |
| Constant | -9.8974*** | -0.9569*** | -9.6996** | -6.4994 | -1.4046*** | -7.5500 |
| | (3.6784) | (0.2436) | (4.2876) | (4.2818) | (0.2898) | (4.6797) |
| Observations | 410 | 410 | 410 | 373 | 387 | 361 |
| R-squared | 0.3053 | 0.2041 | 0.3179 | 0.3384 | 0.2476 | 0.3933 |

Robust standard errors, clustered at the student level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Low previous probability of

# Appendix G: exploring the differences between the experimental vs. non-experimental impact of labor market outcomes on major choice

In this appendix we present examine two potential explanations for why the results from the experimental analyses differ from the results from the non-experimental analyses. In both cases, we do not find satisfactory explanations for the differences in results.

**Relative vs. absolute probability of choosing a major**

First, we examine if the differences are due to modelling decisions. Similar to our strategy for the non-experimental estimates, in this section we predict the difference in probabilities of selecting a major using differences in experimentally assigned labor market outcomes.

$$\left(ln\pi_{h,i} - ln\pi_{b,i}\right) = \beta_0 + \beta_1(lnw_h - lnw_b) + \varepsilon_{k,i}$$

where $\pi_{h,i}$ is the probability student i assigns to majoring in humanities. The subscript b stands for business as everything is relative to business. The key explanatory variable is $lnw_h - lnw_b$, i.e. the difference in salary relative to business.

Table G1: Experimental data, relative to business

| Stated Probability of Choosing a Given Major Category | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| "True" salary, rel. to business | 1.6568*** | | 1.6728*** |
| | (0.1672) | | (0.1682) |
| "True" Prob. of Employ., rel. to business | | -1.5060 | 0.9204 |
| | | (1.7057) | (1.5368) |
| Constant | 0.4221*** | 0.5417*** | 0.4557*** |
| | (0.1475) | (0.1724) | (0.1528) |
| Observations | 397 | 397 | 397 |
| R-squared | 0.1991 | 0.0020 | 0.1998 |

Robust standard errors, clustered at the student level, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1. Observations are at the student-by-scenario level.

The results in Table G1 make it clear that the large estimated differences in effects between the non-experimental and experimental results are not due to the different specifications.

## Focusing on humanities and business in the non-experimental data

In the experiment, we ask students about probability of majoring in only business and humanities majors, whereas in the non-experimental section we ask students about all four categories of majors. Thus, it could be that the difference in results are due to the different categories. We examine this possibility below.

In Table G2 we break the non-experimental results by major category. In the first four columns in Table G2, we regress the log probability of choosing that major on the log salary in the major. We can see that differences in beliefs about salary in a single major across students are not a big predictor of major choice, as salary is never significant. Furthermore, for business and language and arts, the point estimate is also close to zero. The strongest evidence for a role of beliefs is for science, where students with 10% higher estimates of salary in science are 26% more likely to major in science; but this estimate is still not statistically significant.

In columns 5-7, we regress the probability of majoring in that field relative to business. This is now asking if the probability of majoring in a certain field is influenced by relative beliefs. The point estimates are now more definitely positive, showing that students who have higher gaps in salary between a major and business are relatively more likely to major in that. Still, none are statistically significant.

Table G2: Relationship between predicted salary and choice for each broad major category

| | Probability of Choosing a Given Major Category | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Broad Major Category: | business | humanities | language/arts | science | humanities | language/arts | science |
| Predicted Salary | -0.0259 | 0.1160 | 0.0109 | 0.2551 | | | |
| | (0.2127) | (0.2307) | (0.2289) | (0.2017) | | | |
| Pred. Sal., Rel. to Business | | | | | 0.2424 | 0.3391 | 0.6900 |
| | | | | | (0.5095) | (0.5261) | (0.4806) |
| Constant | -3.8711* | -4.6891* | -4.1501* | -5.0767** | 0.7202** | 0.3040 | 1.7757*** |
| | (2.2711) | (2.4088) | (2.3661) | (2.1586) | (0.2971) | (0.3021) | (0.2735) |
| Observations | 266 | 264 | 263 | 269 | 258 | 259 | 258 |
| R-squared | 0.0001 | 0.0012 | 0.0000 | 0.0065 | 0.0010 | 0.0019 | 0.0061 |

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

It is also interesting to note how low the R-squared is in the table, indicating that different beliefs about salary across students are not important drivers of major choice, but relative beliefs explain major choice better than absolute beliefs. There is thus strong evidence that the differences between the experimental and non-experimental results is not driven by the different categories of majors that we ask about.

How can we understand the fact that choices relative to humanities are better predicted than absolute choices? In some way, it makes sense that differences in beliefs explain more than absolute beliefs, because students are choosing between majors, so saying that a student believes business earns a lot

without telling us how much the student thinks other majors earn is leaving out an important part of the picture.

We then repeat Table G2 but add controls for enjoyment and grades. Table G3 below shows the results. The R-squared has increased quite a bit compared to Table G2, and the explanatory power again appears greater in relative terms. In this case, salary is not significant, except when considering science in isolation.

Table G3: Relationship between predicted salary, enjoyment, and grades and choice for each broad major category

| | Probability of Choosing a Given Major Category | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Broad Major Category: | business | humanities | language/arts | science | humanities | language/arts | science |
| Predicted Salary | 0.1120 | 0.1573 | 0.0050 | 0.2575* | | | |
| | (0.1992) | (0.2265) | (0.3143) | (0.1440) | | | |
| Pred. Enjoyment | 0.5976*** | 0.6587*** | 0.6765*** | 0.6770*** | | | |
| | (0.0627) | (0.0829) | (0.0533) | (0.1002) | | | |
| Predicted Grades | 0.4995*** | 0.0769 | 0.3269* | 0.7611*** | | | |
| | (0.1652) | (0.2204) | (0.1802) | (0.2201) | | | |
| Pred. Sal., Rel. to Business | | | | | 0.1111 | 0.6322 | 0.5443 |
| | | | | | (0.4392) | (0.4831) | (0.3615 |
| Pred. Enjoy., Rel. to Bus. | | | | | 0.6987*** | 0.7443*** | 0.7956** |
| | | | | | (0.0879) | (0.0954) | (0.0746 |
| Pred. Grades, Rel. to Business | | | | | 0.9502*** | 0.8455*** | 0.9538** |
| | | | | | (0.2277) | (0.1867) | (0.2862 |
| Constant | -5.7947*** | -4.3187* | -4.2906 | -7.2254*** | -0.3104 | -0.4201 | 0.7452** |
| | (2.1871) | (2.2891) | (3.2510) | (1.9512) | (0.2836) | (0.2766) | (0.2528 |
| Observations | 226 | 224 | 217 | 244 | 205 | 205 | 208 |
| R-squared | 0.3079 | 0.2086 | 0.2255 | 0.3729 | 0.4613 | 0.4544 | 0.3720 |

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

While we learned interesting things here, it is clear that the discrepancy between the experimental and nonexperimental results is not because the experimental results focus on business and humanities alone; indeed, these fields are not really special compared to other fields when looking at nonexperimental data.

## Understanding the bias in nonexperimental estimates due to the correlation of beliefs about labor market outcomes with other determinants of major choice

The basic explanation here is that people who hold higher salary beliefs about some majors relative to others may also hold other unfavorable beliefs (e.g. about enjoyment and grades), which dampens the effect of salary on their choice.

Based on Table G4 below, it appears that it is the enjoyment factor that dampens the impact of salary on major choice, while just controlling for grades does not change the effect of salary on major choice. Interestingly, controlling for both grade and enjoyment (col. 4) leaves a greater impact of salary on major choice than just controlling for enjoyment. The fact that the enjoyment factor dampens the coefficient on salary is interesting in light of the positive but insignificant (P-val=0.28) correlation between enjoyment and salary beliefs (as shown in results for broad major categories above). It appears that some of the positive effect of salary on choice loads onto enjoyment.

Table G4: Relationship between relative beliefs about majors and choice

| | Probability of Choosing a Given Major Category | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Pred. Sal., Rel. to Business | 0.7404*** | 0.5813** | 0.9191*** | 0.7122*** |
| | (0.2804) | (0.2386) | (0.2563) | (0.2426) |
| Pred. Enjoy., Rel. to Business | | 0.9453*** | | 0.7661*** |
| | | (0.0375) | | (0.0516) |
| Pred. Grades, Rel. to Business | | | 1.8144*** | 0.8307*** |
| | | | (0.1066) | (0.1321) |
| Constant | 1.0332*** | 0.2241 | 0.3371** | 0.1048 |
| | (0.1632) | (0.1439) | (0.1656) | (0.1511) |
| Observations | 775 | 690 | 621 | 618 |
| R-squared | 0.0088 | 0.3966 | 0.2600 | 0.4168 |

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The coefficient on salary is biased by omitted variables that determine choice and are correlated either positively or negatively with salary. If we trust our experimental estimates, we must infer that salary is correlated with some determinants that have a negative impact on choosing a certain major. However, these determinants are not grades since controlling for grades doesn't change the impact of salary, and it's not about enjoyment either since controlling for enjoyment rather decreases the impact of salary (if we control for the variable that is responsible for the downward bias in the salary coefficient, the salary coefficient should go up, nor down, hence the inference that enjoyment is not such a variable).