How Do College Students Form Expectations?*

Basit Zafar†

June 29, 2009

Abstract

Students rely on their subjective expectations when choosing a college major. Understanding this process of expectations formation is crucial for policy recommendations. This paper focuses on how college students form expectations about various major-specific outcomes. For this purpose, I collect a unique *panel* dataset of Northwestern University undergraduates that contains their subjective expectations about major-specific outcomes. Though students tend to be overconfident about their future academic performance, I find that they revise their expectations about various major-specific outcomes in systematic ways. For example, students who receive extremely positive information (about their ability) revise upward their prediction of short-term future GPA, and similarly individuals who receive very negative information revise their GPA beliefs downward. Furthermore, students seem to update their probabilistic beliefs in a manner consistent with Bayesian analysis: Prior beliefs about outcomes realized in college tend to be fairly precise, while new information influences prior beliefs about outcomes realized in the workplace. Moreover, students who are more uncertain about the major-specific outcomes in the initial survey make greater absolute revisions in their beliefs. Finally, I present evidence that learning plays a role in the decision to switch majors. Negative revisions in beliefs about graduating in 4 years, enjoying coursework, and expected salary are associated with dropping a major.

JEL Codes: D8, I2, J1, J7

Keywords: college majors; learning; expectations

*I am indebted to Charles Manski and Christopher Taber for extremely helpful discussions and comments. I also thank Meta Brown, Yonatan Eyal, Carlos Madeira, Wilbert van der Klauuw, Bob Willis, and participants at the 2009 CIPREE Workshop on Subjective Expectations in Econometric Models, the Econometric Society 2009 North American Summer Meeting, Northwestern's Labor Lunch, and the Federal Reserve Bank of New York Brown Bag Lunch for feedback and suggestions. Financial Support from a Northwestern University Graduate Research Grant is gratefully acknowledged. I thank all those involved in providing the datasets used in this paper. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System as a whole. All errors that remain are mine.

†Research and Statistics, Federal Reserve Bank of New York, 33 Liberty Street, New York, NY 10045. E-mail: Basit.Zafar@ny.frb.org

1 Introduction

Schooling choices are made under uncertainty–uncertainty about personal tastes, individual abilities, and realizations of choice-related outcomes. Although some theoretical work incorporates the uncertainty associated with schooling choices (Manski, 1989; Altonji, 1993; Malamud, 2007), there is little empirical work in this area (exceptions include Bamberger, 1986; Arcidiacono, 2004; Cunha, Heckman, and Navarro, 2004; Stange, 2008). Moreover, existing empirical studies make non-verifiable assumptions on expectations, assume individuals are rational and form expectations in the same way, and use choice data to infer decision rules conditional on the maintained assumptions about expectations. This approach is problematic for several reasons. First, there is little reason to think that individuals form their expectations in the same way.[1] Second, observed choices may be consistent with several combinations of expectations and preferences (Manski, 1993). Third, the information-processing rule has varied considerably among studies of schooling behavior, and it's not clear which is the correct one to use (given that individuals may use idiosyncratic rules to form their beliefs). A solution to this identification problem is to directly elicit subjective beliefs (Manski, 2004) and incorporate them into choice models (Delavande, 2008a; Zafar, 2008). However, to predict behavior in a new scenario that could possibly affect expectations in nonobvious ways, one would need to understand the process of expectations formation. Moreover, once education is treated as a sequential choice, it is clear that understanding how students perceive pecuniary and non-pecuniary returns to a choice is a prerequisite for informed analysis of schooling decisions. Because few studies collect data on subjective beliefs, and even fewer follow the same respondents over time, little is known about how students form expectations in the context of schooling choices.[2] The main goal of this paper is to fill this gap in the literature.

This paper examines how college students form their expectations about outcomes related to choice of major. Since studying revisions of expectations offers the best hope for learning about the process of expectations formation (Dominitz, 1998), I focus on how students revise

[1] In fact, Madeira (2007) and Arcidiacono et al. (2009) find that black and Hispanic students tend to be more optimistic about their academic performance relative to other groups. Stinebrickner and Stinebrickner (2008) also find that students in the lower part of the ability distribution tend to be more optimistic about their performance. Similarly, in a laboratory experiment, Niederle and Vesterlund (2007) find that men tend to more overconfident about their ability than are women.

[2] There are two exceptions: Stinebrickner and Stinebrickner (2008), using a panel of subjective beliefs about academic ability from low-income college students, study how students update their beliefs and how these beliefs affect their college drop-out decision. Madeira (2007) uses the Beginning School Study to analyze how parents and students (starting at first grade) forecast their academic scores.

their beliefs about major-specific outcomes. For this purpose, I designed and conducted two surveys that elicited subjective expectations from Northwestern University undergraduates regarding their choice of major. The first survey, administered to students in the early part of their sophomore year, collected details on students' demographics and subjective beliefs about major-specific outcomes; these data were used to estimate a choice model of college majors (Zafar, 2008). The second survey, conducted about a year after the first, collected data on how individuals revise their beliefs for major-specific outcomes. Since understanding the mechanisms that lead individuals to revise their beliefs also requires data that directly identify new information, the surveys also contained questions that identified *some* of the new information about their academic ability that individuals had acquired between the two surveys. The major-specific outcomes for which beliefs were elicited include both outcomes realized in college and those realized in the workplace. Examples of the former include graduating in 4 years, enjoying the coursework, and having parents approve of the choice, while examples of the latter include outcomes like finding a job upon graduation and being able to reconcile work and family while working at the jobs. The data are described in Section 2.

Section 3 of the paper analyzes how and why students update their beliefs. Analysis of the panel on beliefs shows that students, in response to new information, modify their beliefs systematically and somewhat rationally. This finding matches with conclusions reached in Bernheim (1988), Dominitz (1998), Hurd and McGarry (2002), and Lochner (2007), all of whom find that expectations are responsive to new information. However, existing studies, due to lack of data that identify new information, cannot pin down the causal explanation for the revision in expectations.[3] In order to understand the mechanisms that lead to revision of beliefs, the first survey elicited beliefs of future GPA over a horizon of one year; these GPA realizations were observed at the time of the second survey. Comparing the beliefs with actual realizations of GPA allows me to develop an "information metric" that identifies some new information about their own academic ability that students acquire between the two surveys. Based on beliefs reported in the first survey, I find that students, on average, tend to be overconfident about their academic performance. However, they adjust their beliefs in response to the new information appropriately. Using local linear regressions, I find that students who

[3]Though some laboratory and field experiments have studied how agents update their beliefs with new information (Viscusi and O'Connor, 1984; El-Gamal and Grether, 1995; Delavande, 2008b; and Houser, Keane, and McCabe, 2004), these studies use extremely stylized settings and focus on learning over short time horizons. It is yet to be seen whether their results would be evident in less standardized environments or over longer time periods.

receive positive information revise upward their predictions of short-term future GPA only if the information content is very positive, and similarly those who receive negative information revise their predictions downward only if the information content is very negative. Students who receive information that is in the intermediate range don't revise their short-term GPA beliefs. Moreover, no effect is found on long-term GPA expectations. I also find a negative relationship between the information metric and revisions in beliefs about number of hours per week that students expect to spend on coursework. This result suggests that students view ability and effort as substitutes in the production of their achievement, which is consistent with Stinebrickner and Stinebrickner (2007), who find a causal effect of studying on academic performance. I do not find a systematic relationship between the information metric and revisions in beliefs for outcomes associated with the workplace.

The updating process is characterized more formally in Section 4. The analysis reveals that priors (beliefs reported in the initial survey) for outcomes such as approval of parents and graduating in 4 years are fairly precise, and that individuals don't revise them by as much as they revise their priors for outcomes realized in the workplace (for example, expected income at the jobs and finding a job upon graduation). These findings are consistent with students adopting a Bayesian learning approach. For outcomes associated with college, one would expect students to have fairly precise information at the time of the initial survey. Conversely, for outcomes realized in the workplace, one would expect students to receive useful information in the period between the two surveys. I also find that individuals who are more uncertain about the major-specific outcomes in the initial survey make greater absolute revisions in their beliefs, and that the population variance of beliefs increases over time for most outcomes.

Over time, students may change their schooling choices (drop out of college or change their field of study) as they learn about their ability, tastes, and quality of match. Dropouts are rare in the current setting: 93% of Northwestern University undergraduate students graduate with a degree within five years of first enrolling. Instead, the phenomenon of switching majors is more common: 12% of the students in my sample switch majors between the two surveys.[4] The analysis in Section 5 suggests that learning plays a role in the decision to switch majors. While I don't find a significant role for the information metric or realized GPA changes in the

[4]Switching of majors is a common occurrence in other settings as well. For example, Arcidiacono (2004) finds that 18% of the students in the NLS72 who attend college switch majors. Similarly, Altonji (1993) documents the discrepancy between planned majors and actual majors.

decision to switch majors, there is evidence that negative revisions in beliefs about graduating in 4 years, enjoying coursework, and expected salary are associated with dropping a major.

Since the literature on schooling choices focuses on monetary returns as the main determinant of the decision, Section 6 explores students' expectations about starting salaries conditional on major. Students seem to be aware of the relative earnings differences across majors (several cross-sectional studies have elicited subjective expectations about monetary returns in the context of higher education: Freeman, 1971; Smith and Powell, 1990; Blau and Ferber, 1991; Betts, 1996; and Dominitz and Manski, 1996). I find that students are more likely to be better informed about expected salaries in the various majors if 1) they are majoring in that field and are more certain about pursuing it, or 2) they have a college-educated father. A more notable finding is that students have a tendency to err in the same direction over time, i.e., students who overestimate or underestimate starting salaries in the initial survey are likely to do the same in the follow-up survey as well, suggesting a persistence in beliefs over time.

Finally, Section 7 of the paper concludes.

2 Data

The data used in this study come from two surveys that were administered to a sample of students in Northwestern University's undergraduate class of 2009. The first survey was administered to students in the early part of their sophomore year over the period from November 2006 to February 2007. I denote this as the *Fall 2006* or *initial* survey for the empirical analysis. Since Northwestern University requires students to officially declare their majors by the beginning of their junior year, the timing of the initial survey corresponds to the period when students are actively thinking about which major to choose. The second survey was administered to a subset of the initial survey-takers at the beginning of their junior year, when students had presumably settled on their final majors.[5] The survey spanned the period from November 2007 to February 2008. I denote it as the *Fall 2007* or *follow-up* survey.

Respondents for the initial survey were recruited by flyers posted around campus and by e-mailing a sample of eligible sophomores whose e-mail addresses were provided by the Northwestern Office of the Registrar. Prospective participants were told that the survey was about the choice of college majors and that they would receive $10 for completing the 45-minute

[5] Students can still change their major after their sophomore year, but they have to go through a formal process to do so.

electronic survey. Respondents were required to come to the Kellogg Experimental Laboratory to take the electronic survey.

A total of 161 sophomores took the first survey, 92 of whom were females. The 45-minute survey consisted of three parts. The first part collected demographic and background information (including parents' and siblings' occupations and college majors, source of college funding, etc.). The second part collected data relevant for the estimation of the choice model (see Zafar, 2008). The third part collected beliefs about future GPA at different time horizons. At the end of the survey, respondents were asked if they were willing to participate in a follow-up survey in a year's time.

Of the 161 respondents who took the initial survey, 156 agreed to be contacted for the follow-up. About a year after the first survey, individuals who gave their consent were contacted by e-mail for the follow-up; the e-mail summarized the findings of the initial survey and the purpose of the follow-up. Students were told that they would be compensated $15 for the 1-hour electronic survey. The follow-up was administered in the PC Laboratory located in the Northwestern Main Library.

Of the 156 initial survey respondents, 117 (75%) took the follow-up survey. The first column of Table 2 shows the characteristics of individuals who took the follow-up survey. For comparison, characteristics of the initial sample and the actual sophomore population are shown in columns (2) and (3), respectively. Respondents to the follow-up survey seem similar to the initial survey respondents in most aspects. Even though the average GPA of follow-up respondents is higher than that of the initial survey-takers, the difference is not statistically significant. Table 3 shows that the distribution of majors in the Weinberg College of Arts and Sciences (WCAS) for the students taking the two surveys is similar, suggesting no differential attrition by field of study. As shown in Table 2, students of Asian ethnicity are overrepresented in the survey samples (both in the initial and follow-up survey) relative to their population proportion. Survey-takers, especially males, have higher average GPAs than their population counterparts. However, for the purposes of this study, it's the selection into the follow-up survey that would be of concern. Based on observables, I don't find any selection in who decides to take the follow-up survey. To the extent that certain ethnicities are overrepresented in my sample relative to the underlying population, this should bias the results only if one believes that the process of belief updating and learning is differentially affected by these traits. Since my sample overrepresents

Asians, for robustness purposes I repeat the analysis in the paper by excluding this group. The results do not change quantitatively.

The follow-up survey consisted of two parts. The first part focused on how individuals revise their beliefs about major-specific outcomes. While the initial survey elicited beliefs about outcomes associated with all majors in the individual's choice set (which could be 8 or 9 majors),[6] the follow-up survey elicited beliefs for major-specific outcomes only for three different major categories in the individual's choice set. Beliefs about the major-specific outcomes were elicited for: 1) the major that the individual was pursuing at the time of the follow-up survey (one's most preferred major or current major), 2) the individual's second major (or the second most preferred major at the time of the follow-up survey if the student did not have a second major), and 3) a major that the individual had once pursued but was no longer pursuing (if this was not applicable, beliefs were elicited for the least preferred major in the individual's choice set at the time of the follow-up survey). The second part of the survey collected data on the individuals' GPA at different points in the past, as well as their beliefs about their academic performance at different points in the future. Individuals were also requested to upload their transcripts; only 41 respondents (35%) permitted access to their transcript data, and hence these data are not used in the analysis.

The set of major-specific outcomes for which beliefs were elicited can be classified as outcomes realized in college, denoted by the vector **a**, and outcomes realized in the workplace, denoted by the vector **c**. The vector **a** includes the outcomes:

a_1 successfully completing (graduating) a field of study in 4 years

a_2 graduating with a GPA of at least 3.5 in the field of study[7]

a_3 enjoying the coursework

a_4 hours per week spent on the coursework

a_5 parents approve of the major

while the vector **c** consists of:

c_1 obtain an acceptable job immediately upon graduation

c_2 enjoy working at the jobs available after graduation

c_3 are able to reconcile work and family while at the available jobs

[6] The College of Arts and Sciences at Northwestern University consists of 41 majors. Similar majors were pooled together. Table 1 shows the categorization of majors.

[7] This outcome is meant to capture the student's belief about academic ability in a major. The cutoff of 3.5 for graduating GPA was arbitrary.

c_4 hours per week spent working at the available jobs

c_5 social status of the available jobs

c_6 income at the available jobs

Note that $\{a_r\}_{r=\{1,2,3,5\}}$ and $\{c_q\}_{q=\{1,2,3\}}$ are binary, while outcomes a_4 and $\{c_q\}_{q=\{4,5,6\}}$ are continuous.[8] The survey elicited the probability of the occurrence of the binary outcomes, i.e., $P_{ikt}(a_r = 1)$ for $r = \{1, 2, 3, 5\}$ and $P_{ikt}(c_q = 1)$ for $q = \{1, 2, 3\}$. Expected value was elicited for the continuous outcomes, i.e., $E_{ikt}(a_4)$ and $E_{ikt}(c_q)$ for $q = \{4, 6\}$. As mentioned earlier, the initial survey elicited these beliefs for *all* majors in the individual's choice set, while the follow-up survey elicited them for three different major categories in the individual's choice set.

Questions eliciting the subjective probabilities of major-specific outcomes were based on the use of percentages. An advantage of asking probabilistic questions relative to approaches that employ a Likert scale or a simple binary response (yes/no or true/false) is that responses are interpersonally comparable and allow the respondent to express uncertainty (see Manski, 2004, for an overview of the literature on subjective expectations). As is standard in studies that collect subjective data, a short introduction was read and handed to the respondents at the start of the survey. The wording of the introduction was similar to that in Delavande (2008a). An excerpt of the survey containing the introduction and list of questions dealing with the major-specific outcomes is presented in the Appendix. The full survey questionnaire is available on request from the author.

It would be impossible to describe patterns in the responses for all outcomes. Table 4 presents only the subjective belief distributions reported in both surveys for graduating with a GPA of at least 3.5 in one's current major and one's least preferred major. The table shows that respondents use the entire scale from zero to 100. Respondents tend to round off their responses to the nearest 5, especially for answers not at the extremes. There is a concern that respondents might answer 50% when they want to respond to the interviewer, but are unable to make any reasonable probability assessment of the relevant question (Bruine de Bruin et al., 2000). However, the 50% response is not the most frequent one in the majority of the cases. Over time, it seems that individuals tend to revise downward their beliefs for graduating with

[8] Social status of available jobs, c_5, was elicited as an ordinal ranking. In hindsight, this question should have been asked in terms of the probabilistic chance of obtaining a high-status job, since the ordinal ranking does not reveal the respondent's uncertainty about the outcome.

a GPA of at least 3.5 for both their current major as well as their least preferred major. For example, in the initial survey, nearly half of the respondents believed there was a greater than 80% chance of graduating with a GPA of at least 3.5 in their current major. In the follow-up survey, the fraction of respondents who believed that to be the case had dropped to about 30%. The table also shows the heterogeneity in beliefs, which questions the accuracy in the literature of restrictions imposed on expectations. The next section explores how students revise their beliefs.

3 Updating Beliefs

One way to understand the process of how individuals form expectations is to study how expectations are revised in response to new information. This area remains relatively unexplored because studying this question requires following individuals over time and obtaining data that directly identify new information. Studies have found that expectations tend to be responsive to changes in the environment, but they cannot determine the causality since the data do not directly identify the new information.[9] The survey questionnaires included questions intended to identify changes in the student's information set. Using responses to these questions, this section analyzes how students revise their beliefs.

3.1 Revisions of GPA beliefs

I first outline a simple model of belief updating. Let X_{it} be individual i's expectation at time t about the value of a variable $\mathbf{X}$ that would be realized at some point in the future. Moreover, let Ω_{it} denote i's information set at time t. For simplicity, I assume that $\mathbf{X}$ is a binary event so that[10]

$$X_{it} = E(\mathbf{X}|\Omega_{it}) = \Pr(\mathbf{X} = 1|\Omega_{it}).$$

Similarly, X_{it+1} is i's expectation about the value of $\mathbf{X}$ at time $t+1$. Individuals are assumed to use all available information in forming expectations; therefore, revisions of expectations are determined solely by new information. I further assume that, at time $t+1$, the individual has access to all information that was available at time t. Therefore, $\Omega_{it+1} = (\Omega_{it}, \omega_{it+1})$, where

[9] For example, Dominitz (1998) finds that revisions to expectations of future earnings are associated with earnings that respondents realize between interviews. Smith et al. (2001) find that HRS respondents revise their longevity expectations sensibly in response to health shocks. Hurd and McGarry (2002) find that individuals revise their survival probabilities downward in response to the onset of cancer or the death of one's spouse. Lochner (2007) finds that individuals revise their arrest probabilities downward if, for example, a sibling engages in a crime.

[10] The same logic also applies to continuous outcomes.

ω_{it+1} is *new* information that becomes available to i between time t and $t+1$. It follows that

$$E(X_{it+1}|\Omega_{it}) = E[E(\mathbf{X}|\Omega_{it}, \omega_{it+1})|\Omega_{it}] = E(\mathbf{X}|\Omega_{it}) = X_{it},$$

which implies that

$$\Pr(\mathbf{X} = 1|\Omega_{it+1}) = \Pr(\mathbf{X} = 1|\Omega_{it}) + \varepsilon_{it+1}, \tag{1}$$

where $E(\varepsilon_{it+1}|\Omega_{it}) = 0$, i.e., ε_{it+1} is a function of new information that becomes available after time t. Equation (1) states that the change in expectations between time t and $t+1$ about some event $\mathbf{X}$ that is realized at some point in the future is a function of new information that becomes available after time t.

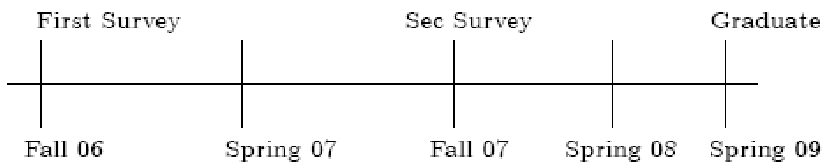

Figure 1: Timeline

In the context of this study, period t refers to the first survey, Fall 2006, and period $t+1$ refers to the follow-up survey, Fall 2007 (see Figure 1 for a visual depiction of the timeline). $\mathbf{X} = 1$ refers to the binary event that the GPA at the end of Spring 2008 (which is realized after the individual takes the follow-up survey) is above a certain threshold. In this case, the threshold is the individual's GPA at the time of the initial survey, so $\Pr(\mathbf{X} = 1|\Omega_{it}) = \Pr(\text{Spring 2008 GPA}_i > \text{Fall GPA}_i|\Omega_{it})$, where Fall GPA$_i$ is the individual's GPA at the time of the initial survey.[11] So $\Pr(\mathbf{X} = 1|\Omega_{it+1}) - \Pr(\mathbf{X} = 1|\Omega_{it})$ is the change in i's subjective belief between the Fall 2006 and Fall 2007 surveys about her Spring 2008 GPA being above her Fall 2006 GPA.

Panel A of Figure 2 depicts the local linear regression estimates of the change in Spring 2008 GPA beliefs on the change in the individual's GPA between the two surveys.[12] The figure also presents the distribution of realized GPA change between the two surveys. Individuals experience GPA changes that vary in the range of -0.45 to 0.4, with -0.01 being the mean. Revisions of Spring 2008 GPA expectations seem to be positively related to changes in realized GPA.

[11]Depending on when the individual took the initial survey, Fall GPA$_i$ refers to the individual's GPA at the *beginning* of Fall 2006 or at the *end* of Fall 2006.

[12]I use a local linear regression estimator instead of a Kernel regression since this avoids the boundary problem. I experimented with different bandwidths, but the figures did not change much.

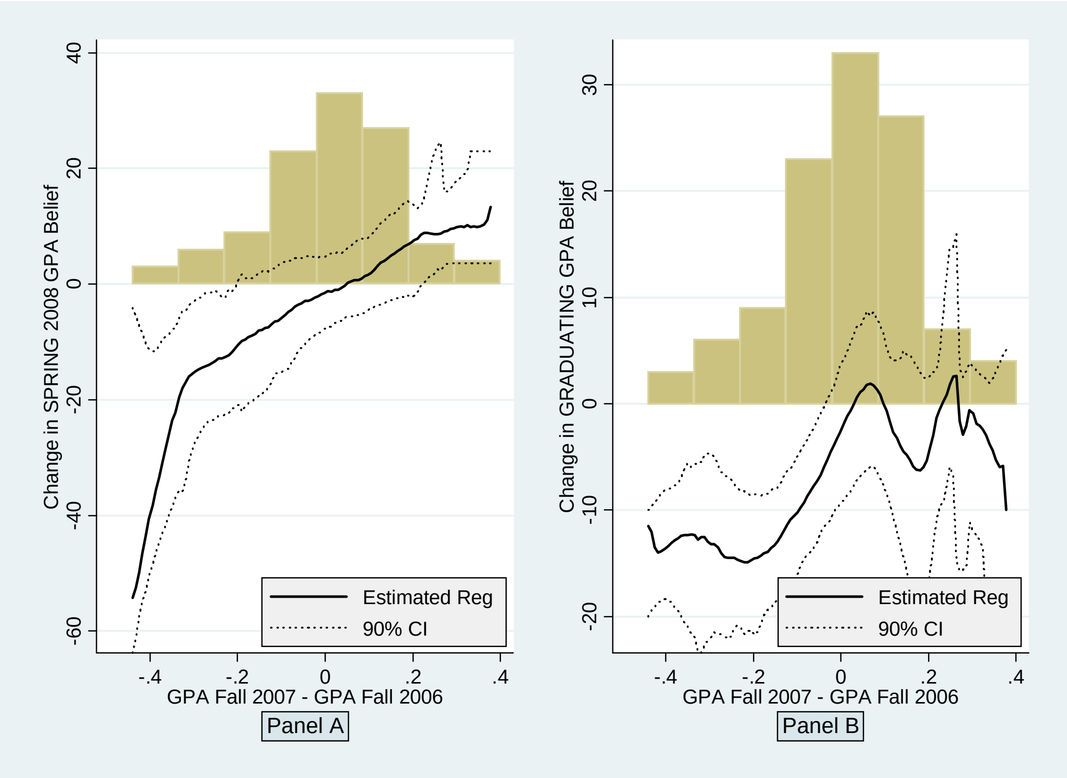

Figure 2: Local linear regressions of the change in Spring 2008 beliefs (Panel A) and Graduation GPA beliefs (Panel B) on changes in GPA between the surveys. Confidence intervals estimated from 200 bootstrap sampling distributions.

The change in beliefs about Spring 2008 GPA in response to positive and negative changes in realized GPA is almost symmetric, except for very negative GPA changes. Similar responsiveness to positive and negative changes in realized GPA may lead one to conclude that increases and decreases in realized GPA between the two surveys contained equally useful information. However, to be able to conclude this, one needs to discern the information content of the GPA realized at the beginning of Fall 2007. More specifically, one needs to know the respondents' prior probability distributions (i.e., their belief in the Fall 2006 survey) about their GPA at the start of Fall 2007.[13] In the absence of this information, one may conclude positive information for negative information when the individual's GPA in Fall 2007 decreases by less than the individual had anticipated. To highlight this point, consider the following example: Individual A's GPA is up by 0.3 point at the beginning of Fall 2007 (relative to Fall 2006 GPA), while that of individual B is down by 0.1 point. Further assume that, when taking the initial survey in Fall 2006, individual A had forecast her GPA at the beginning of Fall 2007 to be up by 0.4 points,

[13] To be more precise, the change in GPA between the two surveys actually is the difference in GPA at the beginning of Fall 2007 (which would be the GPA realized at the end of Spring 2007) and the GPA at the beginning of the quarter when the individual took the initial survey. Therefore, Fall 2007 GPA actually means the GPA realized at the end of Spring 2007. The academic year consists of the Fall, Winter, and Spring quarters (in that order).

while individual B expected his to be down by 0.2 point.[14] In the absence of information on the individuals' beliefs, the researcher would deduce that individual A experienced a positive change and that individual B experienced a negative change, when in fact the converse is true.

Panel B of Figure 2 depicts the local linear regression estimates of the change in graduating GPA beliefs on changes in realized GPA between the two surveys. Both surveys elicited the individuals' beliefs about their GPA at graduation in their major being above 3.5; the dependent variable is now the change in this belief.[15] As depicted in panel B, individuals do not revise upward their belief of graduating GPA in response to positive changes in realized GPA (the estimated confidence intervals are not very precise, though).

In order to understand the responsiveness of beliefs about future GPA, it is important to discern the information content of the realized Fall 2007 GPA. ε_{it+1} in equation (1) can be expressed as a function of new information:

$$\varepsilon_{it+1} = h[\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})].$$

Equation (1) can now be written as:

$$\Pr(X = 1|\Omega_{it+1}) - \Pr(X = 1|\Omega_{it}) = h[\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})], \tag{2}$$

which basically states that the change in an individual's expectation between time t and $t+1$ about some event $\mathbf{X}$ that is realized at some point in the future is a function of surprises between time t and $t+1$. This equation highlights the challenges in studying the updating of expectations; not only does the researcher need data on expectations of an agent over time, but also needs to identify new information between periods. Bernheim (1988) uses assumptions on prior expectations in order to identify a model of revisions of Social Security benefit expectations. However, this approach defeats the purpose of collecting subjective expectations data. Dominitz (1998) faces the same problem in his analysis of revisions of earnings expectations in the SEE and, in the absence of knowledge about what the new information is, cannot pin down the causal explanation for the revision in expectations.

To come up with a metric of new information that wasn't anticipated at time t, I use information on the individual's GPA at the end of Spring 2007 (which is not known at time t

[14]Note that these forecasts are in Ω_t, the individuals' information sets at time t. Thus, any expectations about future events reported at time t are conditional on these forecasts.

[15]Here, $\Pr(\mathbf{X} = 1|\Omega_{it}) = \Pr(\text{Graduation GPA}_i \geq 3.5|\Omega_{it})$. This threshold, unlike the case for the Spring 2008 GPA belief, is not individual specific.

but has been realized at time $t+1$; see Figure 1). I define ω_{it+1} to equal 1 if i's cumulative GPA at the end of Spring 2007 was at least as much as her Fall 2006 GPA, i.e.:

$$\omega_{it+1} = \begin{cases} 1 \text{ if Spring 2007 GPA}_i \; \geq \; \text{Fall 2006 GPA}_i \\ 0 \text{ otherwise.} \end{cases}$$

$E(\omega_{it+1}|\Omega_{it})$ is i's belief elicited at time t (in the Fall 2006 survey) that $\Pr(\omega_{it+1} = 1|\Omega_{it})$. More specifically, in the initial survey, students were asked about the percent chance (probability) that their GPA at the end of Spring 2007 would be at least as much as their Fall 2006 GPA.[16]

Therefore, the metric $\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})$ varies from -1 (this is the case of extreme negative surprise where the individual expected the Spring 2007 GPA to be above the threshold with certainty in the Fall 2006 survey but that did not happen) to 1 (in the case of extreme positive surprise). The histogram in Figure 3 depicts the distribution of the metric in the sample. The metric varies between -1 (extreme negative surprise) to 0.8 in the sample. The mean value of the metric is -0.23, which suggests that individuals tend to be overoptimistic about their future academic performance.[17] The metric is significantly positively correlated with realized GPA changes (a Spearman rank correlation of 0.57 at the 0.01% level).

Panel A of Figure 3 depicts the local linear estimates of Equation (2), i.e., the regression of change in the Spring 2008 GPA beliefs on the new information metric. Revisions of Spring 2008 GPA expectations seem to be positively related to the new information. Individuals who receive positive information revise upward their prediction of Spring 2008 GPA only if the information metric is greater than 0.50, while individuals who receive negative information revise their predictions downward only if the information content is less than -0.50. In the intermediate range, i.e., -0.50 to 0.50, students don't revise their beliefs (the confidence interval cannot reject zero change).

Panel B of Figure 3 estimates the regression function of Equation (2) where the content of new information is defined as before, but $\mathbf{X}$ is now the GPA in one's major at the time of graduation. Panel B shows that all individuals revise downward their beliefs about graduating

[16] Though this belief was elicited on a scale of zero to 100, I normalize it to zero to one to construct the metric.

[17] Though recent studies have found that men tend to be more overconfident about their ability than are women (Niederle and Vesterlund, 2007), that is not the case here: The mean value of the metric is -0.218 for males (with a standard deviation of 0.49) and -0.232 for females (with a standard deviation of 0.53). This suggests that, on average, women in my sample tend to be more overconfident. However, I fail to reject the null that the two means are equal. Similarly, I don't find significant differences in the mean value of the metric for the different ethnic groups.

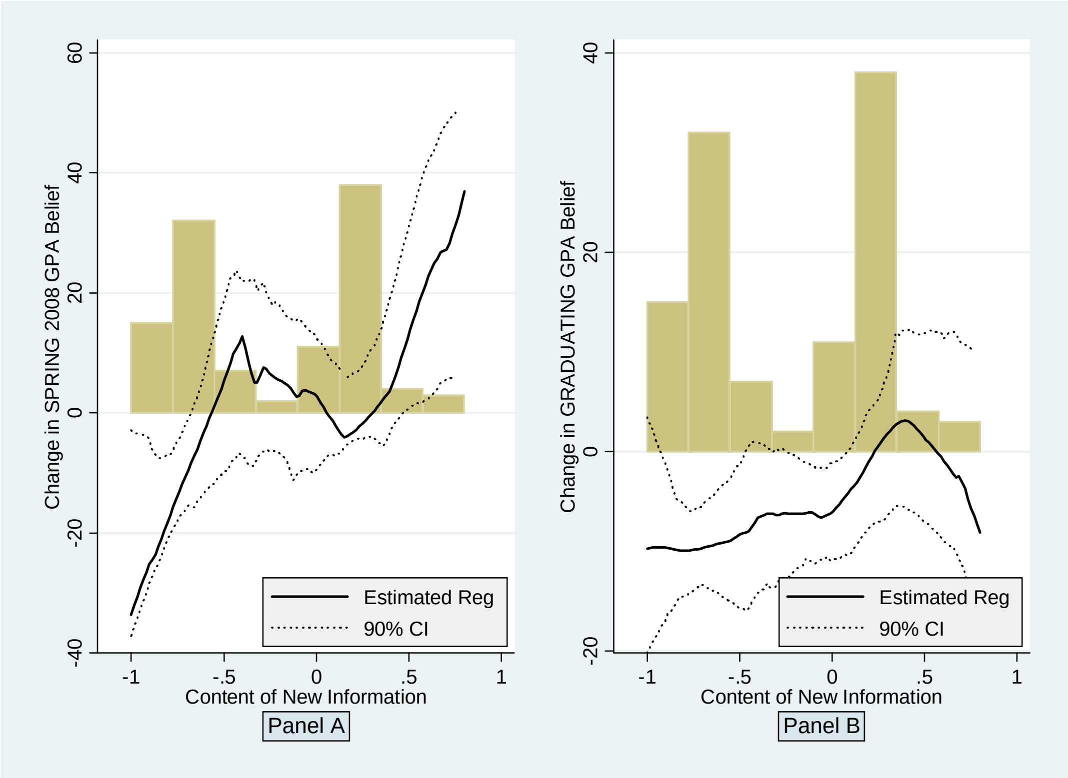

Figure 3: Local linear regressions of the change in Spring 2008 beliefs (Panel A) and Graduation GPA beliefs (Panel B) on *new* information revealed between the surveys. Confidence intervals estimated from 200 bootstrap sampling distributions.

GPA, although those doing better than expected in Spring 2007 revise them down by less. Relative to revisions in Spring 2008 GPA beliefs, individuals revise to a lesser degree their beliefs about their graduating GPA. There could be at least two reasons for this. First, the belief in question here is about the graduating GPA being above 3.5 (instead of an individual-specific threshold, as is the case for the Spring 2008 GPA). For individuals with very high or low GPAs, a threshold of 3.5 will not be binding, and therefore any new information should not cause them to revise their beliefs much. Second, since individuals have another year and a half of classes to take before the graduating GPA outcome is realized (and all these classes will be counted toward the graduating GPA), the mechanical effect of any new information contained in the Spring 2007 GPA should be lower, especially if students believe that Spring 2007 GPA gives them little information about their long-term performance.

Table 5 reports the OLS estimates of regressing the change in Spring 2008 GPA beliefs on realized GPA change and the information metric in columns (1)-(3) as well as the corresponding estimates for the change in Graduation GPA beliefs in columns (7)-(9). As in Figures 2 and 3, revisions in Spring 2008 GPA beliefs and Graduation GPA beliefs are positively related to both realized changes in GPA and the information metric. However, in an equation with both the realized GPA change and the information metric (columns 3 and 9), only the latter is significant (at the 10% level) for revisions in Spring 2008 GPA beliefs. I interpret this to mean that the information metric has an expectational element not captured in the GPA change.

Though GPA is a noisy signal of one's ability, it is also a function of one's field of study. The estimates shown in Table 5 as well as in Figures 2 and 3 would be biased if I don't account for the fact that individuals could switch majors in response to new information.[18] In the sample, 14 of the 117 respondents (~12%) switch majors between the two surveys.[19] Columns (4)-(6) and (10)-(12) in Table 5 report the OLS estimates for the sample excluding respondents who switched majors between the two surveys. Though qualitatively similar to those for the full sample, the estimates are larger in magnitude. This finding suggests that there is indeed some strategic switching of majors on the part of respondents, i.e., students who receive negative information may be switching to easier majors or those who receive positive information may decide to pursue harder majors. Closer examination of students who drop majors shows that the mean value of the metric for them is much lower (-0.265 versus -0.220 for students who don't switch majors), suggesting that negative information is associated with switching majors (Arcidiacono, 2004, also finds that poor performance is correlated with switching majors). This issue is explored in more detail in Section 5. Figure A.1 in the Appendix estimates Equation (2) by excluding those respondents. The overall pattern is similar to that in Figure 3.

Finally, it should be pointed out that I include only the Spring 2007 GPA in ω_{it+1}. It is plausible that individuals are using some other sources of information in updating their beliefs of future academic performance. However, as mentioned earlier, it is nearly impossible to identify all the new information. The analysis in this section shows that, to address the question of how individuals update their beliefs, not only is high-frequency data needed, but the researcher also needs to observe innovations in the individual's information set. Nonetheless, it is certainly reassuring that, despite focusing on an information metric that contains information only about the Spring 2007 GPA, students are found to revise their beliefs in somewhat rational ways.

[18] Another possibility is that students may take easier (harder) elective courses upon receipt of negative (positive) information about their ability. Unfortunately, I cannot address this issue with my data (one would need to observe the courses that a student intended to take in the future as well as the courses the student actually ended up taking, and some measure of the difficulty of the courses). Estimates would most likely be biased downward if this possibility is not considered.

[19] Here, switching a major means that, at the time of the follow-up survey, an individual was pursuing a major different from the one at the time of the first survey and that the individual had also taken at least one course in the new major.

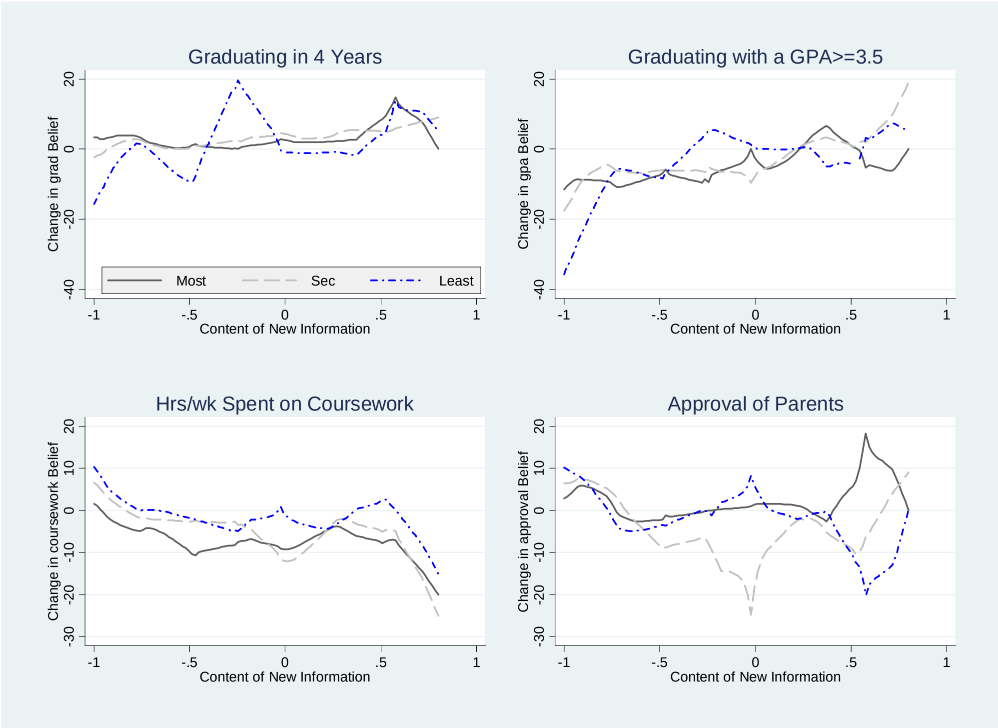

Figure 4: Local linear regressions of the change in beliefs for 1) graduating in 4 years, 2) graduating with a GPA$\geq$3.5, 3) expected hours per week spent on coursework, and 4) approval of parents (in one's *most* preferred/current major, *second* (most preferred) major, and *least* preferred major) on *new* information about ability revealed between the surveys. Standard errors on these regressions are not reported since that would make the graph too cluttered.

3.2 Revisions of various major-specific beliefs

The discussion in Section 3.1 highlights the breadth of data required to understand the revision of expectations in response to new information. Unfortunately, I don't have data for similar metrics of surprise for other determinants. This section investigates how individuals revise their beliefs for other major-specific outcomes in response to new information revealed about academic ability. It's not clear how beliefs for various outcomes in different majors should change in response to new information acquired about ability in a specific major. Beliefs about certain outcomes, such as graduating in 4 years, may change in response to this information. On the other hand, beliefs about outcomes, such as gaining approval of parents, may not change in response to this information. Figure 4 depicts the local linear polynomial estimates of the regression of change in beliefs in the three different major categories for 1) graduating in 4 years, 2) graduating with a GPA$\geq$3.5, 3) expected hours per week spent on coursework, and 4) approval of parents on the new information acquired between the two surveys.

The top-left panel in Figure 4 shows that, for graduating in 4 years, students revise their beliefs only for extreme changes in the information content, and the same relationship is observed for all three major categories. More specifically, students revise downward (upward) their beliefs about graduating in 4 years on receipt of very negative (positive) information. A similar pattern is observed in the case of revised beliefs of graduating with a GPA of ≥ 3.5 (top-right panel of Figure 4). Conversely, as depicted in the bottom-left panel of Figure 4, a negative relationship is observed between revisions of beliefs about coursework hours per week and the information metric. Students who receive positive (negative) information about their academic ability revise their beliefs downward (upward) about expected hours per week spent on coursework in all three major categories.[20] This result are consistent with Stinebrickner and Stinebrickner (2007), who find a causal effect of studying on academic performance. On the other hand, revisions of beliefs for outcomes such as approval of parents (bottom-right panel of Figure 4) don't seem to vary in any particular way with the new information. Revisions of beliefs for other outcomes are reported in Figure A.2; there is no systematic pattern in the revision of these beliefs either.

On the whole, these figures suggest that, at least for some outcomes, there is a clear and logical pattern in which beliefs are revised.

3.3 Variance in beliefs

Another testable implication of the model outlined in Section 3.1 is the change in cross-sectional variance of beliefs over time. From Equation (1), under the assumption that $\varepsilon's$ are iid and that $\varepsilon_{it+1} \perp \Pr(\mathbf{X}=1|\Omega_{it})$ (i.e., any new information is independent of an individual's beliefs about a future outcome), one can write:

$$\begin{aligned} Var[\Pr(\mathbf{X} &= 1|\Omega_{it+1})] = Var[\Pr(\mathbf{X}=1|\Omega_{it})] + Var[\varepsilon_{it+1}] \\ &= Var[\Pr(\mathbf{X}=1|\Omega_{it})] + Var[\Pr(\mathbf{X}=1|\Omega_{it+1}) - \Pr(\mathbf{X}=1|\Omega_{it})] \\ &> Var[\Pr(\mathbf{X}=1|\Omega_{it})], \end{aligned} \tag{3}$$

which implies that the variance of beliefs over time should increase. The intuition of this result is that the population variance in the beliefs should go up as the time to the realization of the outcome nears, since people become more certain about what the realization of the outcome

[20] This pattern between beliefs about coursework hours/week and new information about ability would be obtained if (perceived) ability affects the marginal utility of effort negatively, i.e., students with higher perceived ability spend fewer hours/week on coursework to attain the same GPA. In that case, students who receive a positive signal about ability should decrease the number of hours per week that they expect to spend on coursework.

would be for them. I test this implication for each of the ten outcomes in all the 5 major categories (see Table 7 for the list of outcomes and the majors). In 41 of the 50 tests, the variance in the beliefs elicited in the follow-up survey is greater than the variance in the beliefs in the initial survey.[21] Therefore, this analysis lends support to the test that variance in beliefs increases over time.

Before a formal characterization of the belief-updating process in the next section, I present further evidence that the learning process is consistent with a Bayesian learning approach. I define a dummy, U_i, that equals 1 if, in the initial survey, the individual was *more* uncertain about the occurrence of the major-specific outcome, and zero otherwise. More specifically:

$$U_i = \begin{cases} 1 \text{ if } 25 \leq \Pr(X = 1|\Omega_{it}) \leq 75 \\ 0 \text{ otherwise.} \end{cases}$$

Table 6 regresses $|\Pr(X = 1|\Omega_{it+1}) - \Pr(X = 1|\Omega_{it})|$, the absolute change in beliefs between the two surveys for each of the binary outcomes, on the dummy U_i and a constant term.[22] The coefficient on U_i is positive and statistically significant for each of the major-specific outcomes, suggesting that individuals who are more uncertain about the major-specific outcomes in the initial survey make greater absolute revisions in their beliefs. Both these pieces of evidence, i.e., variance in beliefs increases over time, and larger absolute revisions in beliefs for those who are more uncertain in the initial survey, are consistent with Bayesian updating.

4 Characterizing the Belief-updating Process

This section formalizes the nature of the belief-updating process. I assume that individuals adopt a Bayesian learning approach. If the beliefs of the individuals can be characterized by a beta distribution (which is ideally suited to analyze binary events), the posterior probability P_{ijm}^{t+1} (individual i's probabilistic belief of outcome j happening in the case of major m) is given by (see Viscusi and O'Connor, 1984; and Viscusi, 1997):

$$P_{ijm}^{t+1} = \frac{\alpha}{\alpha+\beta} P_{ijm}^{t} + \frac{\beta}{\alpha+\beta} I_{ijm}, \tag{4}$$

[21] The test is rejected for two outcomes (reconciling work and family at the jobs and expected hours per week at job) for one's current major; for one outcome (enjoying coursework) for the second most preferred major; for three outcomes (graduating in 4 years, enjoying coursework, and reconciling work and family at jobs) for the second major; for one category (expected hours per week at the job) for the dropped major category; and for two outcomes (enjoying coursework and expected salary at the age of 30) for the least preferred major.

[22] Here, I interpret responses in the range of 25-75 (on a scale of 0-100) as exhibiting more uncertainty. Results are robust to alternate definitions as well.

where P_{ijm}^{t} is i's prior belief of outcome j in major m, I_{ijm} is new information that i acquires about this outcome between period t and $t+1$, α is the precision of the prior, and β is the precision of the new information. In this framework, the new information is equivalent to observing additional Bernoulli trials about the occurrence of the various major-specific outcomes. In the context of this study, the prior belief refers to the subjective belief elicited in the initial survey, while the posterior refers to the belief elicited in the follow-up survey. To empirically estimate Equation (4), the researcher needs to determine the individual's information set at both times t and $t+1$, which is almost impossible (Cunha et al., 2004).

In order to estimate Equation (4), I use the information metric introduced in Section 3.1 (the metric that captures the extent of new information that an individual acquires about her academic ability in her current major) as a proxy for the new information. Needless to say, the information metric only partially identifies the new information that individuals receive between the two surveys. Moreover, information about academic ability in one's current major may or may not affect one's beliefs about outcomes associated with other majors or beliefs for outcomes other than academic achievement in the same major. I use the following regression framework for the empirical investigation of (4):

$$P_{ijm}^{t+1} = \gamma P_{ijm}^{t} + \eta I_{ijm} + D_{im} + \varepsilon_{ijm}, \tag{5}$$

where D_{im} is a dummy that equals 1 for major m and zero otherwise, ε_{jm} is a random error term, and:

$$\gamma = \frac{\alpha}{\alpha+\beta}; \quad \eta = \frac{\beta}{\alpha+\beta}.$$

The empirical specification includes a major dummy (D_{im}) to allow for common shocks within a major.[23] In this framework, the coefficients γ and η show the nature of the learning process. One would expect γ to be equal to 1 and η to be equal to 0 if the individual depends solely on her prior information and does not learn any new information about the outcome from the information metric. On the other hand, if the new information is really valuable, γ would be close to zero and η would be large. Equation (5) is estimated for each of the major-specific outcomes and for three different majors in the individual's choice set. The above-mentioned interpretation of the model does not apply to the continuous outcomes (coursework hours per

[23] Regressions that were run excluding major-specific shocks (D_{im}) yield similar results qualitatively, and are available upon request from the author.

week; job hours per week; expected salary); I discuss the updating of expected salary in detail in Section 6.

The results are shown in Table 7. The estimates are between the two extremes, and the prior belief continues to play a significant role in almost all the cases. However, γ is smaller than 1 in most cases, suggesting that the prior belief is not very precise. The table shows that η is small in magnitude and not statistically significant in most cases, suggesting that the information metric is not very useful in predicting the belief-updating process. Another object of interest is the importance of new information relative to the prior, which is denoted as R and given as:

$$R = \frac{\beta}{\alpha} = \frac{1}{\gamma} - 1.$$

Higher values of R would imply greater relative informativeness of the new information. The third row in each panel of Table 7 shows the estimates of R. These estimates indicate that this new information is not very valuable; in most cases, R is less than 1. For outcomes such as approval of parents, new information does not seem to be valuable ($|R| < 0.35$). This finding is plausible because one would expect students to be aware of their parents' perceptions of different majors when they start college, and therefore they should be less likely to receive any valuable information about parents' approval over time. Similarly, priors for outcomes such as graduating in 4 years and graduating with a GPA of at least 3.5 are fairly precise. On the other hand, for outcomes related to the workplace such as finding a job or enjoying working at the jobs, it seems that the prior is less precise: The metric R is larger for these outcomes. These findings are consistent with students adopting a Bayesian learning approach. For outcomes associated with college, one would expect students to have fairly precise information at the time of the initial survey, and any new information should be relatively less informative. Conversely, for outcomes associated with the workplace, one would expect students to receive *useful* information between the two surveys, and hence the relative importance of new information would be higher in that case.

5 Experimenting with Majors

Students may be uncertain about their ability and other outcomes when choosing a major. Over time, when new information arrives, they may choose to drop out of college or switch to a different major that they deem to be a better fit (Manski, 1989; Altonji, 1993; Arcidiacono, 2004; Malamud, 2007; Stinebrickner and Stinebrickner, 2008). In the context of the current setting, Northwestern University, dropouts are not very common. Completion rates for the 2006 and 2007 undergraduate class were 93%. Instead, students are more likely to switch majors during the course of their undergraduate studies. Of the 117 survey respondents, 14 (~12%) switched their majors between the two surveys.

The model I have in mind is as follows. At time t, individual i derives utility $U_{ikt}(\mathbf{a}, \mathbf{c}, X_{it})$ from choosing major k. Utility is a function of a vector of outcomes $\mathbf{a}$ that are realized in college, a vector of outcomes $\mathbf{c}$ that are realized after graduating from college, and individual characteristics X_{it} (outcomes in vectors $\mathbf{a}$ and $\mathbf{c}$ are described in Section 2). Since the outcomes in vectors $\mathbf{a}$ and $\mathbf{c}$ are uncertain at time t, i possesses subjective beliefs $P_{ikt}(\mathbf{a}, \mathbf{c})$ about the outcomes associated with choice of major k for all $k \in C_i$. Individual i chooses major m at time t if

$$m = \arg\max_{k \in C_i} \int U_{ikt}(\mathbf{a}, \mathbf{c}, X_{it}) dP_{ikt}(\mathbf{a}, \mathbf{c}). \tag{6}$$

However, over time, new information may arrive that may lead the individual to update her beliefs about any of the major-specific outcomes. A change in an individual's beliefs about her ability (graduating GPA or probability of completing the major in 4 years), match quality in college (outcomes like enjoying coursework), or match quality in workplace (enjoying working at the jobs or expected earnings at the jobs) may lead the individual to switch to a major that yields higher expected utility.[24] To understand the pattern of switches in major, one would need not only data on the subjective beliefs about major-specific outcomes at several points in time, but also data on how the respondent *believes* the subjective beliefs will evolve over time. For example, as outlined in Section 3.1, one cannot simply infer positive news from observing a GPA increase from one quarter to the next. Instead, one needs to observe how much the student anticipated that her GPA would change over that time horizon. Having very little data on the prior distributions of the respondents' beliefs, I can only conduct a descriptive analysis of why individuals experiment with different majors. Moreover, I focus my analysis primarily on the role of learning about ability in the decision to switch majors.

Individuals who switch majors experience a small average gain of about 0.17 point in their GPA.[25] Fewer than 50% of these individuals experience a positive change in their GPA, sug-

[24] Here, as in Becker and Stigler (1977), I assume that preferences are stationary.

[25] This number comes from directly asking the respondents to report their major-specific GPA for the new major and then comparing it to their GPA in the previous major.

gesting that academic performance is not the only dimension that influences one's choice of major. Respondents were asked to assign weights to different reasons for dropping the major so that they summed to a 100. Table 8 reports the average weight assigned to each reason. Losing interest in the original major, getting interested in something else, and finding the initial major too challenging stand out as the main reasons for dropping the initial major.

Another difficulty in analyzing experimentation with majors is that it is hard to determine when exactly an individual switched majors, since students don't have to formally declare a major to take courses in it. An individual may take a few courses in a new major and then decide to pursue it, or vice versa. Recall that the first survey was conducted during Fall 2006 and the second survey during Fall 2007. Therefore, if an individual switched majors (and also took courses in the new major) between the two surveys, it was most likely in response to information acquired in the Fall 2006 quarter or the Winter 2007 quarter. I define a dummy variable, S_i, that equals 1 if individual i had switched her major between the two surveys, and zero otherwise. I estimate a probit model of the following form:

$$\Pr(S_i = 1) = \Phi(\text{constant} + \alpha * \Delta\text{GPA}_i + \beta * I_i), \tag{7}$$

where ΔGPA_i is the change in an individual's GPA between Winter 2007 and the beginning of Fall 2006, and I_i is the information metric that was introduced in Section 3.1. This metric is only a crude proxy of information relevant for switching majors since it is constructed using information on the Spring 2007 cumulative GPA, by which time an individual had certainly switched majors. However, by construction, it indirectly incorporates any information gained about ability in the quarters starting from the initial survey up until Spring 2007. The purpose of estimating a model of the form in Equation (7) is only to present patterns between major switches and realized GPA changes as well as the information metric.

Table 9 reports the marginal effects (computed at the mean of the independent variable) for the various specifications. The marginal effects seem to be of the *correct* sign, but none of them is statistically different from zero.[26] The marginal effects reveal that a unit increase in the information metric is associated with a decrease of about 1.5% in the probability of switching majors (which is 11.96% in the sample), while a unit increase in ΔGPA is associated with a

[26]This could be either because of the small sample size (I have only 14 major switches in the sample), or because realized changes in GPA and the information metric are actually not predictive of a major switch. Moreover, what is actually needed for the analysis is some measure of the information content acquired in each quarter about each outcome and the exact timing of the major switch.

decrease of about 1% in the probability of switching majors.

Table 10 regresses the change in beliefs for each outcome onto dummies for the different major categories (second preferred major, second major, dropped major, and least preferred major).[27] The coefficients show the direction and magnitude of the mean change in beliefs about the various outcomes for each of the majors. Mean changes in the current major are indicated in the estimate of the constant. In the case of the dropped major, only the change in beliefs for graduating in 4 years is statistically significant. Relative to one's current major, students revise down their beliefs for graduating in 4 years for the dropped major by an additional 8.5 points. Though none of the other changes is significant (presumably because of small sample sizes), changes in beliefs about enjoying coursework and expected salary at the age of 30 seem to be quantitatively different from the corresponding changes in one's current major. If one were to assume that these changes accurately reflect the changes in beliefs at the instant when an individual switched her major, it seems that negative changes in beliefs about graduating in 4 years, enjoying coursework, and expected salary at age 30 are associated with the dropping of a major.[28]

6 Formation of Salary Expectations

Large earnings premiums exist across majors (Daymont and Andrisani, 1984; Garman and Loury, 1995). Although students reported expected income at the age of 30 for various majors, no objective measures exist to which their responses can be compared.[29] Instead, this section analyzes students' responses to questions that asked about the average annual starting salary of Northwestern bachelor's degree graduates of 2007 for three different majors in their choice set. Responses to this question can be compared directly to actual salary realizations of Northwestern graduates. The question asked was: *"What do you think was the average annual starting salary of Northwestern G graduates (of 2007) with Bachelor's Degrees in X?"* where $G = \{\text{Male, Female}\}$. Though there is substantial heterogeneity in the beliefs, I present only

[27] The table reports the change in beliefs *after* the individual has already switched her major. If we really want to understand what led an individual to switch her major, we would need to observe her beliefs right before she made the decision, which I don't have. Nonetheless, it is useful to go through this exercise to see how beliefs changed between the surveys for the dropped major category versus other categories.

[28] It could be that once an individual has decided to drop a major, she devalues the outcomes associated with that major in order to rationalize her choice (cognitive dissonance; see Festinger, 1957). However, estimates in Table 10 indicate that this is not the case. For example, beliefs about enjoying coursework and enjoying work at the jobs are revised downward in all major categories, not only for the dropped major.

[29] This is because Northwestern University does not follow its alumni. Moreover, even if such data existed, one would have to make assumptions about how students believe earnings across majors evolve over time.

the mean responses in Table 11. Analysis of the first six columns of Table 11 shows that respondents are aware of different returns to majors. Moreover, the relative subjective beliefs seem to be consistent with actual trends. There are, however, a few notable patterns. Both males and females underestimate the average salaries (for both genders) for all categories except Natural Science and Ethics and Values. However, compared to their male counterparts, female respondents report higher average starting salaries (for both themselves as well as for males) for Engineering and several majors in the Weinberg College of Arts and Sciences (WCAS). Column (7) shows that the realized wage gap is in favor of males for all WCAS categories except Area Studies and Literature and Fine Arts. The survey respondents (both males and females), on average, believe that the wage gap is in favor of males for all major categories (columns 8 and 9). However, both males and females tend to underestimate the extent of the gender gap in wages for most majors.

Table 11 shows only the average beliefs by gender. Using the demographic information collected from the respondents, I further explore the determinants of the errors in the respondents' beliefs about Northwestern 2007 graduates' salaries. As in Betts (1996), I use the following metric to model the respondents' errors:

$$\ln \left| \frac{\overline{s_{im}^{G}} - s_{obs\_m}^{G}}{s_{obs\_m}^{G}} * 100 \right|, \tag{8}$$

where $\overline{s_{im}^{G}}$ is respondent i's reported average starting salary in major m for gender G ($G = \{\text{Male, Female}\}$), and $s_{obs\_m}^{G}$ is the true average salary for Northwestern 2007 graduates of gender G in major m. For each respondent, there are three values of the metric, one for her current major, one for her second (preferred) major, and one for her least preferred or dropped major. Columns (1) and (2) of Table 12 show the results of regressing this metric on various demographic characteristics for responses about male and female starting salaries, respectively. A random effect is included for each respondent in order to account for random differences in estimates between the respondents. Students with higher SAT math scores make larger errors, while students with higher SAT verbal scores make smaller errors. One would expect individuals majoring in a given field to have better information about their chosen field. The regression includes a variable "Studying Major" that equals 1 if the student is majoring in the major about which the starting salary is reported. The coefficient on this variable is insignificant.

Individuals who were studying the given field for which they reported the starting salary and had also declared their major at the time of the initial survey make smaller errors (though the coefficient is statistically insignificant). This result would be consistent with a story where information acquisition is costly, and individuals seek information about a major only when they are fairly sure about pursuing it. Individuals with a college-educated father make significantly smaller errors, which is consistent with students with college-educated parents having access to more precise information. I find no evidence of individuals with parents who have studied a given major being better informed about starting salaries in that major. One of the more notable findings is that females make significantly larger errors: On average, they make errors that are about 35 log points larger than those of males.[30]

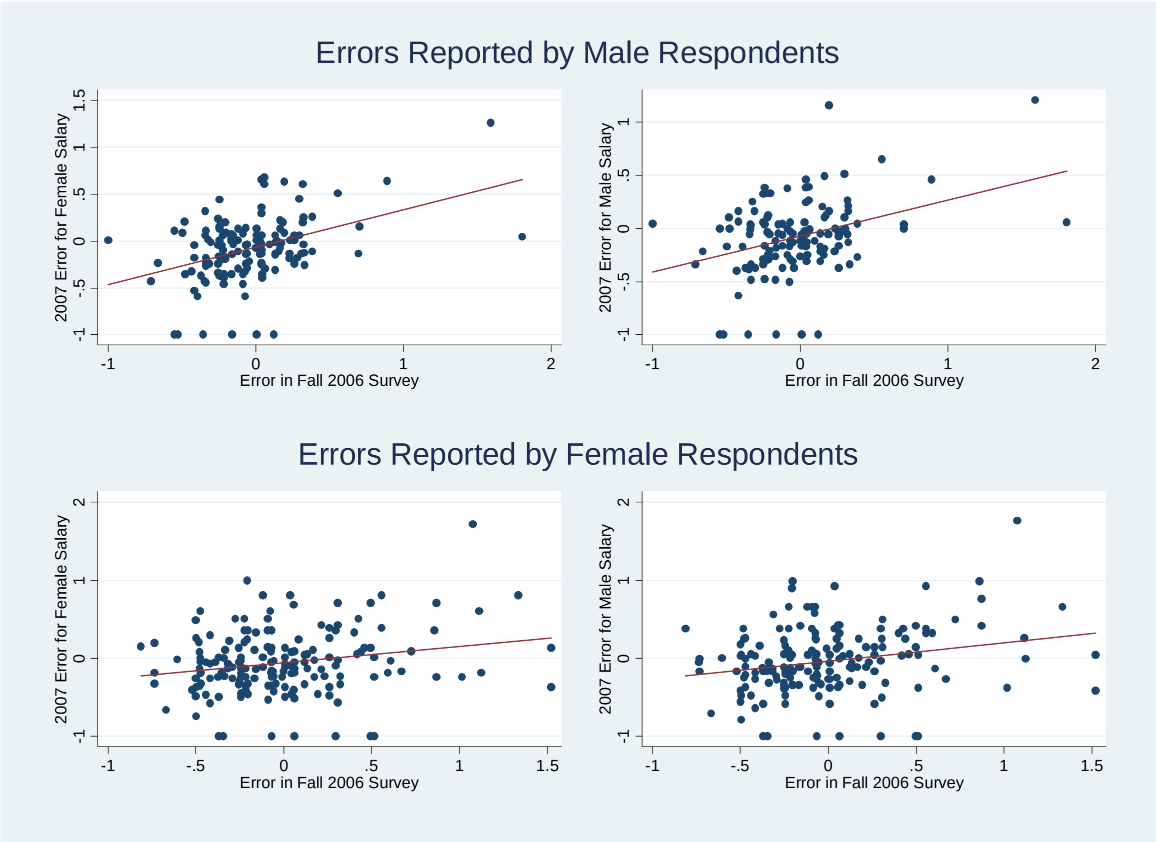

Figure 5: Scatterplots of the errors in salary expectations reported in the two surveys.

The initial survey (Fall 2006) had a similar question, except that students were asked the average annual starting salary for Northwestern 2006 graduates in the various majors *unconditional* on gender. Column (3) of Table 12 shows that demographic characteristics are correlated with the error metric in ways similar to what they were for the responses in the Fall 2007 sur-

[30]This finding contrasts with Betts (1996), who does not find any statistically significant difference in the error patterns between males and females.

vey. With data on students' beliefs about starting salaries at two different points in time, I can study how students' beliefs about starting salaries evolve over time. Figure 5 presents the scatterplots of the errors (i.e., the metric in Equation (8) without taking the absolute value or logs) made in the two surveys for male and female respondents separately. There is a lot of variation in errors for the same person between the surveys. However, as depicted in each of the plots, it seems that students tend to err in the same direction: Both males and females who underestimate (overestimate) starting salaries in the first survey are more likely to underestimate (overestimate) them in the follow-up survey. Indeed, I find a significant positive correlation between the errors in the two surveys for both male and female respondents. This indicates persistence in the students' beliefs about starting salaries.

7 Conclusion

Recent empirical work underscores the importance of uncertainty in the context of schooling choices. For example, Cunha et al. (2004) find that, if students had perfect information, around 30% of them would change their schooling choices. Understanding how individuals make schooling decisions requires one to study how students process information to form expectations and how these subjective expectations and preferences are used to make schooling choices. This paper focuses on the former (see Zafar, 2008, for the latter).

This paper enhances our limited understanding of how students form expectations by focusing on how college students revise expectations for outcomes associated with choice of college major. In the paper, revisions of expectations of future GPA are found to be positively related to changes in GPA between the two surveys. However, unlike in existing studies, I collect data that directly identify some new information that students acquire between the two surveys, which allows me to pin down some of the causal mechanisms that lead individuals to revise their beliefs. By combining elicited expectations of GPA at various points in time with their realizations, I form an information metric about academic ability and find that individuals update their beliefs for various major-specific outcomes in response to this information metric in appropriate ways. For example, individuals who receive positive information about their academic performance revise down their beliefs about number of hours per week that they expect to spend on coursework, and revise their beliefs upward if the information is very negative. Section 3 shows that modifications in expectations about various major-specific outcomes are consistent with a Bayesian learning framework. For instance, priors for outcomes like approval of parents and graduating in 4 years are fairly precise, and individuals don't revise them by as much as they revise their priors for outcomes realized in the workplace. I also find that individuals who are more uncertain about the major-specific outcomes in the initial survey make greater absolute revisions in their beliefs. Learning seems to play a role in the switching of majors. Dropped majors are associated with negative revisions in beliefs about graduating in 4 years, enjoying coursework, and expected salary.

At least two directions can be taken from here. The first deals with the methodological aspect of this paper. As mentioned in Sections 3 and 4, identifying the information set of an individual is an extremely daunting task. This paper focuses only on innovations in information about academic ability since that is the only part of the information set I can identify. To enhance our understanding of expectations formation, it is crucial to collect repeated data on subjective expectations over a short time horizon and to identify changes in one's information set. However, as argued in Manski (2004), rich longitudinal data on subjective expectations may not suffice to help in understanding expectations formation, and probing students to learn how they perceive their environments may be informative.

From an applied aspect, it seems that students are forming their beliefs for various major-specific outcomes even before they come to college. For most outcomes, the prior belief continues to be important. In attempting to understand the choice of college majors, it might be useful to focus on students at earlier stages of their schooling (for example, in high school) and analyze their subjective beliefs.

References

Altonji, Joseph. 1993. "The Demand for and Return to Education when Education Outcomes are Uncertain." *Journal of Labor Economics*, 11: 48-83.

Arcidiacono, Peter. 2004. "Ability Sorting and the Returns to College Major." *Journal of Econometrics*, 121(1-2): 343-375.

Arcidiacono, Peter, Esteban Aucejo, Hanming Fang, and Kenneth Spenner. 2009. "Does Affirmative Action Lead to Mismatch? A New Test and Evidence." Duke University Working Paper.

Bamberger, Gustavo. 1986. "Occupation Choice: The Role of Undergraduate Education." Ph.D. Dissertation, University of Chicago.

Becker, Gary, and George Stigler. 1977. "De Gustibus Non Est Disputandum." *American Economic Review*, 67: 76-90.

Betts, Julian. 1996. "What do Students Know about Wages? Evidence from a Survey of Undergraduates." *Journal of Human Resources*, 31(1): 27-56.

Bernheim, B. Douglas. 1988. "How Do Elderly Form Expectations? An Analysis of Responses to New Information." NBER Working Paper No. 2719.

Blau, Francine, and Marianne Ferber. 1991. "Career Plans and Expectations of Young Women and Men." *Journal of Human Resources*, 26: 581-607.

Bruine de Bruin, Wandi, Baruch Fischhoff, Bonnie Halpern-Felsher, and Susan Millstein. 2000. "Expressing Epistemic Uncertainty: It's a Fifty-Fifty Chance." *Organizational Behavior and Human Decision Processes*, 81: 115-131.

Cunha, Flavio, James Heckman, and Salvador Navarro. 2004. "Separating Heterogeneity from Uncertainty in Life Cycle Earnings." *Oxford Economic Papers*, 57(2): 191-261

Daymont, Thomas, and Paul Andrisani. 1984. "Job Preferences, College Major, and the Gender Gap in Earnings." *Journal of Human Resources*, 19: 408-428.

Delavande, Adeline. 2008a. "Pill, Patch or Shot? Subjective Expectations and Birth Control Choice." *International Economic Review*, 49(3): 999-1042.

Delavande, Adeline. 2008b. "Measuring Revisions to Subjective Expectations." *Journal of Risk and Uncertainty*, 36(1): 43-82.

Dominitz, Jeff and Charles Manski. 1996. "Eliciting Student Expectations of the Returns to Schooling." *Journal of Human Resources*, 31: 1-26

Dominitz Jeff. 1998. "Earning Expectations, Revisions, and Realizations." *Review of Economics and Statistics*, 80: 374-388

El-Gamal, Mahmoud, and David Grether. 1995. "Are People Bayesian? Uncovering Behavioral Strategies." *Journal of the American Statistical Association*, 90 (432): 1137-1145.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.

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

Garman, David, and Linda Loury. 1995. "College Selectivity and Earnings." *Journal of Labor Economics*, 13: 289-308.

Houser, Daniel, Michael Keane, and Kevin McCabe. 2004. "Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm." *Econometrica*, 72(3): 781-822.

Hurd, Michael, and Kathleen McGarry. 2002. "The Predictive Validity of Subjective Probabilities of Survival." *Economic Journal*, 112(482): 966-985.

Lochner, Lance. 2007. "Individual Perceptions of the Criminal Justice System." *American Economic Review*, 97(1): 444-460.

Madeira, Carlos. 2007. "Forecasting Academic Achievement." Northwestern University Working Paper.

Malamud, Ofer. 2007. “Breadth vs. Depth: The Effect of Academic Specialization on Labor Market Outcomes.” Harris School Working Paper Series: 08.08.

Manski, Charles. 1989. "Schooling as Experimentation: A Reappraisal of the Post-Secondary Drop-out Phenomenon." *Economics of Education Review*, 8(4): 305-312.

Manski, Charles. 1993. “Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?” In *Studies of Supply and Demand in Higher Education*, ed. Charles Clotfelter and Michael Rothschild, 43-57. Chicago: University of Chicago Press.

Manski, Charles. 2004. "Measuring Expectations." *Econometrica*, 72(5): 1329-1376.

Niederle, Muriel, and Lise Vesterlund. 2007. “Do Women Shy away from Competition? Do Men Compete too Much?” *Quarterly Journal of Economics*, 122(3): 1067-1101.

Smith, Herbert, and Brian Powell. 1990. "Great Expectations: Variations in Income Expectations Among College Seniors." *Sociology of Education*, 63 (July): 194-207.

Smith, V. Kerry, Donald Taylor, and Frank Sloan. 2001. "Longevity Expectations and Death: Can People Predict Their Own Demise." *American Economic Review*, 91(4): 1126-1134.

Stange, Kevin. 2008. "An Empirical Examination of the Option Value of Schooling." University of Michigan Working Paper.

Stinebrickner, Todd, and Ralph Stinebrickner. 2007. "The Causal Effect of Studying on Academic Performance." University of Western Ontario Working Paper.

Stinebrickner, Todd, and Ralph Stinebrickner. 2008. "Learning about Academic Ability and the College Drop-out Decision." University of Western Ontario Working Paper.

Viscusi, W. Kip, and Charles O’Connor. 1984. "Adaptive Responses to Chemical Labeling: Are Workers Bayesian Decision Makers?" *American Economic Review*, 74(5): 942-956.

Viscusi, W. Kip. 1997. "Alarmist Decisions with Divergent Risk Information." *The Economic Journal*, 107(445): 1657-1670.

Zafar, Basit. 2009. "College Major Choice and the Gender Gap", Federal Reserve Bank of New York Staff Report No. 364.

8 Appendix

8.1 Survey Excerpt

The following introduction was read and handed to the respondents at the start of the survey:

> *"In some of the survey questions, you will be asked about the PERCENT CHANCE of something happening. The percent chance must be a number between zero and 100. Numbers like 2 or 5% indicate “almost no chance,” 19% or so may mean “not much chance,” a 47 or 55% chance may be a “pretty even chance,” 82% or so indicates a “very good chance,” and a 95 or 98% mean “almost certain.” The percent chance can also be thought of as the NUMBER OF CHANCES OUT OF 100.*

The following set of questions was asked for each of the relevant categories. The questions below were asked for Natural Sciences.

Q1 If you were majoring in Natural Sciences, what would be your most likely major?

Q2 If you were majoring in Natural Sciences, what do you think is the percent chance that you will successfully complete this major in 4 years (from the time that you started college)? (Successfully complete means to complete a bachelors)

NOTE: In answering these questions fully place yourself in the (possibly) hypothetical situation. For example, for this question, your answer should be the percent chance that you think you will successfully complete your major in Natural Sciences in 4 years IF you were (FORCED) to major in it.

Q3 If you were majoring in Natural Sciences, what do you think is the percent chance that you will graduate with a GPA of at least 3.5 (on a scale of 4)?

Q4 If you were majoring in Natural Sciences, what do you think is the percent chance that you will enjoy the coursework?

Q5 If you were majoring in Natural Sciences, how many hours per week on average do you think you will need to spend on the coursework?

Q6 If you were majoring in Natural Sciences, what do you think is the percent chance that your parents and other family members would approve of it?

Q7 If you were majoring in Natural Sciences, what do you think is the percent chance that you could find a job (that you would accept) immediately upon graduation?

Q8 If you obtained a bachelors in Natural Sciences, what do you think is the percent chance that you will go to graduate school in Natural Sciences some time in the future?

Q9 What do you think was the average annual starting salary of Northwestern MALE graduates (of 2007) with Bachelor's Degrees in Natural Sciences?

Q10 What do you think was the average annual starting salary of Northwestern FEMALE graduates (of 2007) with Bachelor's Degrees in Natural Sciences?

Now look ahead to when you will be 30 YEARS OLD. Think about the kinds of jobs that will be available for you and that you will accept if you successfully graduate in Natural Sciences.

NOTE that there are some jobs that you can get irrespective of what your Field of Study is. For example, one could be a janitor irrespective of their Field of Study. However, one could not get into Medical School (and hence become a doctor) if they were to major in Journalism.

Your answers SHOULD take into account whether you think you would get some kind of advanced degree after your bachelors if you majored in Natural Sciences.

Q10 What kind of jobs are you thinking of?

Q11 Look ahead to when you will be 30 YEARS OLD. If you majored in Natural Sciences, what do you think is the percent chance that you will enjoy working at the kinds of jobs that will be available to you?

Q12 Look ahead to when you will be 30 YEARS OLD. If you majored in Natural Sciences, what do you think is the percent chance that you will be able to reconcile work and your social life/ family at the kinds of jobs that will be available to you?

Q13 Look ahead to when you will be 30 YEARS OLD. If you majored in Natural Sciences, how many hours per week on average do you think you will need to spend working at the kinds of jobs that will be available to you?

When answering the next two questions, please ignore the effects of price inflation on earnings. That is, assume that one dollar today is worth the same as one dollar when you are 30 years old and when you are 40 years old.

Q14 Look ahead to when you will be 30 years old. Think about the kinds of jobs that will be available to you and that you will accept if you graduate in Natural Sciences. What is the average amount of money that you think you will earn per year by the time you are 30 YEARS OLD?

Q15 Now look ahead to when you will be 40 years old. Think about the kinds of jobs that will be available to you and that you will accept if you graduate in Natural Sciences. What is the average amount of money that you think you will earn per year by the time you are 40 YEARS OLD?

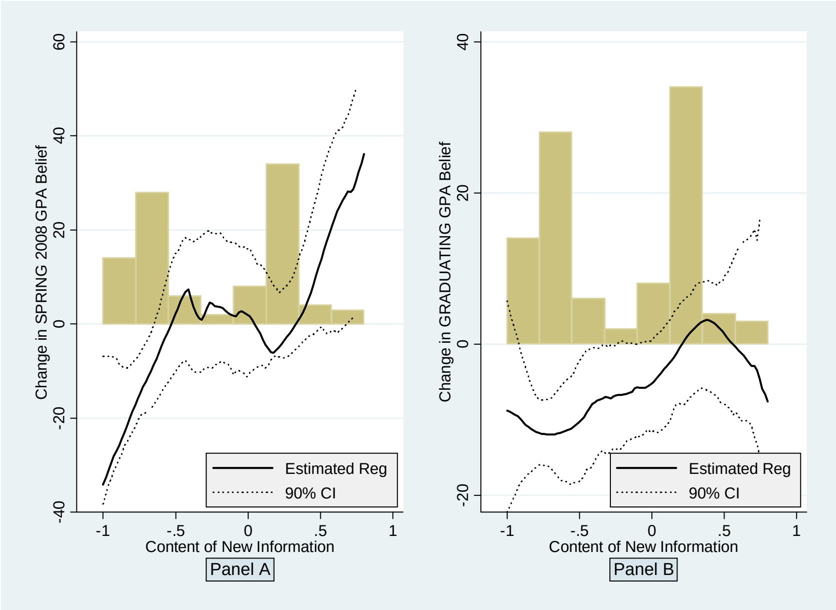

Figure A.1: Local linear regressions of the change in Spring 2008 beliefs (Panel A) and Graduation GPA beliefs (Panel B) on *new* information revealed between the surveys. Confidence intervals estimated from 200 bootstrap sampling distributions. Sample only includes respondents who had the same major in the two surveys.

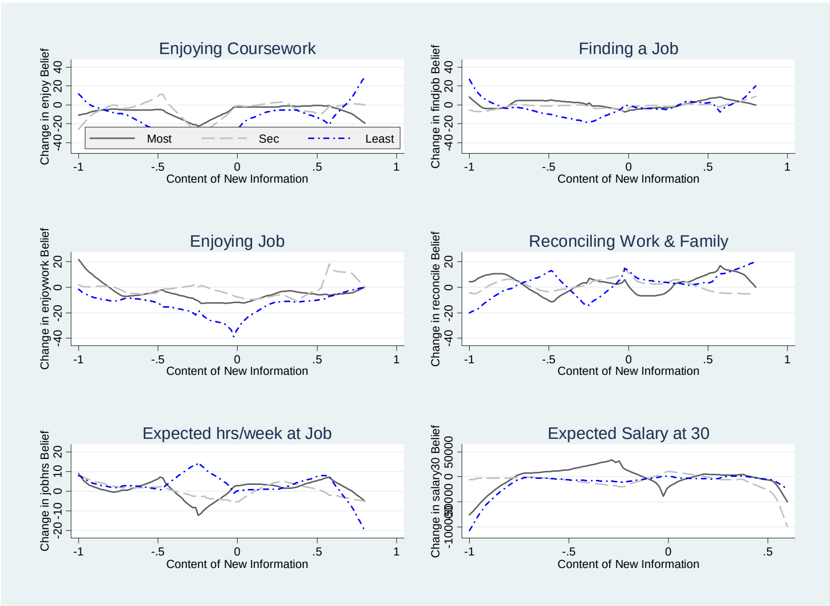

Figure A.2: Local linear regressions of the change in beliefs for various outcomes (in one's *most* preferred major, *second* most preferred major, and *least* preferred major) on *new* information about ability revealed between the surveys.

Table 1: List of Majors

The following is the classification of majors into categories:

a Natural Sciences
- Biological Sciences
- Chemistry
- Environmental Sciences
- Geography*
- Geological Sciences
- Integrated Science
- Materials Science
- Physics

b Mathematical and Computer Sciences
- Cognitive Science
- Computing and Information Systems
- Mathematics
- Statistics

c Social Sciences I
- Anthropology
- Gender Studies*
- History
- Linguistics
- Political Science
- Psychology
- Sociology

d Social Sciences II
- Economics
- Mathematical Methods in Social Sciences*

e Ethics and Values
- Legal Studies*
- Philosophy
- Religion
- Science in Human Culture*

f Area Studies
- African American Studies
- American Studies
- Asian & Middle East Languages & Civilization
- European Studies
- International Studies*
- Slavic Languages and Literatures

g Literature and Fine Arts
- Art History
- Art Theory and Practice
- Classics
- Comparative Literary Studies
- Drama
- English
- French
- German
- Italian
- Spanish

h Music Studies[1]
- Jazz Studies
- Music Cognition
- Music Composition
- Music Education
- Music Technology
- Music Theory
- Musicology
- Piano Performance
- String Performance
- Voice and Opera Performance
- Wind and Percussion Performance

i Education and Social Policy[2]
- Human Development and Psychological Services
- Learning and Organizational Change
- Secondary Teaching
- Social Policy

j Communication Studies[3]
- Communication Studies
- Dance
- Human Communication Science
- Interdepartmental Studies
- Performance Studies
- Radio/Television/ Film
- Theater

k Engineering[4]
- Applied Mathematics
- Biomedical Engineering
- Chemical Engineering
- Civil Engineering
- Computer Engineering
- Computer Science
- Electrical Engineering
- Environmental Engineering
- Industrial Engineering
- Manufacturing and Design Engineering
- Materials Science and Engineering
- Mechanical Engineering

L Journalism[5]
- Journalism

∗ *Adjunct majors (these do not stand alone)*

1 Majors in the School of Music
2 Majors in the School of Education and Social Policy
3 Majors in the School of Communication
4 Majors in the McCormick School of Engineering
5 Majors in the Medill School of Journalism

Table 2: Sample Characteristics

Characteristics	Follow-up Survey[a] **Freq.(Percent)**	Initial Survey[b] **Freq.(Percent)**	Population[c] **Freq.(Percent)**
	(1)	(2)	(3)
Gender			
Male	51 (43.5)	69 (43)	465 (46)
Female	66 (56.5)	92 (57)	546 (54)
Total	117	161	1011
Ethnicity			
Caucasian	66 (56)	79 (49)	546 (54)
African American	10 (9)	11 (7)	71 (7)
Asian	35 (30)	56 (35)	232 (23)
Hispanic	1 (1)	5 (3)	61 (6)
Other	5 (4)	10 (6)	101 (10)
Declared Major?[d]			
Yes	61 (52)	90 (56)	477[h] (47)
No	56 (48)	71 (44)	534 (53)
Second Major?[e]			
Yes	55 (47)	78 (48.5)	–
No	62 (53)	83 (51.5)	–
International Student?[f]			
Yes	5 (4)	8 (5)	40 (4)
No	112 (96)	153 (95)	971 (96)
Second-Gen Immigrant?[g]			
Yes	43 (37)	66 (41)	–
No	74 (63)	95 (59)	–
Average GPA[*]			
Male	3.51	3.48	3.26
Female	3.43	3.40	3.31

a Individuals who participated in the follow-up (second) survey
b Individuals who participated in the initial survey
c Population statistics for the sophomore class. (Source: Northwestern Office of the Registrar)
d Whether the respondent has declared a major at the time of the INITIAL survey
e Whether the respondent was pursuing a second major at the time of the INITIAL survey
f Whether the respondent is an international student
g Whether at least one of the respondent's parents is foreign born and the respondent was born in the U.S.
h Statistic obtained from Registrar's Office at the end of the Fall 2006 quarter (during/middle of first survey)
$*$ Difference in GPAs within gender between the two surveys is insignificant (2-tailed t-test)

Table 3: Distribution of WCAS Majors in the Two Surveys

WCAS Majors[a]	**Follow-up Survey**[b] **Freq**	**(%)**	**Initial Survey** **Freq**	**(%)**
Natural Sciences	22	(19)	31	(19)
Math & Computer Sci.	2	(1.5)	4	(2.5)
Social Sciences I	33	(28)	41	(25.5)
Social Sciences II	35	(30)	48	(30)
Ethics and Values	1	(1)	4	(2.5)
Area Studies	8	(7)	13	(8)
Literature & Fine Arts	16	(13.5)	20	(12.5)
Total	117	(100)	161	(100)

a Majors that appear in each category are listed in Table 1.
b In cases where the survey respondent has more than one major in WCAS, only the first one is included.

Table 4: Beliefs of Graduating with a GPA of at Least 3.5

Percent chance of graduating with a GPA≥3.5 in:								
	Current Major				**Least Preferred Major**			
Reported in:	Follow-up Survey		Initial Survey		Follow-up Survey		Initial Survey	
Subj. Belief:	Freq.	Cum. %	Freq.	Cum. %	Freq.	Cum. %	Freq.	Cum. %
0	1	0.9	1	0.9	6	5.8	-	0
1	-	0.9	-	0.9	-	5.8	3	2.9
2	-	0.9	-	0.9	2	7.8	-	2.9
3	-	0.9	-	0.9	-	7.8	1	3.9
5	4	4.4	-	0.9	1	8.7	2	5.9
10	1	5.3	1	1.8	5	13.6	1	6.9
12	-	5.3	-	1.8	-	13.6	1	7.8
15	-	5.3	-	1.8	1	14.6	2	9.8
20	-	5.3	2	3.7	12	26.2	4	13.7
21	-	5.3	-	3.7	1	27.2	-	13.7
25	2	7.1	1	4.6	5	32.0	1	14.7
30	-	7.1	1	5.5	4	35.9	5	19.6
33	-	7.1	-	5.5	-	35.9	1	20.6
35	-	7.1	-	5.5	1	36.9	3	23.5
40	1	8.0	2	7.3	5	41.8	6	29.4
45	2	9.7	1	8.3	2	43.7	3	32.4
50	16	23.9	4	11.9	7	50.5	10	42.2
55	1	24.8	1	12.8	1	51.5	1	43.1
60	7	31.0	9	21.1	7	58.2	8	51.0
65	4	34.5	3	23.9	2	60.2	3	53.9
68	-	34.5	-	23.9	1	61.2	-	53.9
70	10	43.4	8	31.3	6	67.0	10	63.7
73	-	43.4	1	32.1	-	67.0	-	63.7
75	15	56.6	7	38.5	3	69.9	3	66.7
76	-	56.6	1	39.5	-	69.9	-	66.7
79	-	56.6	1	40.4	-	69.9	-	66.7
80	13	68.1	13	52.3	7	76.7	5	71.6
82	1	69.0	2	54.1	-	76.7	1	72.6
85	7	75.2	9	62.4	-	76.7	5	77.5
87	-	75.2	1	63.3	-	76.7	-	77.5
88	-	75.2	-	63.3	-	76.7	1	78.4
89	-	75.2	2	65.1	-	76.7	-	78.4
90	11	85.0	10	74.3	9	85.4	9	87.3
91	1	85.8	1	75.2	-	85.4	2	89.2
92	-	85.8	2	77.1	-	85.4	-	89.2
95	3	88.5	10	86.2	7	92.2	2	91.2
96	-	88.5	1	87.2	-	92.2	-	91.2
98	1	89.4	4	90.8	-	92.2	3	94.1
99	2	91.2	2	92.7	2	94.2	2	96.1
100	10	100	8	100	6	100	4	100

Table 5: Updating GPA Beliefs

Dependent Variable:	Change in **Spring 2008 GPA** beliefs						Change in **Graduation GPA** beliefs					
	All			Same Major[a]			All			Same Major		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
△ GPA between the surveys	52.78***	-	32.64	56.12***	-	36.66*	28.43**	-	15.00	30.19**	-	12.12
	(17.26)		(20.75)	(17.91)		(21.34)	(13.31)		(16.07)	(14.12)		(16.73)
Information Content	-	15.88***	10.46*	-	16.56***	10.48*	-	9.47**	6.97	-	11.74***	9.73*
		(5.06)	(6.10)		(5.36)	(6.30)		(3.89)	(4.72)		(4.15)	(5.00)
R-Squared	0.080	0.084	0.105	0.095	0.092	0.120	0.041	0.052	0.060	0.046	0.078	0.084
No. of Observations	109	109	109	96	96	96	109	109	109	96	96	96

Standard errors in parentheses. * sig at 10%; ** sig at 5%; *** sig at 1%

[a] Sample restricted to respondents who have the same major in both surveys

Table 6: Absolute Change in Beliefs as a Function of Uncertainty in Prior Belief

Dependent Variable: Absolute Change in belief for:

	Grad in 4 years	**Grad w/ GPA $\geq$ 3.5**	**Enjoy Courses**	**Parents Approve**	**Find Job**	**Enjoy Work**	**Work Flexible**
Constant	7.84***	13.59***	12.78***	10.23***	14.96***	15.08***	12.13***
	(1.01)	(1.09)	(0.94)	(0.86)	(1.55)	(1.13)	(1.37)
U^a	18.07***	3.52**	7.27***	9.35***	3.26*	4.41***	6.81***
	(3.28)	(1.59)	(1.63)	(1.88)	(1.85)	(1.61)	(1.74)

Each column corresponds to one regression. Regressions include random effects.

Each regression has 341 observations with 117 groups (students).

Cluster standard errors in parentheses. * sig at 10%; ** sig at 5%; *** sig at 1%

[a] $U = 1$ if prior belief (belief in first survey) is ≥ 25 or ≤ 75.

Table 7: Updating in Response to New Information

Dependent Variable: The posterior belief (i.e. belief in the follow-up survey)	**Dropped Mj**[a] N=14	**Least Pref.** N=102	**Next Pref. Mj**[b] N=58	**Second Mj**[c] N=58	**Current Mj** N=109
	Dependent Variable: New Belief about Graduating in 4 years				
Initial Belief (γ)	1.49***(0.27)	0.79***(0.031)	0.89***(0.026)	0.73***(0.038)	0.87***(0.029)
New Info (η)	0.34***(0.070)	0.033 (0.021)	0.012(0.023)	0.035***(0.011)	0.021*** (0.009)
Imp of I (R)	-0.33	0.27	0.12	0.37	0.15
	Dependent Variable: New Belief about Graduating with a GPA of more than 3.5				
Initial Belief (γ)	0.56***(0.089)	0.83***(0.028)	0.86***(0.028)	0.62***(0.052)	0.62***(0.034)
New Info (η)	0.37***(0.064)	0.098***(0.017)	-0.050**(0.024)	0.058***(0.020)	0.11***(0.016)
Imp of I (R)	0.79	0.21	0.17	0.61	0.60
	Dependent Variable: New Belief about Enjoying Coursework				
Initial Belief (γ)	0.84***(0.083)	0.55***(0.033)	0.77***(0.029)	0.53***(0.057)	0.58***(0.034)
New Info (η)	0.18***(0.064)	0.014(0.018)	-0.094***(0.024)	0.039*(0.020)	0.023*(0.011)
Imp of I (R)	0.19	0.81	0.31	0.90	0.73
	Dependent Variable: New Belief about Coursework hrs/week				
Initial Belief (γ)	0.35*(0.19)	0.71***(0.038)	0.49***(0.029)	0.78***(0.074)	0.58***(0.042)
New Info (η)	-0.11***(0.041)	-0.029***(0.010)	-0.013(0.010)	-0.0043(0.018)	-0.0110(0.011)
Imp of I (R)	1.82	0.41	1.01	0.28	0.72
	Dependent Variable: New Belief about Approval of Parents				
Initial Belief (γ)	1.49***(0.098)	0.74***(0.029)	0.93***(0.032)	0.76***(0.054)	0.79***(0.042)
New Info (η)	-0.067(0.050)	-0.010(0.018)	-0.013(0.027)	-0.062***(0.024)	0.0070(0.013)
Imp of I (R)	-0.33	0.36	0.081	0.32	0.26
	Dependent Variable: New Belief about Finding a job				
Initial Belief (γ)	0.33*(0.17)	0.55***(0.033)	0.66***(0.039)	0.54***(0.048)	0.44***(0.048)
New Info (η)	0.097(0.12)	0.010(0.018)	-0.078***(0.026)	0.062***(0.019)	-0.0070(0.018)
Imp of I (R)	2.01	0.83	0.51	0.86	1.25
	Dependent Variable: New Belief about Enjoying working at the jobs				
Initial Belief (γ)	0.31***(0.084)	0.55***(0.032)	0.75***(0.032)	0.42***(0.048)	0.59***(0.043)
New Info (η)	0.26***(0.057)	-0.029(0.017)	-0.022(0.025)	0.0067(0.018)	-0.0082(0.014)
Imp of I (R)	2.29	0.80	0.33	1.39	0.69
	Dependent Variable: New Belief about Reconciling work and family at the jobs				
Initial Belief (γ)	0.13(0.21)	0.79***(0.029)	0.92***(0.042)	0.39***(0.048)	0.54***(0.038)
New Info (η)	0.60***(0.065)	0.060(0.018)	-0.11(0.030)	-0.069***(0.017)	-0.057**(0.015)
Imp of I (R)	6.98	0.27	0.088	1.51	0.87
	Dependent Variable: New Belief about Job hrs/week				
Initial Belief (γ)	0.54***(0.15)	0.86***(0.024)	0.66***(0.040)	0.64***(0.067)	0.47***(0.034)
New Info (η)	0.12**(0.047)	-0.0023(0.009)	0.0025(0.021)	0.019(0.013)	0.011(0.0086)
Imp of I (R)	0.87	0.17	0.51	0.57	1.13
	Dependent Variable: New Belief about Expected Salary at the age of 30				
Initial Belief (γ)	2.83***(0.25)	0.50***(0.022)	0.69***(0.044)	0.56***(0.15)	0.56***(0.041)
New Info (η)	-1044.5***(196.8)	32.67(25.60)	-42.43(38.66)	-343.4***(120.8)	-52.59(42.34)
Imp of I (R)	-0.65	0.99	0.43	0.78	0.78

Standard errors in parentheses. * sig at 10%; ** sig at 5%; *** sig at 1%

Each column within a panel corresponds to one regression.

The posterior beliefs and the initial beliefs are on a scale of 0-100 for the binary outcomes.

a A major that the individual had once pursued

b The second most preferred major for individuals without a second major

c The individual's second major

Table 8: Why Do Students Switch Majors?

Reasons for dropping majors	
The initial major was too challenging	14.10[a]
	(22.62)[b]
The initial major was too easy	1.70
	(6.19)
I did not find the major interesting any more	29.80
	(28.97)
I got interested in something else	29.90
	(29.89)
My parents wanted me to change majors	0.80
	(2.63)
There was peer pressure to change majors	0.80
	(3.40)
Others	31.00
	(29.60)
Number of Observations	14

[a] Each cell is the AVERAGE contribution of the reason for switching majors. Students were asked to assign an integer between 0 and 100 to each reason so that their responses all summed to a 100.
[b] Standard deviation in parentheses

Table 9: Understanding Switching of Majors

Dependent Variable: S_i (dummy =1 if individual switched major)			
	(2)	(3)	(5)
$\triangle$GPA[a]	-0.014[b]	-	-0.008
	(0.082)		(0.087)
Information Content		-0.016	-0.014
		(0.060)	(0.063)
No. of Observations	117	117	117

S_i equals 0.1196 in the sample.
[a]GPA at end of Winter 2007 - GPA at beginning of Fall 2006
[b]Table reports the marginal effects for a unit change in the independent variables. Standard errors in parentheses. * sig at 10%; ** sig at 5%; *** sig at 1%

Table 10: The Nature of Change in Beliefs for Outcomes

Dependent Variable: Change in belief for:	**Grad in 4 Years**	**Grad w/ GPA $\geq$ 3.5**	**Enjoy Courses**	**Course Hrs/Wk**	**Parents Approve**	**Find Job**	**Enjoy Work**	**Work Flexible**	**Job Hrs/Wk**	**Salary at 30**
Constant	1.48	-5.32***	-4.11***	-5.53***	0.39	-0.92	-4.55***	2.05	2.13	14549***
	(1.13)	(2.05)	(1.50)	(1.24)	(1.51)	(2.25)	(1.75)	(2.01)	(1.19)	(5227)
Second Pursued Major	2.36*	3.07	3.72	1.15	-0.19	-0.18	-1.37	1.82	0.097	7155
	(1.42)	(2.50)	(2.44)	(1.41)	(2.82)	(3.13)	(2.69)	(2.56)	(1.73)	(12976)
Second Preferred Major	-1.72	0.78	1.42	2.52*	-5.51*	-1.89	0.55	2.34	-1.03	-13982**
	(1.47)	(2.78)	(3.16)	(1.16)	(2.89)	(2.84)	(3.38)	(3.13)	(2.22)	(6797)
Dropped Major	-8.48*	6.84	-6.20	-0.31	-2.31	0.67	-2.01	1.98	-1.79	-20526
	(5.14)	(4.91)	(4.34)	(3.24)	(3.44)	(2.74)	(7.02)	(7.11)	(3.23)	(27206)
Least Preferred Major	-2.07	-1.23	-5.05*	5.31***	-0.76	-1.24	-8.23***	-1.45	0.45	-19453***
	(2.23)	(2.81)	(2.72)	(1.29)	(2.58	(2.74)	(2.95)	(2.87)	(1.39)	(5565)

Cluster standard errors in parentheses. * sig at 10%; ** sig at 5%; *** sig at 1%

Regressions include random effects. Each of these regressions has 341 observations with 117 groups (students).

The binary outcomes (all outcomes excluding coursework hrs/wk; job hrs/week; salary at 30) are on a 0-100 scale.

Table 11: Statistics on Average Annual Starting Salaries of Northwestern 2007 Graduates

Variable: Avg. Starting Salary of 2007 NU Graduates in Each Category:	Avg. Salary of Males			Avg. Salary of Females			% Diff in Salaries[d]		
Reported by:	**Grad '07**[a]	**Males**[b]	**Females**[c]	**Grad '07**	**Males**	**Females**	**Grad '07**	**Males**	**Females**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Category:									
Natural Sciences	36,286	42,690	50,340	33,200	38,347	45,422.5	8.50	10.17	9.38
	(11,898)	(17,028)	(17,614)	(12,785)	(15,930)	(16,093)		(8.14)	(8.93)
Math & Computer Sc.	68,000	48,929	51,667	54,000	45,857	49,667	20.59	5.73	5.02
	(12,550)	(11,958)	(49,667)	–	(10,523)	(18,016)		(7.34)	(7.76)
Social Sciences I	47,928	38,643	35,666	39,694	36,083	32,758	17.18	6.14	7.41
	(25,337)	(14,940)	(13,337)	(19,716)	(13,669)	(12,100)		(6.47)	(8.72)
Social Sciences II	57,877**	51,518	49,286	52,000	49,185	44,643	10.15	4.39	8.01
	(14,780)	(9,720)	(11,906)	(7,552)	(9,140)	(7,958)		(4.51)	(9.11)
Ethics and Values	31,667	32,900	35,231	25,000	32,500	32,461	2.11	1.45	7.31
	(2,886)	(6,118)	(12,001)	–	(6,587)	(10,316)		(7.44)	(9.04)
Area Studies	40,000	36,318	36,043	46,250	33,476	33,418	−15.62	7.21	6.79
	(12,062)	(11,838)	(12,062)	(15,654)	(11,199)	(11,052)		(8.74)	(7.68)
Literature & Fine Arts	30,142	31,601	32,345	36,800	30,601	29,814	−22.08	2.62	6.59
	(14,485)	(12,775)	(12,259)	(14,953)	(12,688)	(10,124)		(18.33)	(10.55)
Music Studies	30,000	–	30,000	30,400	–	30,000	−1.33	–	0
	(8,660)		–	(8,384.5)		–			
Educ & Social Policy	40,000	50,000	35,000	39,696	45,000	35,000	0.75	10.00	0
	(12,884)	–	–	(12,221)	–	–			
Communication Std	46,456.5	–	40,250	42,362	–	37,000	8.81	–	7.69
	(53,001)		(12,120)	(14,611)		(10,456)			(3.41)
Engineering	56,706*	50,053	54,732	52,671	49,103	51,429	7.11	2.73	5.85
	(10,211)	(15,448)	(20,962)	(12,309)	(21,314)	(20,913)		(18.87)	(11.42)
Journalism	47,875	30,000	30,000	35,694	35,000	30,000	2.54	−16.67	14.28
	(24,375)	–	–	(12,648)	–	–		–	–

Standard errors in parentheses; gender difference is significant at 10%; ** significant at 5%; *** significant at 1% (2-tailed T-test).

a ACTUAL average starting salary of the (male) graduating class of 2007 majoring in that category (Source: Northwestern Graduation Survey 2007)

b Response of Male survey respondents to the question: "What do you think was the average annual starting salary of Northwestern MALE graduates (of 2007) with Bachelor's Degrees in X?"

c Response of female survey respondents to the same question as in *b*

d The AVERAGE % difference between the female and male salary, i.e., $\frac{\text{male salary - female salary}}{\text{male salary}} * 100$

Table 12: Correlates of Errors in Beliefs About Expected Salary

Dependent Variable: Log Absolute Error in Beliefs about:$^{\oplus}$	Starting Salaries for MALES	Starting Salaries for FEMALES	Initial Survey$^{\bigstar}$
	(1)	(2)	(3)
Major Declared[a]	0.298	0.130	-0.096
	(0.187)	(0.168)	(0.182)
Cumulative GPA	-0.344*	-0.235	0.189
	(0.206)	(0.233)	(0.235)
SAT Math	0.0021**	0.0034***	0.00033
	(0.0011)	(0.0012)	(0.0016)
SAT Verbal	0.00041	-0.0022*	-0.0013
	(0.0011)	(0.0013)	(0.0014)
Female	0.365**	0.404**	0.212
	(0.158)	(0.161)	(0.190)
NU Credits	-0.010	0.014	-0.0232
	(0.0237)	(0.022)	(0.0292)
Asian	-0.115	-0.098	-0.106
	(0.248)	(0.301)	(0.280)
Foreign	-0.123	-0.234	-0.020
	(0.314)	(0.366)	(0.569)
Sec-Gen Immigrant	0.163	0.228	0.361
	(0.193)	(0.258)	(0.254)
Studying Major[b]	0.144	-0.105	-0.190
	(0.148)	(0.186)	(0.161)
Studying Major × Decl	-0.288	-0.360	-0.014
	(0.224)	(0.238)	(0.251)
Private High School	-0.093	-0.022	0.121
	(0.164)	(0.161)	(0.156)
Low Parents' Income[c]	0.131	0.068	-0.153
	(0.154)	(0.155)	(0.176)
Father Attended College	-0.616***	-0.289	-0.107
	(0.228)	(0.261)	(0.366)
Mother Attended College	0.351*	0.277	-0.638*
	(0.195)	(0.229)	(0.334)
Father Studied Major[d]	0.044	-0.332	-0.205
	(0.207)	(0.227)	(0.241)
Mother Studied Major[e]	-0.077	-0.194	0.383**
	(0.167)	(0.205)	(0.192)
Resp. Random Eff.	Yes	Yes	Yes
No. of Observations	338	344	341
No. of Clusters	117	117	117
R-Squared	0.0715	0.1127	0.0783

Estimates correspond to OLS estimation. Cluster errors in parentheses;
* sig at 10%; ** sig at 5%; *** sig at 1%

$\oplus$Dep. var is $\ln\left|\frac{\overline{s^G_{im}} - s^G_{obs\_m}}{s^G_{obs\_m}} * 100\right|$; $\overline{s^G_{im}}$ is respondent's belief of the avg. salary of NU 2007 grads of gender G in major m, and $s^G_{obs\_m}$ is the actual avg. salary of 2007 graduates in m.

$\bigstar$ Responses from the first survey. Average salaries were reported unconditional on gender, so dependent variable is $\ln\left|\frac{\overline{s_{im}} - s_{obs\_m}}{s_{obs\_m}} * 100\right|$; $\overline{s_{im}}$ is the respondent's belief of the average salary of NU 2006 graduates in major m, and $s^G_{obs\_m}$ is the observed salary.

a dummy =1 if the respondent had declared her major at the time of the initial survey.
b dummy =1 if the respondent's intended major is same as category m in the question.
c dummy =1 if parents' annual income is less than \$150,000.
d (e) dummy =1 if father's (mother's) field of study is the same as in the salary question.