Federal Reserve Bank of New York
Staff Reports

Can Subjective Expectations Data Be Used in Choice Models? Evidence on Cognitive Biases

Basit Zafar

Staff Report no. 454
June 2010

Can Subjective Expectations Data Be Used in Choice Models?
Evidence on Cognitive Biases
Basit Zafar
Federal Reserve Bank of New York Staff Reports, no. 454
June 2010
JEL classification: D8, I2, J1, J7

Abstract

A pervasive concern with the use of subjective data in choice models is that the data are biased and endogenous. This paper examines the extent to which cognitive biases plague subjective data, specifically addressing 1) whether cognitive dissonance affects the reporting of beliefs, and 2) whether individuals exert *sufficient* mental effort when probed about their subjective beliefs. For this purpose, I collect a unique panel data set of Northwestern University undergraduates that contains their subjective expectations about outcomes specific to different majors in their choice set. I do not find evidence of cognitive biases systematically affecting the reporting of beliefs: By analyzing patterns of belief updating, I can rule out cognitive dissonance being a serious concern in the current setting. Moreover, there seems to be no systematic (nonclassical) measurement error in the reporting of beliefs. In the reported beliefs for the various majors, I find no systematic patterns in mental recall of previous responses or in the extent of rounding. Comparison of subjective beliefs with objective measures suggests that students have well-formed expectations. Overall, the results paint a favorable picture for the use of subjective expectations data in choice models.

Key words: college majors, expectations, cognitive biases, endogeneity, dissonance

---

Zafar: Federal Reserve Bank of New York (e-mail: basit.zafar@ny.frb.org). The author is indebted to Charles Manski and Christopher Taber for extremely helpful discussions. He also thanks Meta Brown, Yonatan Eyal, Carlos Madeira, Wilbert van der Klaauw, Bob Willis, and participants in the 2009 CIRPÉE (Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi) Workshop on Subjective Expectations in Econometric Models, the Econometric Society's 2009 North American summer meeting, and seminars at Northwestern University and the Federal Reserve Bank of New York. The author thanks all those involved in providing the data sets used in this paper. Financial support from a Northwestern University graduate research grant is gratefully acknowledged.

1 Introduction

Understanding any decision made under uncertainty requires one to study how expectations and preferences are used to make the choice. In the absence of data on expectations, existing empirical studies make non-verifiable assumptions on expectations, assume individuals are rational and use the same information-processing rule, and use choice data to infer decision rules conditional on the maintained assumptions about expectations. This approach is problematic because 1) there is little reason to think that individuals with similar information form their expectations in the same way,[1] 2) observed choices may be consistent with several combinations of expectations and preferences (Manski, 1993), and 3) 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), and different rules yield vastly different predictions (Buchinsky and Leslie, 2009).

A solution to this identification problem is to directly elicit subjective beliefs and incorporate them into choice models. However, despite the fact that economists have increasingly been collecting and describing subjective data in the last decade or so (Manski, 2004), few studies incorporate subjective data into choice models (Lochner, 2007; Bellemare, Kroger, and van Soest, 2008; Delavande, 2008; Zafar, 2009). One reason for this is that the expectations data needed as inputs into choice models have rarely been available until recently. The other suggested explanation for this is that subjective data are endogenous (Bertrand and Mullainathan, 2001; Bound et al, 2001; Benitez-Silva et al., 2004). In particular, when estimating choice models that incorporate subjective expectations, the researcher needs to elicit beliefs for outcomes associated with the choice that the individual has made as well as for outcomes associated with the other options in the individual's choice set. One concern is that if the respondent is asked about his beliefs, he is likely to exaggerate them in order to rationalize his choice (cognitive dissonance). Other cognitive biases such as respondents making little mental effort in answering questions, lack of beliefs existing in a coherent form, and social desirability may also affect the way in which individuals report their beliefs. Some of these concerns have been studied in cross-sectional analyses of subjective beliefs (Manski, 2004, and references therein). However, studying issues such as cognitive dissonance requires the researcher to have data on how beliefs

[1] In fact, Stinebrickner and Stinebrickner (2008) and Niederle and Vesterlund (2007) find that, conditional on the available information, expectations about performance differ systematically by ability and gender, respectively.

evolve over time for outcomes associated with the choice that the individual made as well as for outcomes associated with choices that the individual did not choose- such data usually don't exist.[2]

In order to address these questions, I designed and conducted two surveys that elicited subjective expectations from Northwestern University undergraduates regarding the choice of major. The first survey, administered to students in the early part of their sophomore year, collected details on respondents' demographics and data relevant for the estimation of the choice model; these data were used to estimate a choice model of college majors (Zafar, 2009). The second survey, conducted about a year after the first, collected data on how individuals revise their beliefs for major-specific outcomes. The major-specific outcomes for which beliefs were elicited include both outcomes that are realized in college and those that are realized in the workplace. Examples of the former include graduating in 4 years, enjoying the coursework, and parents approving of the choice, while examples of the latter include outcomes such as finding a job upon graduation, and being able to reconcile work and family at the jobs.

One concern with subjective data is that individuals could report beliefs that are consistent with their behavior, i.e., cognitive dissonance may affect the subjective data (Bound et al., 2001; Mullainathan and Washington, 2008). Here, it would imply that if an individual never pursued a major, they would tell themselves that they never liked it anyway. Therefore, one would observe *unfavorable* changes in beliefs for outcomes in majors that an individual never pursued, and similarly large *favorable* changes in beliefs for outcomes for the major that the individual has stuck with. This systematic biasing of beliefs would be especially problematic if subjective data were used to estimate choice models. In section 2, I show that this would cause the estimated parameters to be biased upwards. The panel on subjective beliefs allows me to check if cognitive dissonance affects the students' beliefs. Analysis of changes in beliefs suggests that this is not a serious concern: Average changes in beliefs for most outcomes in

[2] There are, however, longitudinal studies that find that beliefs respond to changes in one's environment in meaningful ways. Examples include Bernheim (1988), Dominitz (1998), Dominitz and Manski (2005), Hurd and McGarry (2002) and Lochner (2007) who study revisions to expectations about social security benefits, income, mutual-fund investments, survival, and arrest, respectively.

Cross-sectional dispersion in data from two different cohorts, one that has already made the choice and one that is yet to make the choice, could also be used to study cognitive dissonance. However, this would require making an assumption of stationarity of the cross-sectional distribution of beliefs across cohorts. Moreover, it's not clear how much of the difference in cross-sectional dispersion between the two cohorts could be attributed to one cohort (the one that has already made the choice) having more information than the other.

One alternative to using survey data to test for cognitive biases in expectations data is to use an experimental approach. For example, Offerman, Sonnemans, and Schram (1996) use an experimental public goods design to examine biases in expectations. Such an approach is not feasible in surveys.

a student's non-pursued majors (in this case, the student's least preferred major and second preferred major) are not too different from those in the student's pursued major(s). It should be pointed out that this would not be a valid test for cognitive dissonance if students already reported beliefs that were consistent with their choices in the initial survey. Since, nearly half of the students were still undecided on their major when first surveyed, I also analyze the revisions in beliefs across different majors conditional on whether the student had declared his major when first surveyed. I fail to find strong evidence of cognitive dissonance affecting the reporting of beliefs for either of the groups.

The second bias that this paper focuses on is insufficient mental effort and the lack of beliefs existing in a coherent form. Subjective expectations data, like most data, have measurement error. However, any idiosyncratic rounding would yield biased estimates if such data were used directly to estimate choice models as in Delavande (2008) and Zafar (2009). The empirical analysis shows that individuals adopt similar rounding practices when reporting their beliefs for outcomes associated with the various majors in their choice set, implying that such data can be used in choice models. Most students successfully recall their previous responses, and I do not find evidence of systematic biases in recall for outcomes associated with one's own major and with the other majors in one's choice set. As a final test to assess whether students exert sufficient mental effort when reporting their beliefs, I compare the subjective data with objective measures. Since realizations data for students that are surveyed do not exist, I compare their beliefs to realizations of previous cohorts and find that the subjective data match up well with objective measures.[3] For example, in the case of expected salary in the various majors, responses match up well with objective realities and students seem to be aware of income differences across majors.

On the whole, the results in this paper bode well for the use of subjective data in choice models. I do find any strong evidence of cognitive dissonance confounding the data. I fail to find evidence of insufficient mental effort on part of the respondents: (1) students revise their beliefs in meaningful ways; (2) there are no systematic biases in recall of past responses, and

[3]Several studies have explored the accuracy of subjective data. One approach is to compare expectations data with realizations data: For example, Dominitz (1998), Smith, Taylor and Sloan (2001), and Hurd and McGarry (2002) show that expectations tend to be useful predictors of future outcomes/behavior. The second approach compares subjective data with objective measures: For example, Bruine de Bruin et al. (2000) show that teen expectations for various events tend to match up well with objective measures; Delavande, Gine and McKenzie (2010) provide a review of various studies that collect subjective data in developing country settings and conclude that the evidence is mixed with regards to the accuracy of such data. I use this second approach in this paper.

(3) there is no differential rounding of responses. Since it's not possible to directly observe individuals' thinking, I cannot refute the argument that individuals don't reveal their true beliefs.[4] However, I show that individuals give internally consistent and sensible responses. However, one should be careful in generalizing these findings because the setting in this study is such that cognitive biases should be minimized: (1) students are surveyed at a time when they're actively thinking about the choice of major. Since the decision of what to major in has consequences, one would imagine students to have seriously thought about the likelihood of major-specific outcomes and to have well-defined expectations, and (2) the choice of major is reversible, which would imply that biases such as cognitive dissonance would be less of an issue. Nonetheless, this is one of the few studies that has the data needed to address these issues without making any assumptions on the underlying data-generating process, and offers a framework with which one can examine some of the cognitive issues. More studies that collect and examine subjective data from different settings are needed before a definite word can be reached on this subject.

The paper is organized as follows: Section 2 outlines the choice model framework and formalizes the different biases. Section 3 describes the sources of data used in this study. Section 4 empirically investigates the extent to which various cognitive biases affect subjective data. Finally, Section 5 concludes.

2 Theoretical Framework

The choice decision that I consider in this paper is that of college major. However, the general framework would apply to any choice under uncertainty. Individual i is confronted with the decision to choose a college major from his choice set C_i. Individuals are forward-looking, and their choice depends not only on the current state of the world but also on what they expect will happen in the future. Individual i derives utility $U_{ik}(\mathbf{b}, \mathbf{d})$ from choosing major k. Utility is a function of a vector of choice-specific discrete outcomes $\mathbf{b}$ and a vector of continuous outcomes $\mathbf{d}$.[5] Examples of outcomes in $\mathbf{b}$ include graduating from college within four years, gaining approval of parents, and ability to reconcile family and work at the jobs. Examples of outcomes in $\mathbf{d}$ include future income and number of hours spent on coursework. Both vectors, $\mathbf{b}$ and $\mathbf{d}$, are uncertain at the time of the choice, and individual i possesses

[4] This concern, however, is not specific to expectations data, and also applies to other survey research.

[5] One could allow the utility to be a function of individual characteristics as well.

subjective beliefs $P_{ik}(\mathbf{b}, \mathbf{d})$ about the outcomes associated with choice of major k $\forall k \in C_i$.[6] If an individual chooses major m, then standard revealed preference argument (assuming that indifference between alternatives occurs with zero probability) implies that:

$$m \equiv \arg\max_{k \in C_i} \int U_{ik}(\mathbf{b}, \mathbf{d}) dP_{ik}(\mathbf{b}, \mathbf{d}). \tag{1}$$

The goal is to infer the preference parameters from observed choices. However, the expectations of the individual about the choice-specific outcomes are also unknown. The most one can do is infer the decision rule conditional on the assumptions imposed on expectations. This would not be an issue if there were reasons to think that prevailing expectations assumptions are correct. However, not only has the information-processing rule varied considerably among studies of schooling behavior, but most assume that individuals process information in the same way. First, there is little reason to think that individuals with the same information set form the same expectations. Second, different combinations of preferences and expectations may lead to the same choice (Manski, 2002). A solution to this identification problem is to elicit subjective beliefs directly from individuals. Because economists have been skeptical of using subjective data, few studies have used this approach in the estimation of choice models (see, for example, Lochner, 2007; Bellemare, Kroger, and van Soest, 2008; Delavande, 2008; Zafar, 2009). One reason for this skepticism is that cognitive biases may severely affect the reporting of beliefs (Bertrand and Mullainathan, 2001; Bound et al., 2001). Since identification of the preference parameters in equation (1) requires the elicitation of the respondent's expectations regarding the major that he has chosen as well as expectations regarding the *other* majors in his choice set that he could have chosen, concerns about the validity of subjective data are further exacerbated. Below I discuss some of the cognitive problems in more detail.

For simplicity, I assume that utility in linear and separable in outcomes. In that case, $U_i(\mathbf{b}, \mathbf{d}) = \sum_{r=1}^{R} u_r(b_r) + \sum_{q=1}^{Q} \gamma_q d_q + \varepsilon_{ik}$, where $u_r(b_r)$ is the utility associated with the binary outcome b_r, γ_q is a constant for the continuous outcome d_q, R (Q) is the number of binary (continuous) outcomes, and ε_{ik} is a random term. Equation (1) can now be written as:

$$m \equiv \arg\max_{k \in C_i} \left(\sum_{r=1}^{R} \int u_r(b_r) dP_{ik}(b_r) + \sum_{q=1}^{Q} \gamma_q \int d_q dP_{ik}(d_q) + \varepsilon_{ik} \right).$$

[6] The vectors $\mathbf{b}$ and $\mathbf{d}$ are the set of outcomes common to all majors. It is the joint probability distribution of these outcomes $P_{ik}(\mathbf{b}, \mathbf{d})$ which is indexed by major k.

The additive separability of the utility function implies that only the marginal distribution of beliefs about the outcomes enter the expected utility. For the binary outcomes ($\{b_r\}_{r=1}^{R}$):

$$\begin{aligned}\int u_r(b_r)dP_{ik}(b_r) &= P_{ik}(b_r=1)u_r(b_r=1)+[1-P_{ik}(b_r=1)]u_r(b_r=0)\\ &= P_{ik}(b_r=1)\triangle u_r+u_r(b_r=0),\end{aligned}$$

where $\triangle u_r \equiv u_r(b_r=1)-u_r(b_r=0)$, i.e., it is the difference in utility between outcome b_r happening and not happening. The linearity assumption of the utility function implies that only the expected value of the continuous outcomes matters since $\int U_i(\mathbf{b},\mathbf{d})dP_{ik}(\mathbf{b},\mathbf{d}) = U_i(\int \mathbf{b},\mathbf{d}\ dP_{ikt}(\mathbf{b},\mathbf{d}))$. Thus, for the continuous outcomes ($\{d_q\}_{q=1}^{Q}$), $\int d_q dP_{ik}(d_q)$ equals $E_{ik}(d_q)$, the expected value of the outcome. The expected utility that individual i derives from choosing major m is:

$$\begin{aligned}&U_{im}(\mathbf{b},\mathbf{d},\{P_{im}(b_r=1)\}_{r=1}^{R},\{E_{im}(d_q)\}_{q=1}^{Q}) =\\ &\sum_{r=1}^{R}P_{im}(b_r=1)\triangle u_r+\textstyle\sum_r u_r(b_r=0)+\displaystyle\sum_{q=1}^{Q}\gamma_q E_{im}(d_q)+\varepsilon_{im}.\end{aligned} \tag{2}$$

An individual i with subjective beliefs $\{P_{ikt}(b_r),P_{ikt}(d_q)\}_{r,q}\ \forall k\in C_i$ chooses major m with probability:

$$\begin{aligned}&\Pr(m|\{P_{ikt}(b_r),E_{ikt}(d_q)\}_{r,q;\ k\in C_i}) =\\ &\Pr\left(\begin{array}{l}\displaystyle\sum_{r=1}^{R}P_{im}(b_r=1)\triangle u_r+\sum_{q=1}^{Q}\gamma_q E_{im}(d_q)+\varepsilon_{im}\\ \displaystyle>\sum_{r=1}^{R}P_{ik}(b_r=1)\triangle u_r+\sum_{q=1}^{Q}\gamma_q E_{ik}(d_q)+\varepsilon_{im}\end{array}\right)\\ &\qquad\qquad\qquad\qquad\forall k\in C_i,\ m\neq k.\end{aligned} \tag{3}$$

In equation (3), $\{\triangle u_r\}_{r=1}^{R}$ and $\{\gamma_q\}_{q=1}^{Q}$ are the parameters of the utility function that need to be estimated; $\triangle u_r$ is the change in utility from the occurrence of outcome b_r, while γ_q is the parameter in the utility function for the continuous outcome d_q. $\{P_{ik}(b_r=1)\}_{r=1}^{R}$ and $\{E_{ik}(d_q)\}_{q=1}^{Q}$ are elicited directly from the respondent $\forall k\in C_i$.

The first concern with regards to using subjective data in choice models is **cognitive dissonance** (Festinger, 1957). Cognitive dissonance implies that individuals report attitudes that are consistent with their behavior. In the current setting, this bias would imply that if an individual has chosen a major m, he is likely to upgrade beliefs for outcomes associated with that major and to devalue beliefs about outcomes associated with other majors in his choice set. Let

$P_{ik}(b_r = 1)$ denote i's true belief about the likelihood of outcome b_r $\forall k \in C_i$, and $P^*_{ik}(b_r = 1)$ the reported value. For desirable (undesirable) outcomes, cognitive dissonance would imply that $P^*_{im}(b_r = 1)$ is greater (less) than $P_{im}(b_r = 1)$, and $P^*_{ik}(b_r = 1)$ is less (greater) than $P_{ik}(b_r = 1)$ $\forall k \in C_i$ and $k \neq m$. Therefore, a consequence of this bias would be that the estimated coefficients (in this case, $\triangle u_r$) would be *upward* biased.

A second potential problem is respondents making **insufficient mental effort** when reporting their beliefs or having **undefined and ambiguous beliefs** (opposed to well-formed expectations), in particular for outcomes associated with majors they did not choose. Subjective expectations data, like most data, have measurement error. Studies using subjective expectations have documented the fact that respondents tend to round to the nearest 5 when reporting their beliefs on a scale of zero to 100 (Manski, 2004). However, the concern is that the respondent's beliefs for outcomes associated with rejected majors in his choice set may exhibit greater noise or may be all noise if they are undefined or if they do not exist in a coherent form. If such data were used in choice models, this *systematic rounding* (noise) would yield biased estimates. The bias arising from lack of mental effort would depend on the rounding practice of the respondent.[7]

Other concerns with using subjective data include social desirability, i.e., respondents giving the socially acceptable response in order to avoid looking bad in front of the interviewer. However, in the data used in this paper, response distortion due to social desirability was mitigated by making the questionnaires anonymous, and by making respondents answer the survey online so that they didn't have to answer directly to an interviewer. Therefore, this bias is not empirically relevant for the setting in this paper.

Another potential concern is that the expectations data may be statistically correlated with the unobserved error term in a random utility model (the ε term above), and hence may be endogenous. One way to deal with this endogeneity is to allow the reported belief about the likelihood of outcome b_r for choice k, i.e, $P^*_{ik}(b_r = 1)$, to deviate from the true belief, $P_{ik}(b_r = 1)$,

[7] Here, I do not deal with the issue of non-systematic measurement error in the reporting of beliefs. Non-systematic measurement error would lead to bias towards zero in the case of classical measurement error, and no bias if the respondent reports his best estimate based only on the noisy measure of the mismeasured variable itself (see Hyslop and Imbens, 2001, for details).

There are two alternatives to dealing with (non-systematic) measurement error in subjective data: 1) infer the respondent's rounding practice, and interpret the numerical responses as intervals (Manski and Molinari, 2008), and then conduct the statistical inference by treating the subjective data as interval data (Manski and Tamer, 2002), or 2) modelling the preferences and beliefs jointly. This approach has been used by Lochner (2007) and Bellemare et al. (2008). Such an approach is, however, not feasible in models incorporating beliefs for several outcomes (as in Delavande, 2008, and Zafar, 2009).

because of idiosyncratic rounding error, μ_{ik}. In that case, $P^*_{ik}(b_r = 1) = P_{ik}(b_r = 1) + \mu_{ik}$. Unbiased estimates of the preference parameters can then be obtained if the idiosyncratic error term, μ_{ik}, is allowed to be arbitrarily correlated with the unobserved random term in the utility model, ε_{im}. This is primarily an econometric issue, and since this paper deals with cognitive biases, I don't discuss it in detail here.[8]

The rest of the paper empirically explores the extent to which these cognitive issues plague the data.

3 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.[9] 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 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 two 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, 2009).

[8] Interested readers should refer to Lochner (2007) and Bellemare et al. (2008) who address this issue in their econometric models that include subjective data.

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

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.[10]

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 1 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 who took 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 1 also shows that the respondents' distribution of (intended) majors in the Weinberg College of Arts and Sciences (WCAS) is similar in the two surveys, suggesting no differential attrition by field of study. As shown in Table 1, 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 since I am primarily interested in analyzing how beliefs change over time, 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.

The follow-up survey primarily 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),[11] 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 ma-

[10] When taking the initial survey, students were not aware that there could be a potential follow-up survey, and their contact information was collected at the end of the survey only if they wanted to be contacted for a follow-up.

[11] The College of Arts and Sciences at Northwestern University consists of 41 majors. Similar majors were pooled together. Table A1 shows the categorization of majors.

jor 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 set of major-specific outcomes consists of binary outcomes, $\mathbf{b}$, and continuous outcomes, $\mathbf{d}$. The vector $\mathbf{b}$ includes the outcomes:

b_1 successfully completing (graduating) a field of study in 4 years

b_2 graduating with a GPA of at least 3.5 in the field of study[12]

b_3 enjoying the coursework

b_4 parents approve of the major

b_5 obtain an acceptable job immediately upon graduation

b_6 enjoy working at the jobs available after graduation

b_7 able to reconcile work and family while at the available jobs

while the vector $\mathbf{d}$ consists of:

d_1 hours per week spent on the coursework

d_2 hours per week spent working at the available jobs

d_3 social status of the available jobs[13]

d_4 income at the available jobs

The survey elicited the probability of the occurrence of the binary outcomes, i.e., $\{P_{ikt}(b_r = 1)\}_{r=1}^{7}$ and the expected value for the continuous outcomes, i.e., $\{E_{ikt}(d_q)\}_{q=1}^{4}$. As mentioned above, 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,

[12] 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.

[13] The initial survey elicited social status of available jobs 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.

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 (2008). An excerpt of the survey containing the introduction and list of questions dealing with the major-specific outcomes can be found in the Appendix. The full survey questionnaires are available on request from the author.

Before analyzing the extent to which the various cognitive issues plague the data, I briefly describe the subjective data. Table 2 shows the mean belief reported in the initial survey for each of the eleven outcomes for the eight main major categories. The table reports the means conditional on whether the student is majoring in that category or not. There is substantial variation in mean belief for the same outcome across the various major categories, indicating that students do perceive differences in the occurrence for these outcomes across majors. For example, the mean belief of being able to graduate in 4 years varies from about 0.82 (on a 0-1 scale) for Engineering to 0.97 for Literature and Fine Arts. The table also shows that, for most outcomes and particularly for those realized in college, students majoring in the category report higher mean beliefs, i.e., they are more sure about the likelihood of the outcomes relative to students not majoring in that category. More optimistic beliefs about outcomes in one's own major do not necessarily imply cognitive dissonance. If students were sorting into majors (Arcidiacono, 2004), one would similarly observe students who decided in favor of a major to report higher positive values while those deciding against that major reporting lower values. Therefore, it shouldn't be surprising that students report more favorable beliefs about positive outcomes for their more preferred (chosen) majors. Thus, analyzing the level of beliefs would not be useful in discerning the extent of the various biases. Instead, in order to analyze the various cognitive biases, the analysis in the next section would primarily focus on the change in reported beliefs between the first survey and the follow-up.

The mean beliefs reported in Table 2 mask the heterogeneity in responses across respondents for the *same* outcome. Figure 1 presents the histogram and cumulative belief distribution of enjoying coursework in one's current major, while the bottom panel presents the corresponding distribution for one's least preferred major. The figure shows that: (1) Students' beliefs exhibit substantial heterogeneity and they use the entire scale from zero to 100; (2) Students revise their beliefs between the two surveys. For both the most and least preferred major, the belief distribution in the initial survey first order stochastically dominates the belief distribution in the final survey, i.e., over time, students revise their beliefs for enjoying coursework downward; and (3) As one would expect, the belief distribution in the case of the most preferred major is skewed left relative to the distribution for one's least preferred major, i.e., students have relatively more favorable beliefs about outcomes in their more preferred major. Figure 2 presents the belief distribution about expected salary at age 30. As in Figure 1, students exhibit substantial heterogeneity and revise their beliefs between the two surveys. However, in this case, the final survey belief distribution first order stochastically dominates the distribution from the initial survey for both major categories; students revise their beliefs upward about expected income for both their most preferred major as well as their least preferred major.

4 Empirical Analysis

This section explores the empirical nature of the various issues mentioned in Section 2.

4.1 Cognitive Dissonance

Cognitive dissonance would imply that individuals revise their beliefs to preserve the positions that they are most committed to, and in a way that is consistent with their behavior (Festinger, 1957). For example, Mullainathan and Washington (2008) find evidence of cognitive dissonance in political attitudes; they find that opinion ratings of politicians reported by people eligible to vote exhibit greater polarization that those of comparable ineligibles. In the current context, cognitive dissonance would imply that one would observe larger unfavorable changes in beliefs between the two surveys for outcomes in majors that an individual never pursued (the least preferred major and second preferred major), and similarly larger favorable (or at least, less unfavorable) changes in beliefs for outcomes for the major(s) that the individual has stuck with. As mentioned in Section 2, distortion of responses due to cognitive dissonance would cause model estimates to be upward biased.

The first row in each panel of Table 3 reports the mean change in the belief for the binary outcomes (disaggregated by how the individual ranks the major). Table 4 reports the corresponding statistics for the continuous outcomes.[14] The tables also report the mean absolute

[14] The belief about the "social status of the jobs" is excluded from the analysis because it was elicited differently in the two surveys. The first survey elicited an ordinal ranking of the majors according to social status, while the follow-up survey elicited the belief (on a 0-100 scale) of being able to get a job with a high social status.

change of beliefs in parentheses and the fraction of responses which have remain *unchanged* since the initial survey in square brackets.[15] The mean change in beliefs for almost all the binary outcomes is less than 10. Changes in beliefs for outcomes such as graduating in 4 years and parents' approval for a given major are smaller than those for other binary outcomes. In fact, one would expect individuals to have fairly precise beliefs about the occurrence of these outcomes at the time of the initial survey and, therefore, to give similar responses in the follow-up. Similarly, as one would expect, mean changes in beliefs for outcomes associated with a dropped major are larger in magnitude than the corresponding changes in other major categories.

The average change in beliefs of outcomes in an individual's least preferred major and current major are not too different from those in other major categories. Moreover, the direction in which beliefs are revised are fairly consistent across the major categories. As depicted in Table 3, beliefs about enjoying coursework and enjoying working at the jobs are revised downward in all major categories, and not only for the least preferred major. This is reassuring since cognitive dissonance would have implied favorable revisions for outcomes associated with one's current major.

Panel A of Table 5 presents the information shown in Table 3 in a slightly different way. The change in beliefs for each outcome is regressed onto dummies for the different major categories (second preferred major, second major, dropped major, least preferred major). 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. On average, students revise their beliefs of graduating with a GPA of more than 3.5, enjoying coursework and enjoying work downward in their current major. For the dropped major, individuals revise their beliefs downward for all outcomes except graduating with a GPA of more than 3.5 and work flexibility, suggesting that students report and revise their beliefs meaningfully. The coefficients on the "Least Preferred Major" and "Second Preferred Major" dummies are of interest to test for the presence of cognitive dissonance. Presence of cognitive dissonance would imply that these estimates are significantly different from zero. Of the twenty estimates, only three (coursework hours/week, enjoying work at the jobs, and expected salary at age 30 for the least preferred major) are significantly different from zero at confidence levels

[15] I define the belief of an outcome to have remain *unchanged* between the two surveys if: (1) the absolute change in beliefs is less than 5 points (on a scale of 0-100) for binary outcomes; (2) the absolute change in beliefs is less than 5 for hrs/week spent on coursework or job; (3) the absolute change in beliefs for salary is less than \$5000.

of 95% or higher. This suggests that biases arising from cognitive dissonance in the data are not severe.

It should be pointed out that if, in the initial survey, students reported beliefs that were already consistent with their choices, then revisions in beliefs for various outcomes across different majors would be similar and analyzing temporal patterns in changes in beliefs would not be useful in determining the presence of cognitive dissonance. In the current setting, it is highly implausible that students reported choice-consistent beliefs in the first survey because, when initially surveyed, their choice of college major was reversible as they could easily switch majors until the end of the sophomore year. However, the data allow me to actually test for this concern. As noted in Table 1, 61 of the 117 respondents had already declared their major at the time of the initial survey. Therefore, the revision of beliefs can be analyzed separately for these two groups. Panel B (C) of Table 5 reports the estimates of the change in beliefs regressed onto the different dummies for the group of students who had declared (not declared) their major at the time of the initial survey. Mean revisions for outcomes across all majors tend to be larger for the group of students that were undeclared when initially surveyed, consistent with them being more uncertain at the time of the initial survey. Since students who had already chosen their majors when surveyed for the initial survey may have reported beliefs consistent with their choices, the purpose of breaking up the sample is to focus on the changes in beliefs for the group that had not declared their major (Panel C of Table 5). The estimates on the least preferred and second preferred dummy indicate that cognitive dissonance is not a major concern for this group either: Of the twenty estimates, only three are significantly different from zero (coursework hours/week and enjoying work at the jobs for the least preferred major, and approval of parents for the second preferred major).

4.2 Insufficient Mental Effort

As mentioned in Section 2, differential mental effort when reporting beliefs for outcomes for different majors would lead to biased model estimates. This section checks for this bias in different ways.

4.2.1 Differential Rounding

Subjective data, like other survey data, have measurement error associated with them. Studies that use subjective data have documented the fact that responses tend to be rounded to the nearest tenth or fifth (Manski, 2004; Manski and Molinari, 2010). If the respondent's rounding practice can be inferred, statistical inference can be conducted by treating the subjective data as interval data. However, if the respondent uses differential rounding practices when reporting beliefs for his chosen major and for the unchosen majors in his choice set, estimation of decision models that employ subjective data would be extremely challenging. The fourth row in each panel of Tables 3 and 4 reports the proportion of responses in the follow-up survey which are not multiples of 5 in curly brackets for the binary and continuous outcomes, respectively. There appears to be a fair amount of rounding to the nearest fifth. Responses seem to be rounded more for a given outcome as one moves from the right-most column to the left, i.e., more responses are rounded for outcomes associated with the least preferred major than for the second most preferred major, which in turn has more responses rounded than for outcomes associated with the current major. This kind of rounding practice would be consistent with individuals making less mental effort when reporting their attitudes for outcomes associated with majors that they've not pursued.

Several studies have documented the reporting of responses at one-percent intervals at the extremes, i.e., 0, 1, 2 and 98, 99, 100 (Dominitz and Manski, 1997; Manski, 2004). One would expect to observe more responses at the extremes if there's less uncertainty associated with that outcome; this is plausible in the case of outcomes associated with one's actual major(s). The fifth row of each panel in Table 3 reports the proportion of responses in the follow-up survey that are not multiples of 5 in italicized curly brackets, i.e. *{.}*, after excluding responses in the extremities. For this purpose, responses that are ≤ 5 or ≥ 95 on a 0-100 scale are excluded from the analysis. The extent of rounding is now similar across the various major categories. This suggests that there is no evidence of systematic rounding on part of the respondents, and that it is safe to conclude that such data can be directly used for choice analysis.

4.2.2 Mental Recall

Another way in which I assess whether respondents exert sufficient mental effort when reporting their beliefs is by checking if they are aware of how their response in the follow-up survey compares to their response in the initial survey. The idea behind this exercise is that if respondents don't exert enough effort in reporting their beliefs, they would not be able to successfully report how their beliefs have changed over time. More specifically, individuals were asked about their *perceptions* of how their beliefs had changed since the initial survey.[16] The wording of the questions was as follows:

> "This question asks you to recall your beliefs i.e. responses to the questions you answered about Major X in the previous survey. Try to recall your beliefs from a year ago about the various outcomes in Major X, and then report whether YOU THINK your current beliefs are HIGHER, LOWER, or ABOUT THE SAME as the old (a year-ago) beliefs."

The elicited beliefs for the various outcomes in the two different surveys tells us how the beliefs *actually* changed between the two surveys. Table 6 presents a matrix of perceived changes in beliefs versus actual changes in beliefs.[17] Each cell shows the number of responses (out of 117) that fall in that category. Perfect recall of earlier responses would imply that all off-diagonal cells would be zero- that is not the case. However, the number of occurrences of *absolute* error in recall (i.e., an individual perceiving their response to have increased when in fact it decreased, and vice versa) are only a small fraction of total responses. For example, in the case of beliefs about graduating in 4 years, the number of such occurrences is zero for the current major and second major, and 7 for the least preferred major. More importantly, there is no systematic pattern in which individuals make errors. One would be concerned if more errors in recall were made when reporting beliefs for the least preferred major or second major relative to those for the current major since this could imply that individuals make less mental effort when reporting their beliefs for outcomes not associated with their own major. That is, however, not the case. For example, in the case of beliefs of finding a job, the number of *absolute* errors is 3, 10, and 13 for the least preferred major, the second major, and the current major respectively. Table 6, however, shows that reported beliefs *actually* changed for a large fraction of respondents who perceived no change in their beliefs. One possible explanation for this is that respondents tend to round their beliefs.

[16] This question was asked *after* the respondent had reported his beliefs for the major-specific outcomes. This order negates the concern that students might become more conscious and attentive when reporting their beliefs if they know they will be asked about recalling their beliefs (and hence decrease the incidence of cognitive dissonance).

[17] For this table, I have pooled responses for the least preferred major and dropped major into a single category, and similarly responses for the second major and second preferred major into a single category.

4.2.3 Undefined Expectations

Another concern is that individuals may not have well-formed expectations. One way in which this may affect the responses is that, when forced to respond to the interviewer about some question for which the respondent has not made a reasonable probability assessment, he might answer 50%. This is what Bruine de Bruin et al. (2000) call epistemic uncertainty.[18] Analysis of the data reveals that the 50% response is the not the most frequent one in majority of the cases. This can also be seen in the histogram reported in Figure 1 which shows the distribution of beliefs about enjoying coursework in the two surveys for the most and least preferred major. The 50% response is not the unique mode in any of them.

Another way to test whether students have well-formed expectations is to check for their "accuracy", i.e., how they compare with objective realities. It should be pointed out that subjective data need not be "accurate" to be used in choice models. However, if they line up well with objective measures, that would suggest that individuals exert sufficient mental effort when reporting their beliefs and that their expectations are well-defined.[19] At least two different ways have been used to assess the validity of subjective expectations: (1) comparing elicited expectations with future realizations,[20] and (2) comparing elicited expectations with historical realizations.[21] This study uses the second approach since I don't observe realizations for most outcomes about which expectations are elicited. Moreover, it is not possible to assess the accuracy of non-pecuniary outcomes such as approval of parents or enjoying coursework since no objective measures exist for these outcomes.

When evaluating how the subjective data compare with the various objective measures that I use, it should be pointed out that there are at least four legitimate reasons why respondents' expectations may be different from them. First, Northwestern University undergraduates are a very specific demographic and the comparison groups that I've used might not be appropriate. Second, respondents might think that future distributions for the event of interest will differ from the current (or past) ones. Third, respondents may have private information about themselves which justifies them having different expectations. Fourth, the objective measures

[18] This bias is an example of what Bertrand and Mullainathan (2001) term as lack of attitude in a coherent form.

[19] Another reason why subjective data may be compared to objective measures is to test how well assumptions like rational expectations explain real decision makers. That is, however, not the purpose of this paper.

[20] This is the approach used by, for example, Dominitz (1998) and Hurd and McGarry (2002). The former study finds that income expectations predict actual income realizations, while the latter study finds that subjective survival probabilities predict actual survival.

[21] This approach has been used, for example, by Bruine de Bruin et al. (2000) who study teen expectations for several significant life events and conclude that they are sensible.

correspond to outcomes for students who choose to pursue that major. In this study, since beliefs are elicited from an individual about the occurrence of the various outcomes in his current major as well as for other majors in his choice set which he considered but did not choose, using data on realizations of students who choose that major may not be the *correct* objective measure. However, since these are the only data available, I use them for comparison purposes. To address the concern of self-selection, I also report the mean beliefs of respondents conditional on majoring in the category.

Table 7 compares the mean belief about graduating with a GPA of at least 3.5 and about expected income at the age of 30 in the various majors with realizations of bachelor graduates from institutions that are similar to Northwestern University. Ideally one would like to see a similar comparison for other outcomes but such data are not readily available. Column (1a) of Table 7 shows the mean GPA by major category of bachelor graduates (of selective Doctoral/Research universities) in the 2001 Baccalaureate & Beyond Longitudinal Study (B&B 2001), and column (1b) ranks the majors according to their GPA. Columns (2a) and (2b) provide the survey respondents' mean belief of being able to graduate with a GPA of at least 3.5 and the ranking of the majors in this dimension, respectively. The relative ranks of majors according to their GPA are similar in my sample and the B&B 2001, suggesting that students are aware of the relative difficulty of the various majors. Columns (2c) and (2d) report the mean beliefs of graduating with a GPA of at least 3.5 for students with a major in the category and students with no major in that category, respectively. Compared to the mean belief in column (2d), the corresponding belief is higher (more favorable) in column (2c), i.e., students with a major in the category tend to report more favorable beliefs. This is consistent with self-selection into majors. However, for both subsamples, the relative beliefs match up well with the relative difficulty of majors as reported in column (1b). Columns (3)-(4) of the table report the corresponding statistics for expected income at the age of 30. The objective measure in this case is the 2003 average annual salary of 1993 college graduates of selective colleges (Carnegie code 4) from the B&B 1993/2003 Study. The relative ranking of majors by income reported by respondents are similar to that computed using the B&B sample, indicating that students correctly perceive income differences across majors. In the case of expected income, comparison of columns (4c) and (4d) shows that students majoring in the category do not always report more favorable beliefs relative to students not majoring in the category. However, the relative beliefs are similar to the objective measures.

Item non-response would be another indication of beliefs not existing in a coherent form or respondents not exerting sufficient mental effort when probed about their beliefs. Non-response is not an issue in this study since there are hardly any instances of students not answering a question.

Overall, the results in this section are consistent with students having well-formed expectations and do not support the hypothesis that students exert insufficient mental effort when reporting their beliefs.

5 Conclusion

This paper investigates a very specific question: Can subjective expectations data be used in choice models? This question is motivated by recent empirical work that underscores the importance of expectations in situations that involve uncertain outcomes, in particular schooling choices (Cunha, Heckman, and Navarro, 2004). Economic models of schooling choices usually make assumptions about how students form expectations. This is problematic because different information-processing rules yield significantly different predictions about individuals' schooling choices (Buchinsky and Leslie, 2009). A solution to this problem is to directly elicit subjective expectations data from the individuals. Though economists have increasingly undertaken the task of collecting subjective expectations data (Manski, 2004), concerns still remain about their use in decision models. This paper specifically addresses these concerns by investigating the extent to which cognitive biases such as social desirability, cognitive dissonance and insufficient mental effort plague subjective data. For this purpose, I collect a unique panel dataset of Northwestern University undergraduates which contains their subjective expectations about various major-specific outcomes.

The results in this paper bode well for the use of subjective expectations. Students revise their beliefs for various outcomes in meaningful ways (for example, beliefs for outcomes associated with dropped majors are revised down), and revisions of beliefs associated with different majors tend to be in the same direction. Analysis of how beliefs evolve over time reveals that biases like systematic rounding and cognitive dissonance do not confound the data. Comparison of elicited beliefs with objective measures (in this case, realizations of previous cohorts) reveals that students are aware of income differences across college majors as well as of differences in how academically challenging the various majors are, suggesting that students have well-defined expectations. The finding of no systematic biases in recall of previous beliefs also lends support to the hypothesis that students exert sufficient mental effort when reporting their beliefs.

This paper adds to the literature on the validity and extent of bias of self-reported survey data (Bound, Brown, and Mathiowetz, 2001; Buchinsky and Leslie, 2009). To date, there is little agreement in the literature as to the validity and unbiasedness of such data. The results in the paper show that such data can be used successfully to understand how individuals make choices. However, this paper is clearly not the last word on this subject and more studies of this nature need to be conducted, especially because the data used in this study come from a very stylized setting. Moreover, the particular setting is one where, at the time of the first survey, the student's decision of what major to choose was reversible at a low cost. One would expect the impact of biases on beliefs to be stronger in settings where the decision is irreversible (or more costly to reverse).

References

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

Bellemare, Charles, Sabine Kroger, and Arthur van Soest. 2008. "Measuring Inequality Aversion in Heterogeneous Population Using Experimental Decisions and Subjective Probabilities." *Econometrica*, 76(4): 815-839.

Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, Sofia Cheidvasser, and John Rust. 2004. "How Large is the Bias in Self-Reported Disability?" *Journal of Applied Econometrics*, 19: 649-670.

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

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Do People Mean What They Say? Implications for Subjective Survey Data", *American Economic Review Papers & Proceedings*, 91(2): 67-72.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." Chapter in *Handbook of Econometrics*, V. 5, eds. Edward Leamer and James Heckman, 3705-3843. Elsevier Science.

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.

Buchinsky, Moshe, and Phillip Leslie. 2009. "Educational Attainment and the Changing U.S. Wage Structure: Dynamic Implications on Young Individuals' Choices." Working Paper.

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

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

Delavande, Adeline, Xavier Gine, and David McKenzie. 2010. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics*, Forthcoming.

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

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

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

Hyslop, Dean, and Guido Imbens. 2001. "Bias From Classical and Other Forms of Measurement Error." *Journal of Business and Economics Statistics*, 19(4): 475-481.

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

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. 2002. "Identification of Decision Rules in Experiments on Simple Games of Proposal and Response." *European Economic Review*, 46(4-5): 880-891.

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

Manski, Charles, and Francesca Molinari. 2010. "Rounding Probabilistic Expectations in Surveys." *Journal of Business and Economic Statistics*, 28(2): 219-231.

Manski, Charles, and Elie Tamer. 2002. "Inference on Regressions with Interval Data on a Regressor or Outcome." *Econometrica*, 70(2): 519-546.

Mullainathan, Sendhil and Ebonya Washington. 2008. "Sticking with your Vote: Cognitive Dissonance and Political Attitudes." *American Economic Journal: Applied Economics*, Forthcoming.

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.

Offerman, Theo, Joep Sonnemans, and Arthur Schram. 1996. "Value Orientations, Expectations and Voluntary Contributions in Public Goods." *The Economic Journal*, 106(437): 817-845.

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

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

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

6 Appendix

6.1 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?

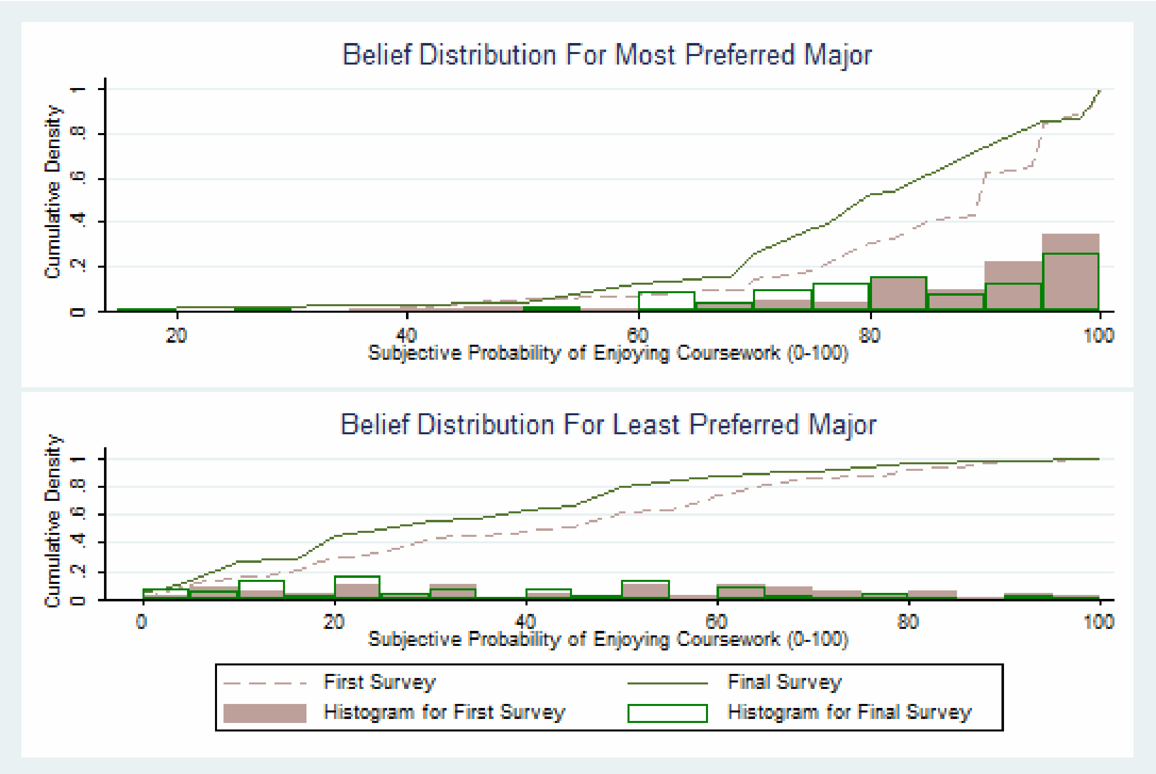

Figure 1: Belief Distribution of Enjoying Coursework

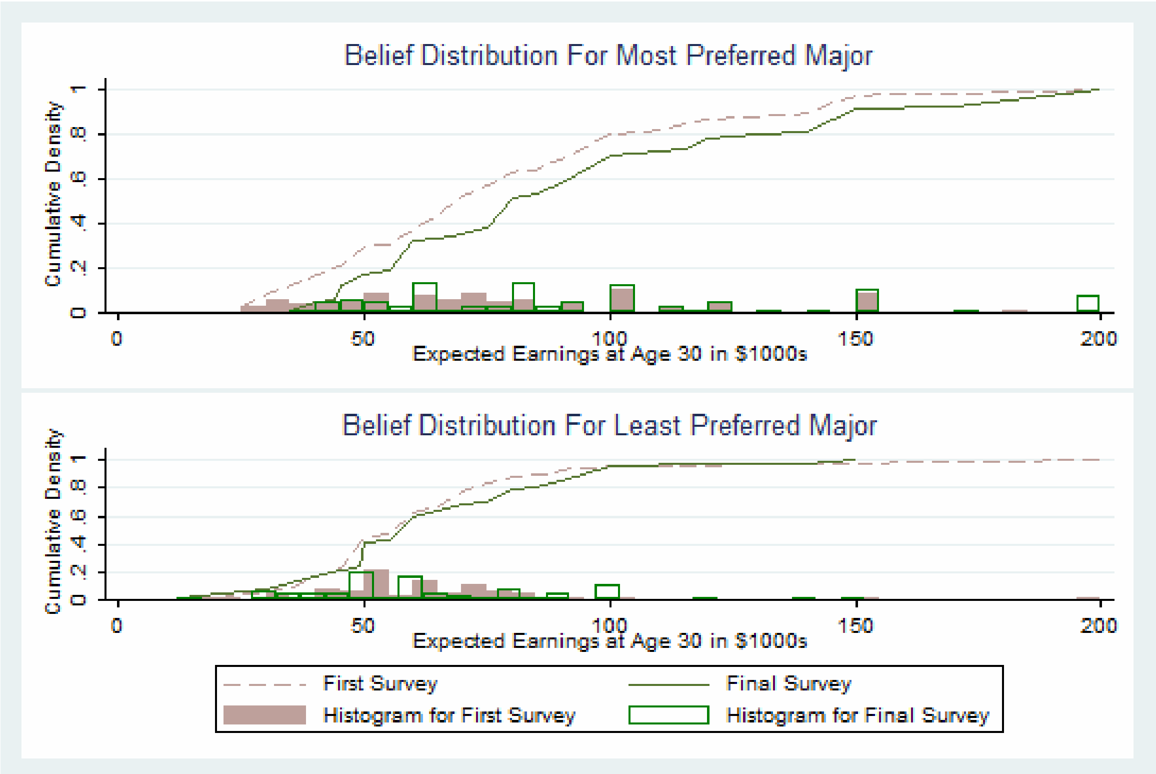

Figure 2: Belief Distribution of Expected Income at Age 30

Table 1: 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[g] (47)
No	56 (48)	71 (44)	534 (53)
Second Major?[e]			
Yes	55 (47)	78 (48.5)	–
No	62 (53)	83 (51.5)	–
Average GPA*			
Male	3.51	3.48	3.26
Female	3.43	3.40	3.31
WCAS Majors[f]			
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)	–
Lit & Fine Arts	16 (13.5)	20 (12.5)	–

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 Major distribution of students. In cases where the survey respondent has more than one major in WCAS, only the first one is included. Majors that appear in each category are listed in Table A1.
g 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 2: Summary Statistics: Mean and Standard Deviation of Elicited Beliefs

		Graduate in 4 years	GPA $\geq$ 3.5	Enjoy Courses	Courses hrs/wk	Parents' Approval	Find Job	Enjoy Work	Reconcile Family	Job hrs/wk	Status of Jobs	Income At 30
Natural Sciences	In[a] (25)	0.94**	0.70***	0.79***	27.32	0.93**	0.79	0.87***	0.67*	50.88	0.84**	111.40
		(0.11)	(0.28)	(0.20)	(13.22)	(0.10)	(0.19)	(0.09)	(0.24)	(14.63)	(0.09)	(67.79)
	Out (92)	0.83	0.52	0.52	28.34	0.86	0.72	0.60	0.58	50.74	0.75	91.13
		(0.21)	(0.27)	(0.24)	(11.86)	(0.16)	(0.21)	(0.23)	(0.21)	(11.87)	(0.18)	(86.36)
Math & Comp Sci	In (6)	0.95	0.84**	0.82***	21.33	0.93*	0.82	0.76**	0.84**	41.67	0.60	58.67
		(0.04)	(0.13)	(0.16)	(14.99)	(0.09)	(0.14)	(0.19)	(0.05)	(5.01)	(0.09)	(25.97)
	Out (111)	0.83	0.56	0.48	27.05	0.75	0.72	0.53	0.67	44.19	0.59	74.41
		(0.21)	(0.27)	(0.25)	(12.00)	(0.22)	(0.19)	(0.22)	(0.20)	(10.00)	(0.18)	(52.45)
Social Sciences I	In (44)	0.94	0.79	0.87***	20.98	0.79*	0.65	0.79**	0.71	47.73***	0.58	76.25
		(0.07)	(0.16)	(0.11)	(10.24)	(0.19)	(0.20)	(0.14)	(0.18)	(11.67)	(0.15)	(37.73)
	Out (73)	0.94	0.81	0.76	22.63	0.72	0.61	0.70	0.76	42.11	0.55	64.99
		(0.10)	(0.16)	(0.19)	(10.35)	(0.25)	(0.20)	(0.20)	(0.15)	(7.53)	(0.17)	(38.22)
Social Sciences II	In (42)	0.93***	0.73***	0.77***	25.69	0.88***	0.85**	0.73***	0.66	54.10**	0.73***	124.60*
		(0.10)	(0.20)	(0.18)	(14.23)	(0.12)	(0.12)	(0.17)	(0.15)	(11.05)	(0.16)	(81.64)
	Out (75)	0.85	0.59	0.57	27.40	0.80	0.78	0.54	0.60	49.47	0.59	97.40
		(0.14)	(0.26)	(0.23)	(11.35)	(0.17)	(0.15)	(0.23)	(0.22)	(11.88)	(0.17)	(82.10)
Ethics and Values	In (2)	0.98	0.80	0.88	21.00	0.68	0.68	0.60	0.80	37.50	0.35	47.50
		(0.04)	(0.28)	(0.18)	(4.24)	(0.46)	(0.46)	(0.57)	(0.28)	(3.54)	(0.07)	(3.54)
	Out (115)	0.90	0.77	0.70	23.21	0.64	0.55	0.63	0.69	44.95	0.42	70.10
		(0.12)	(0.17)	(0.18)	(11.94)	(0.26)	(0.20)	(0.19)	(0.20)	(11.32)	(0.19)	(42.52)
Area Studies	In (18)	0.95*	0.86*	0.86***	20.89	0.76**	0.64	0.81***	0.65	43.17	0.46***	68.39
		(0.04)	(0.10)	(0.17)	(6.48)	(0.26)	(0.20)	(0.15)	(0.18)	(6.80)	(0.23)	(36.08)
	Out (99)	0.90	0.78	0.70	22.83	0.60	0.56	0.61	0.71	43.17	0.34	57.60
		(0.11)	(0.19)	(0.20)	(11.00)	(0.26)	(0.20)	(0.19)	(0.15)	(9.81)	(0.15)	(24.01)
Lit & Fine Arts	In (20)	0.97**	0.84*	0.95***	26.65	0.86***	0.64**	0.83***	0.87***	40.00	0.48***	54.75
		(0.05)	(0.22)	(0.05)	(11.82)	(0.12)	(0.24)	(0.17)	(0.11)	(10.18)	(0.27)	(29.60)
	Out (97)	0.90	0.77	0.65	23.43	0.51	0.51	0.60	0.74	41.31	0.32	54.59
		(0.12)	(0.15)	(0.23)	(11.47)	(0.28)	(0.22)	(0.21)	(0.16)	(9.60)	(0.13)	(25.85)
Engineering	In (4)	0.97	0.63	0.79***	25.75	0.93	0.89	0.85***	0.79*	43.75	0.87	76.25
		(0.05)	(0.34)	(0.09)	(14.10)	(0.08)	(0.10)	(0.12)	(0.03)	(4.79)	(0.06)	(7.50)
	Out (106)	0.82	0.48	0.41	27.42	0.84	0.83	0.56	0.63	48.53	0.71	94.57
		(0.21)	(0.26)	(0.26)	(11.62)	(0.17)	(0.15)	(0.21)	(0.18)	(10.45)	(0.16)	(65.41)

(.) Standard deviations in parentheses

*** Difference in means between "In" and "Out" sample is significant at the 1% level (2-tailed t-test); ** significant at 5%; * significant at 10%.

Binary outcomes (all outcomes except coursework hrs/wk, job hrs/wk, social status of jobs; income at 30) are on a 0-100 scale (and then divided by 100);

Coursework hrs/wk and job hrs/wk elicited on a scale of 0-70; Social status elicited on a 0-0.9 scale; Income at age 30 is expressed in 1000s.

[a]"In" is the subsample that has a major in that category; "Out" is the subsample that does not have a major in that category. Number of observations in brackets.

All fields other than Engineering have 117 responses. 7 respondents didn't answer beliefs about outcomes associated with Engineering because of a survey glitch.

Table 3: Summary Statistics about Subjective Beliefs about Binary Outcomes

	Dropped Major[a]	**Least Pref. Major**[b]	**Next Pref. Major**[c]	**Sec Major**[d]	**Current Major**
Graduate in 4 years	-8.5	-0.30	-0.97	4.43	1.71
	(12.5)	(15.85)	(9.65)	(6.47)	(4.75)
	[7.69%]	[37.61%]	[28.20%]	[32.48%]	[75.21%]
	{1.7%}	{7.7%}	{7.7%}	{6.0%}	{17.1%}
	{0%}	*{0.85%}*	*{0.85%}*	*{0%}*	*{0%}*
GPA of ≥ 3.5	2.57	-6.66	-5.24	-1.60	-5.32
	(13.71)	(16.09)	(16.37)	(12.26)	(14.99)
	[5.98%]	[29.91%]	[11.96%]	[23.08%]	[33.33%]
	{0.9%}	{5.1%}	{0.9%}	{4.3%}	{4.3%}
	{0.85%}	*{1.7%}*	*{0.85%}*	*{0%}*	*{1.7%}*
Enjoy Coursework	-10.21	-9.15	-2.91	-0.26	-4.09
	(14.35)	(18.97)	(16.43)	(14.74)	(11.89)
	[3.41%]	[23.93%]	[15.38%]	[18.80%]	[35.04%]
	{0%}	{5.1%}	{1.7%}	{2.6%}	{9.4%}
	{0%}	*{2.6%}*	*{0%}*	*{1.7%}*	*{3.4%}*
Approval of Parents	-2.64	-0.26	-4.48	-0.45	0.42
	(8.21)	(14.37)	(15.28)	(15.79)	(10.13)
	[6.84%]	[35.04%]	[16.24%]	[18.80%]	[48.71%]
	{1.7%}	{1.7%}	{2.6%}	{5.1%}	{8.5%}
	{1.7%}	*{0%}*	*{0.85%}*	*{0.85%}*	*{2.6%}*
Finding a job	2.29	-2.39	-2.31	-1.41	-1.19
	(18.00)	(18.33)	(16.17)	(15.00)	(16.16)
	[5.13%]	[29.91%]	[21.37%]	[17.95%]	[31.62%]
	{0.9%}	{3.4%}	{0.9%}	{3.4%}	{5.1%}
	{0.85%}	*{2.6%}*	*{0.85%}*	*{2.6%}*	*{2.6%}*
Enjoying work at jobs	-7.78	-12.66	-3.97	-5.98	-4.51
	(23.35)	(21.38)	(17.59)	(16.22)	(13.19)
	[2.56%]	[17.95%]	[16.24%]	[12.82%]	[39.31%]
	{0.9%}	{2.6%}	{0.9%}	{1.7%}	{3.4%}
	{0.85%}	*{2.6%}*	*{0%}*	*{1.7%}*	*{2.6%}*
Reconcile work & family	2.21	0.78	5.07	3.26	2.26
	(20.07)	(18.33)	(16.38)	(12.36)	(14.68)
	[2.56%]	[17.94%]	[19.66%]	[20.51%]	[35.04%]
	{1.7%}	{0.9%}	{2.6%}	{0.9%}	{4.3%}
	{0%}	*{0.85%}*	*{1.7%}*	*{0.85%}*	*{2.6%}*
No. of Observations	14	103	58	59	117

(.) mean absolute change in belief between the two surveys

[.] proportion of respondents for whom change in beliefs is $\leqslant 5$ for binary outcomes

{.} Proportion of responses in the follow-up survey that are not a multiple of 5

{.} Proportion of responses in the follow-up that are not a multiple of 5 EXCLUDING extremities (≤ 5; $\geq$ 95)

a A major that the student was pursuing when first surveyed, but dropped at the time of the second survey

b An individual's least preferred major at the time of the second survey

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

d The individual's second major

Table 4: Summary Statistics about Expectations for Continuous Outcomes

	Dropped Major[a]	**Least Pref. Major**[b]	**Next Pref. Major**[c]	**Sec Major**[d]	**Current Major**
Coursework hrs/week	-8.21	0.26	-3.02	-4.29	-5.49
	(12.07)	(9.32)	(8.63)	(10.50)	(9.95)
	[4.27%]	[47.00%]	[24.79%]	[23.93%]	[41.02%]
	{2.6%}	{12.8%}	{6.0%}	{16.2%}	{22.2%}
Job hrs/week	0	2.69	1.72	1.84	2.26
	(7.14)	(7.81)	(10.93)	(7.71)	(8.34)
	[9.41%]	[50.42%]	[26.49%]	[29.06%]	[47.86%]
	{0%}	{0%}	{0%}	{0.9%}	{2.7%}
Salary at the age of 30	-4928.57	-4757.99	-100	22878.97	12856.38
	(56071.43)	(26718.8)	(21927.6)	(37460.7)	(40097.75)
	[1.71%]	[17.09%]	[16.23%]	[16.24%]	[16.24%]
	{0%}	{0%}	{0%}	{0%}	{0%}
No. of Observations	14	103	58	59	117

(.) mean absolute change in belief between the two surveys

[.] proportion of respondents for whom change in beliefs is $\leqslant$ 5 for hrs/week; $\leqslant$5000 for inc

{.} Proportion of responses in the follow-up survey that are not a multiple of 5

a A major that the student was pursuing when first surveyed, but dropped at the time of the second survey

b An individual's least preferred major at the time of the second survey

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

d The individual's second major

Table 5: 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**
Panel A: All (117 individuals; 341 observations)										
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)
Panel B: Declared Major (61 individuals; 179 observations)										
Constant	2.14	-4.83*	-3.51	-3.83**	-0.54	-2.77	-4.78*	2.67	-0.82	8377.96
	(1.23)	(2.77)	(2.42)	(1.73)	(2.60)	(3.07)	(2.78)	(2.54)	(1.38)	(6776.89)
Second Pursued Major	3.74	1.39	5.85*	1.28	0.63	0.25	-1.15	0.35	3.22	304.52
	(2.70)	(4.05)	(3.54)	(2.11)	(3.86)	(4.24)	(4.37)	(3.86)	(2.13)	(9215.49)
Second Preferred Major	-3.19	2.99	0.54	2.35	-0.65	0.81	0.62	0.04	-1.24	-17325.46*
	(3.04)	(4.59)	(4.02)	(2.41)	(4.38)	(4.83)	(4.93)	(4.37)	(2.40)	(10510.08)
Dropped Major	-16.22***	-0.55	-8.14	2.46	6.71	8.23	9.58	11.22	6.69	27982.38
	(6.08)	(9.36)	(8.19)	(5.05)	(8.89)	(9.95)	(9.81)	(8.80)	(4.82)	(21700.45)
Least Preferred Major	0.74	0.00	-4.09	2.69	-0.48	-2.94	-3.61	-1.93	0.70	-18031.65**
	(2.34)	(3.48)	(3.05)	(1.78)	(3.33)	(3.62)	(3.80)	(3.34)	(1.84)	(7862.48)
Panel C: Not Declared Major (56 individuals; 162 observations)										
Constant	0.77	-5.84**	-4.72	-7.53***	1.50	1.16	-4.31	1.23	5.43**	20749.42*
	(3.04)	(2.80)	(3.19)	(1.91)	(2.85)	(3.37)	(3.07)	(3.44)	(2.23)	(10772.76)
Second Pursued Major	0.86	5.39	0.49	0.94	-1.19	-0.74	-1.58	3.51	-4.16	19881.47
	(3.83)	(4.65)	(5.59)	(2.20)	(4.71)	(4.84)	(5.13)	(5.46)	(3.24)	(17684.27)
Second Preferred Major	-1.18	-0.75	2.13	2.95	-9.53**	-4.12	0.38	4.73	-1.19	-12461.85
	(3.34)	(4.10)	(4.97)	(1.92)	(4.15)	(4.23)	(4.52)	(4.80)	(2.84)	(15580.99)
Dropped Major	-3.10	11.11	-6.21	-1.64	-7.89	-4.60	-8.04	-2.07	-6.32	-49732.6*
	(5.79)	(6.84)	(8.09)	(3.34)	(6.94)	(7.25)	(7.53)	(8.10)	(4.85)	(26073.68)
Least Preferred Major	-5.45*	-2.76	-6.05	8.55***	-1.13	0.84	-13.73***	-0.94	0.06	-20242.00
	(2.93)	(3.67)	(4.48)	(1.68)	(3.71)	(3.74)	(4.05)	(4.27)	(2.51)	(13904.87)

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 6: Are individuals aware of changes in beliefs?

	Least Pref Major			Second Major			Current Major		
Perceived Change: Actual Change	Inc.	Unchg	Dec.	Inc.	Unchg	Dec.	Inc.	Unchg	Dec.
				Graduate in 4 years					
Increase	3*	22	5	8	23	0	6	16	0
Unchanged	5	46	1	4	66	1	10	74	4
Decrease	2	20	9	0	9	6	0	2	1
				Graduate with a GPA $\geq$ 3.5					
Increase	5	17	7	9	14	6	11	8	3
Unchanged	4	28	8	8	25	8	9	21	9
Decrease	1	29	15	4	29	14	9	20	23
				Enjoy Coursework					
Increase	5	13	8	11	17	4	10	11	5
Unchanged	0	21	8	5	30	4	16	25	0
Decrease	8	38	13	7	20	18	11	25	10
				Coursework hrs/week					
Increase	9	12	3	5	7	1	4	9	1
Unchanged	10	44	4	14	36	7	13	30	5
Decrease	7	20	5	11	27	8	20	21	10
				Approval of Parents					
Increase	5	32	0	5	27	2	5	29	0
Unchanged	3	39	4	7	29	5	2	51	4
Decrease	4	20	7	5	34	3	3	15	4
				Finding a job					
Increase	11	21	1	6	19	6	9	20	6
Unchanged	3	32	4	5	29	12	6	23	8
Decrease	2	31	9	4	25	11	7	21	13
				Enjoying work at jobs					
Increase	3	12	9	9	21	3	12	8	5
Unchanged	1	19	4	7	22	5	15	26	5
Decrease	7	46	13	2	35	13	6	30	6
				Reconciling work & family					
Increase	10	37	4	6	28	5	9	24	10
Unchanged	3	18	3	8	31	8	10	25	6
Decrease	2	33	4	4	23	4	4	12	13
				Job hrs/week					
Increase	9	23	1	11	17	2	21	14	1
Unchanged	10	57	1	7	54	4	16	35	5
Decrease	1	8	4	3	13	5	6	14	1

*Each cell shows the number of respondents who fall in that category. Recall there are a total of 117 respondents.

Table 7: Comparing GPA and Income Beliefs with Objective Measures

	Average GPA[a]		Belief GPA ≥ 3.5[b]				Average Salary[c]		Belief of Income at 30[d]			
	Mean	**Rank**[e]	**Mean**	**Rank**	**In**[f]	**Out**[g]	**Mean**	**Rank**	**Mean**	**Rank**	**In**	**Out**
	(1a)	(1b)	(2a)	(2b)	(2c)	(2d)	(3a)	(3b)	(4a)	(4b)	(4c)	(4d)
Natural Sciences	3.22	5	0.56	7	0.70***	0.52	75.86	3	95.46	2	111.40	91.13
			(0.28)		(0.28)	(0.27)			(82.89)		(67.79)	(86.36)
Math & Comp Sci	3.21	6	0.58	6	0.84**	0.56	73.32	4	73.6	4	58.67	74.41
			(0.27)		(0.13)	(0.27)			(51.48)		(25.97)	(52.45)
Social Sciences I	3.29	1	0.8	1	0.79	0.81	72.73	5	69.22	6	76.25	64.99
			(0.16)		(0.16)	(0.16)			(38.27)		(37.73)	(38.22)
Social Sciences II	3.09	8	0.64	5	0.73***	0.59	78.1	2	107.16	1	124.60*	97.40
			(0.25)		(0.20)	(0.26)			(82.63)		(81.64)	(82.10)
Ethics and Values	3.29	1	0.77	4	0.80	0.77	68.23	6	69.71	5	47.50	70.10
			(0.17)		(0.28)	(0.17)			(42.25)		(3.54)	(42.52)
Area Studies	3.29	1	0.79	2	0.86*	0.78	68.23	6	59.25	7	68.39	57.60
			(0.18)		(0.10)	(0.19)			(26.32)		(36.08)	(24.01)
Lit & Fine Arts	3.29	1	0.78	3	0.84*	0.77	62.87	8	54.62	8	54.75	54.59
			(0.16)		(0.22)	(0.15)			(26.39)		(29.60)	(25.85)
Engineering	3.11	7	0.48	8	0.63	0.48	89.26	1	93.9	3	76.25	94.57
			(0.27)		(0.34)	(0.26)			(64.30)		(7.50)	(65.41)

[a] Mean GPA of bachelor graduates of Doctoral/Research Universities in 2001 (Source: 2001 Baccalaureate and Beyond Longitudinal Study).
[b] Belief of survey respondents about graduating with a GPA ≥ 3.5 (on a scale of 0-100) and divided by 100. (N= 117 except for Engineering which has 110 observations).
[c] Average salary in 1000s (in 2007 dollars) in 2003 of college graduates of 1993. Restricted to selective Doctoral/Research Universities with Carnegie Code 4 (Source: 1993/03 Baccalaureate and Beyond Longitudinal Study).
[d] Expected salary at the age of 30 (in 1000s) elicited from survey respondents.
[e] Majors are ranked from highest GPA or expected salary (rank 1) to lowest (rank 8).
[f] Sample restricted to students with a major in that category.
[g] Sample restricted to students with NO major in that category.
*** Difference in means is significant at the 1% level (2-tailed t-test); ** significant at 5%; * significant at 10%.

Table A1: Categorization of Majors

Category Name	**Majors in Category**
WCAS Majors[+]	
a Natural Sciences	Biological Sciences; Chemistry; Environmental Sciences; Geography*; Geological Sciences; Integrated Science; Materials Science; Physics
b Math & 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 the Social Sciences
e Ethics and Values	Legal Studies*; Philosophy; Religion; Science in Human Culture*
f Area Studies	African American Studies; American Studies; Asian and Middle East Languages & Civilization; European Studies; International Studies*; Slavic Languages and Literatures
g Literature & Fine Arts	Art History; Art Theory and Practice; Classics; Comparative Literary Studies; Drama; English; French; German; Italian; Spanish
Non-WCAS Majors	
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 & Design Engineering; Materials Science & Engineering; Mechanical Engineering
l Journalism[5]	Journalism

*Adjunct majors (these do not stand alone)

[+] Majors in the Weinberg College of Arts & Sciences (WCAS)

[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