Federal Reserve Bank of New York
Staff Reports

Double Majors: One for Me, One for the Parents?

Basit Zafar

Staff Report no. 478
November 2010

Double Majors: One for Me, One for the Parents?
Basit Zafar
Federal Reserve Bank of New York Staff Reports, no. 478
November 2010
JEL classification: D8, I2, J1

Abstract

At least a quarter of college students in the United States graduate with more than one undergraduate major. This paper investigates how students decide on the composition of their paired majors—in other words, whether the majors chosen are substitutes or complements. Since students use both their preferences and their expectations about major-specific outcomes when choosing their majors, I collect innovative data on subjective expectations, drawn from a sample of Northwestern University sophomores. Despite showing substantial heterogeneity in beliefs, the students seem aware of differences across majors and have sensible beliefs about the outcomes. Students believe that their parents are more likely to approve majors associated with high social status and high returns in the labor market. I incorporate the subjective data in a choice model of double majors that also captures the notion of specialization. I find that enjoying the coursework and gaining approval of parents are the most important determinants in the choice of majors. The model estimates reject the hypothesis that students major in one field to pursue their own interests and in another for parents' approval. Instead, I find that gaining parents' approval and enjoying a field of study both academically and professionally are outcomes that students feel are important for both majors. However, I do find that students act strategically in their choice of majors by choosing ones that differ in their chances of completion and difficulty and in finding a job upon graduation.

Key words: college majors, uncertainty, subjective expectations, preferences

---

Zafar: Federal Reserve Bank of New York (e-mail: basit.zafar@ny.frb.org).

1 Introduction

At least a quarter of college students have more than one undergraduate major (2003 National Survey of College Graduates), and the share of college students choosing more than one major is increasing at a very fast rate (Lewin, 2002). It has been postulated that parents inadvertently drive their children to have more than one major, i.e., students major in one field to satisfy their own interests, and in another field that meets parents' approval.[1] Other explanations for double majors include students majoring in one field associated with their professional specialty and another that reflects a very different interest (like an Engineering major who's also majoring in French), or students hedging their chances in the labor market by preparing to work in more than field. However, evidence for all these explanations remains anecdotal (Lewin, 2002; Gomstyn, 2003), and there is little systematic evidence on how students choose the composition of their double majors. This paper provides, to the best of my knowledge, the first direct evidence on how students, conditional on having a double major, choose the composition of majors.[2]

Students choose a college major (or pair of majors) in order to influence the occurrence of choice-specific outcomes that enter their utility function. These outcomes include, for example, being able to successfully complete a field of study, gaining parents' approval, finding a job upon graduation, enjoying coursework or earnings at the job. Since these outcomes are uncertain at the time the student makes his choice, he has a belief distribution of the probability for the occurrence of these outcomes conditional on each major in his choice set. Therefore, a student uses both his preferences and subjective beliefs in choosing his college major(s). The researcher usually only observes the major(s) that the student chooses, and has to make non-verifiable assumptions on expectations/ beliefs to infer the parameters of the utility function (preferences). The basic difficulty is that observed choices may be consistent with several combinations of expectations and preferences, and the list of underlying assumptions on expectations may not be valid (see Manski, 1993, for a discussion of this inference problem in the context of how

[1] For example, Lewin (2002) remarks, "Occasionally, combinations represent a compromise: where the mother is pushing for law school, for example, and the son wants to pursue ethnomusicology, a "one for me, one for Mom" double major in political science and music can keep the whole family happy."

[2] There is a literature on college majors which primarily focuses on the choice of a *single* field of study: Altonji (1993), Arcidiacono (2004), Zafar (2009), and Arcidiacono, Hotz, and Kang (2010).

students infer returns to schooling). A solution to this identification problem is to use additional data on expectations (Manski, 2004), and that is precisely what I do. I survey a group of 69 Northwestern University students pursuing double majors and elicit their subjective beliefs about major-specific outcomes.

In my relatively homogenous sample, there is substantial heterogeneity in beliefs for outcomes within a major as well as across majors, indicative that there exists tremendous heterogeneity in beliefs in the population of college students. Analysis of beliefs for the same outcome across majors suggests that students have *sensible* beliefs about the occurrence of outcomes conditional on major. For example, the belief distribution of reconciling work and family at jobs available in Literature and Fine Arts first order stochastically dominates the corresponding distribution in Natural Sciences (in which most pre-med students major). Comparison of beliefs of being able to graduate with a GPA of at least 3.5 and expected income at the age of 30 with objective measures reveals that students are aware of differences across majors. I find that students believe that their parents are more likely to approve majors associated with high social status and returns in the labor market. For example, the mean belief of gaining parents' approval for majoring in Engineering is 0.87 (on a scale of 0-1) compared with 0.59 for Literature and Fine Arts.

The subjective data elicited from the students is employed directly in a structural model of double major choice.[3] Though I assume that students are forward-looking and care about future outcomes when making their choices, I do not have data needed to estimate a dynamic model. I instead assume that individuals maximize *current* expected utility, and estimate a static choice model. The heterogeneity in beliefs allows me to identify the preferences for each outcome considered in the model. Since students may choose more than one major to either expand the set of options they have or hedge along a certain outcome (i.e., they may choose majors that differ in the likelihood of that outcome), the model specification captures both these motivations. I find that enjoying the coursework and gaining approval of parents are the most

[3]This approach adds to the recent literature which employs expectations data in econometric models to conduct inference on behavior: Lochner (2007); Bellemare, Kroger, and van Soest (2008); Delavande (2008a); Zafar (2009); and Arcidiacono et al (2010). van der Klaauw (2000) and van der Klaauw and Wolpin (2008) employ expectations data to improve the precision of estimates in their structural dynamic models while maintaining the assumption of rational expectations to identify the model.

important determinants in the choice of double majors in my sample.[4] However, contrary to existing anecdotal evidence (Lewin, 2002; Gomstyn, 2003), I do not find that students major in one field to gain the approval of their parents and in another to satisfy their interests. Instead, gaining parents' approval and enjoying studying and working in a field of study are outcomes that are important for *both* majors in an individual's major pair. I do, however, find that students act strategically by choosing majors that differ in their chances of completion, in their level of difficulty, and in finding a job upon graduation. So, students basically pair an easy major with a hard one. This pattern of specialization is also consistent with anecdotal evidence that students choose double majors to hedge their prospects in the labor market.

On the whole, the results in the paper suggest that students with double majors pursue their interests at college while taking into account parents' approval, and they also act strategically in their choices by choosing majors that differ in their chances of completion and in finding a job upon graduation. It should be pointed out that this paper investigates how students choose the composition of their majors *conditional* on deciding to pursue a double major. I do not attempt to explain why some students may choose to pursue more than one major. However, I present evidence that double major respondents are similar to single major respondents: Their subjective belief distributions for most outcomes are similar to the corresponding distributions of single major respondents, and the two groups of students have similar preferences for the various outcomes. For example, gaining parents' approval is an equally important determinant in the choice of both single and double major respondents. I do find that, compared to single major respondents, the double major respondents in my sample arrive in college with more AP credits (suggesting that they need to satisfy fewer requirements for completing a major) and have higher GPAs at the time of the survey (indicative of selection along ability). The limited available data prevent me from answering the question of what drives certain students to choose more than one major. This paper also does not have anything to say about the costs and benefits to double majors. There is concern that studying more than one major in college may result in too little depth within one's main field of study and a decrease in the breadth

[4] This finding is similar to Zafar (2009) who estimates students' preferences for college majors and restricts the analysis to single majors only. He finds that enjoying coursework and gaining parents' approval are the most important determinants of (single) major choice.

of general knowledge. On the other hand, students with double majors have been found to have higher earnings (Del Rossi and Hersch, 2008). In the absence of data on student outcomes and addressing the issue of selection into double majors, it is not possible to evaluate the net benefits of pursuing more than one major.

This paper is organized as follows. Section 2 outlines the choice model and the identification strategy. Section 3 describes the data collection methodology, the institutional setup at Northwestern University, and the subjective data in detail. Section 4 presents the estimation results for the choice model and some robustness checks. Finally, Section 5 concludes.

2 Choice Model

Student i derives utility $\mathbf{U}_{ik}(\mathbf{a}, \mathbf{c})$ from choosing a major k (if the student chooses a dual major, k denotes a pair of majors consisting of majors k_1 and k_2). Students are assumed to be forward-looking, so their choice of major(s) depends not only on the current state of the world but also on what they expect will happen in the future. Utility is a function of a vector of major-specific outcomes $\mathbf{a}$ that are realized in college and a vector of outcomes $\mathbf{c}$ that are realized after graduating from college. The vector $\mathbf{a}$ includes the outcomes:

a_1 successfully complete (graduate in) a field of study in four years

a_2 graduate with a GPA of at least 3.5 in the field of study

a_3 enjoy the coursework

a_4 hours per week spent on the coursework

a_5 gain parents' approval of the major

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

c_1 get an acceptable job immediately upon graduation

c_2 enjoy working at the jobs available after graduation

c_3 able to reconcile work and family at the available jobs

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

The outcomes $\{a_r\}_{r=\{1,2,3,5\}}$ and $\{c_q\}_{q=\{1,2,3\}}$ are binary (for example, in the case of a_1, a student either graduates in four years or not), while outcomes a_4 and $\{c_q\}_{q=\{4,5,6\}}$ are continuous. I change the notation slightly and define $\mathbf{b}$ to be a 7×1 vector of all binary outcomes, i.e., $\mathbf{b} = \{a_1, a_2, a_3, a_5, c_1, c_2, c_3\}$, and $\mathbf{d}$ to be a 4×1 vector of all continuous outcomes, i.e., $\mathbf{d} = \{a_4, c_4, c_5, c_6\}$. The vectors $\mathbf{b}$ and $\mathbf{d}$ are uncertain at the time of the choice, and individual i possesses subjective beliefs $P_{ik}(\mathbf{b}, \mathbf{d})$ about the outcomes associated with major k for all $k \in S_i$, where S_i is i's choice set.[5]

Before specifying the structural form of the utility function describing choice of majors for double major students, it is useful to outline the objective function of a student with a single major. Students are assumed to maximize their *current* expected utility.[6] If an individual chooses a major m, then a standard revealed preference argument (assuming that indifference between alternatives occurs with zero probability) implies that:

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

The goal is to infer the preference parameters from observed choices. However, the expectations of the individual about the choice-specific outcomes, $P_{ik}(\mathbf{b}, \mathbf{d})$, are usually not directly observable. The standard approach in the literature is to 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 form their expectations in the same way.[7] First, there is little reason to think that individuals form their expectations in the same way.[8] Second, different combinations of preferences and expectations may lead to the same choice (Manski, 2002). To cope with

[5] 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.

[6] Under the assumption that individuals maximize current expected utility, I don't need to take into account that individuals may find it optimal to experiment with different majors (i.e., they may switch major combinations or simply switch from pursuing a double major to a single major). However, experimentation could be important in this context as students may learn about ability and match quality (Malamud, 2006; Stinebrickner and Stinebrickner, 2008; Zafar, 2010a). It is beyond the scope of this paper.

[7] Consider, for example, income expectations conditional on schooling choices. Freeman (1971) assumes that students have myopic expectations, Willis and Rosen (1979) hypothesize that expectations are rational, while Arcidiacono (2004) assumes that students condition their expectations on ability, GPA, average ability of other students enrolled in the college and some demographic controls. It's not clear which of these rules is the correct one.

[8] In fact Delavande (2008b) finds heterogeneity in the way women revise their expectations about effectiveness of contraception methods with receipt of the same information, and Zafar (2009) finds heterogeneity in how students update their beliefs about ability in response to new information.

the problem of joint inference on preferences and expectations, I elicit subjective probabilities directly from individuals. An additional advantage of this approach is that it allows me to explicitly account for the various non-pecuniary determinants of the choice.[9] Since it would be difficult to elicit the joint probability distribution $P_{ik}(\mathbf{b}, \mathbf{d})$, I assume that utility is linear and separable in outcomes, so that:

$$\mathbf{U}_{ik}(\mathbf{b}, \mathbf{d}) = \alpha_k + \sum_{r=1}^{7} u_r(b_r) + \sum_{q=1}^{4} \gamma_q d_q + \varepsilon_{ik},$$

where α_k is a major-specific constant, $u_r(b_r)$ is the utility associated with the binary outcome b_r, γ_q is a constant for the continuous outcome d_q, and ε_{ik} is a random term. Equation (1) can now be written as:

$$m \equiv \arg\max_{k \in S_i}(\alpha_k + \sum_{r=1}^{7} \int u_r(b_r) dP_{ik}(b_r) + \sum_{q=1}^{4} \gamma_q \int d_q dP_{ik}(d_q) + \varepsilon_{ik}).$$

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 ($\{b_r\}_{r=1}^{7}$):

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

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 $\int \mathbf{U}_i(\mathbf{b}, \mathbf{d}) dP_{ik}(\mathbf{b}, \mathbf{d}) = \mathbf{U}_i(\int \mathbf{b}, \mathbf{d}\ dP_{ikt}(\mathbf{b}, \mathbf{d}))$. Thus, for the continuous outcomes ($\{d_q\}_{q=1}^{4}$), $\int d_q dP_{ik}(d_q)$ equals $E_{ik}(d_q)$, the expected value of the outcome. The *expected* utility that individual i derives from choosing major m is:

$$\begin{aligned} &U_{im}(\mathbf{b}, \mathbf{d}, \{P_{im}(b_r = 1)\}_{r=1}^{7}, \{E_{im}(d_q)\}_{q=1}^{4}) = \\ &\alpha_m + \textstyle\sum_{r=1}^{7} P_{im}(b_r = 1) \triangle u_r + \sum_{r=1}^{7} u_r(b_r = 0) + \sum_{q=1}^{4} \gamma_q E_{im}(d_q) + \varepsilon_{im}. \end{aligned} \tag{2}$$

In equation (2), α_m, $\{\triangle u_r\}_{r=1}^{7}$, and $\{\gamma_q\}_{q=1}^{4}$ are the parameters of the utility function that

[9] In the absence of data on non-pecuniary outcomes, existing studies are constrained to infer the importance of non-pecuniary and psychic factors from residuals that help explain the choices after the model is imposed (see, for example, Cunha, Heckman, and Navarro, 2005). The approach used in this paper allows me to label the various components of these unspecified psychic and non-pecuniary factors.

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. $\{P_{ik}(b_r = 1)\}_{r=1}^{7}$ and $\{E_{ik}(d_q)\}_{q=1}^{4}$ are elicited directly from the respondent $\forall k \in S_i$.

The additive separability of the utility function rules out complementarities across outcomes. This might be a strong assumption, but one that is necessary since it is not feasible to elicit joint subjective distributions from respondents for such a large number of outcomes. The linearity assumption implies that respondents are risk-neutral. This assumption is primarily made so that it would suffice to elicit the expected value for the continuous outcomes. In the absence of this assumption, one would have to elicit multiple points on the subjective distribution for each of the continuous outcomes for each major in one's choice set (as in Dominitz and Manski, 1996). This assumption is clearly restrictive since evidence suggests that educational choices are affected by labor income risk (Nielsen and Vissing-Jorgensen, 2006). However, relaxing this assumption would not have been feasible for the purposes of this study.

Now consider the case of double majors. The relevant set of major-specific outcomes, vectors **b** and **d**, remain the same as in the case of a single major. However, the utility function now becomes more complex. It may be useful to think about why an individual may decide to choose two majors. Respondents pursuing more than one major were asked to explain their reasons; selected responses are shown in Section A.2 of the Appendix. Two main reasons emerge. First, two majors appropriately differentiated can provide a broader mix of options than a single major. Second, it might be the case that no single major meets the needs of the individual. To capture the enhanced options and specialization of function that two majors provide, I assume that the utility of a pair of majors depends on the attributes of each major separately, as well as on the attributes of a composite major combining the best of both majors. The expected utility function of a pair of majors p consisting of majors p_1 and p_2 takes the form:[10]

$$U_{ip} = U_{ip_1}(\mathbf{b}, \mathbf{d}, \{P_{ip_1}(b_r = 1)\}_{r=1}^{7}, \{E_{ip_1}(d_q)\}_{q=1}^{4})$$
$$+U_{ip_2}(\mathbf{b}, \mathbf{d}, \{P_{ip_2}(b_r = 1)\}_{r=1}^{7}, \{E_{ip_2}(d_q)\}_{q=1}^{4}) \quad (3)$$
$$+U_{i\widetilde{p}}(\mathbf{b}, \mathbf{d}, \{\max[P_{ip_1}(b_r = 1), P_{ip_2 t}(b_r = 1)]\}_{r=1}^{7},\ \{\max[E_{ip_1}(d_q), E_{ip_2}(d_q)]\}_{q=1}^{4}),$$

where $\widetilde{p}$ refers to the composite major, and $U_{ip_1}(.)$ is as defined in equation (2). Because there is no way of specifying a "primary" and a "secondary" major, I use the same functional form for the utility of each major in one's major pairing, i.e., $U_{ip_1} = U_{ip_2}$. Since $U_{ip_1}(.)$ is linear-in-parameters, the average characteristics of the two majors appear in the utility function. Thus, equation (3) can be written as:[11]

$$U_{ip}(\mathbf{b}, \mathbf{d}, \{P_{ip_1}(b_r), E_{ip_1}(d_q), P_{ip_2}(b_r), E_{ip_2}(d_q)\}_{r\in\{1,..,7\}, q\in\{1,..,4\}})$$

$$= \alpha_{p_1} + \alpha_{p2} + \sum_{r=1}^{7}\{\tfrac{P_{ip_1}(b_r=1)+P_{ip_2}(b_r=1)}{2}\}\triangle u_{r1} + \sum_{q=1}^{4}\gamma_{q1}\{\tfrac{E_{ip_1}(d_q)+E_{ip_2}(d_q)}{2}\}$$
$$+\sum_{r=1}^{7}\max[P_{ip_1}(b_r = 1), P_{ip_2}(b_r = 1)]\triangle u_{r2} \quad (4)$$
$$+\sum_{q=\{1,2\}}\gamma_{q2}\min[E_{ip_1}(d_q), E_{ip_2}(d_q)] + \sum_{q=\{3,4\}}\gamma_{q2}\max[E_{ip_1}(d_q), E_{ip_2}(d_q)] + \varepsilon_{ip}$$

$$= U_{ip} + \varepsilon_{ip},$$

where $\{\triangle u_{r1}, \triangle u_{r2}\}_{r=1}^{7}$ and $\{\gamma_{q1},\ \gamma_{q2}\}_{q=1}^{4}$ are the parameters of the utility function that need be estimated. The probability that an individual i with a choice set S_i and subjective beliefs $\{P_{ik}(b_r)\}_{r\in\{1,..,7\}}, P_{ik}(d_q)_{q\in\{1,..,4\}}\}$ for $\forall k \in S_i$ chooses a major pair p is then:

$$\Pr(p|\{P_{ik}(b_r), P_{ik}(d_q)\}_{r\in\{1,..,7\}, q\in\{1,..,4\};\ k\in S_i}) = \Pr(U_{ip} + \varepsilon_{ip} > U_{is} + \varepsilon_{is}) \quad \forall s \in S_i,\ p \neq s, \quad (5)$$

where U_{ip} is defined in equation (4), and s is the set of all possible major pairs and single majors in i's choice set.

[10] Manski and Sherman (1980) use a similar approach to model the composition of motor vehicles in two-vehicle households.

[11] The vector $\mathbf{b}$ consists of b_1= graduating in 4 years, b_2= graduating with a GPA of at least 3.5, b_3= enjoying the coursework, b_4= parents approving of the major, b_5= getting a job on graduation, b_6= enjoying work at the jobs, and b_7= being able to reconcile work and family at the jobs. The vector $\mathbf{d}$ consists of d_1= average hours per week spent on coursework, d_2= average hours per week spent at the job, d_3= social status of the job, and d_4= expected income at the age of 30.

The composite-major representation captures the notion of functional specialization as follows. Say, an individual with a pair of majors chooses one major with a low-completion probability (i.e., $P(b_1)$ is low) because of some of its other attributes and a second major where the completion probability is the most important consideration. Given the specification above, one would expect $\triangle u_{11} \approx 0$ and $\triangle u_{12} > 0$ in this case of extreme specialization. On the other hand, for an individual who values the completion probabilities associated with both his majors equally, one would expect $\triangle u_{12} \approx 0$ and $\triangle u_{11} > 0$. Thus the ratio $\triangle u_{12}/\triangle u_{11}$ (and similarly $\{\triangle u_{r2}/\triangle u_{r1}\}_{r=1}^{7}$, $\{\gamma_{q2}/\gamma_{q1}\}_{q=1}^{4}$) is a measure of the extent to which an individual desires to functionally specialize his majors along the given outcome.

The exact parametric restrictions on the random terms required for identifying the model parameters are outlined in section 4, which discusses the estimation. In the baseline model outlined in this section and estimated in section 4.1, I assume that the utility function is identical for all individuals up to a random term ε_{ip}. This assumption will be relaxed later when I allow for unobserved heterogeneity in Section 4.2.

3 Data

To estimate a choice model of double majors, the subjective beliefs about the outcomes associated with a major, $P_{ik}(\mathbf{b}, \mathbf{d})$, need to be elicited for each major ($\forall k \in S_i$) in individual i's choice set. Since the range of majors available to students and institutional details vary considerably across institutions, one standard survey cannot be used to collect data in different settings. As a first step towards understanding how college students choose the composition of double majors and whether they hedge along certain dimensions, I focus on Northwestern University and use data on a sample of students who were pursuing more than one major at the time of the survey. This section describes the institutional details at Northwestern University, the data collection method, and the subjective data.

3.1 Institutional Details

For the purposes of this study, I focus on students who are in the process of choosing their majors but have not necessarily chosen one. There are several reasons for this criteria: Students who are in the process of choosing a major are actively thinking about the occurrence of outcomes associated with the majors, and hence their responses to subjective questions related to majors are likely to have less measurement error, and more likely to be meaningful. Second, interviewing students who have already chosen their majors raises the issue of cognitive dissonance (Festinger, 1957), i.e., students who have already chosen their major could rationalize their choice of major by devaluing their beliefs for outcomes associated with the majors they considered but rejected, and upgrading their beliefs for outcomes associated with the major that they chose. This systematic measurement error in elicited subjective beliefs would be problematic, and would result in biased estimates of preference parameters. Northwestern University requires students to declare their major by the end of their sophomore year. Surveying juniors and seniors would exacerbate issues arising from cognitive dissonance. On the other hand, freshmen may have little idea of what majors they want to pursue when they first arrive in college, and may not have seriously thought about the likelihood of the various outcomes conditional on choice of major. Therefore, I restrict my sample to Northwestern University sophomores.[12]

The study is further restricted to schools at Northwestern University that accord students flexibility in choosing a major. For example, a student in the School of Journalism has to declare his major at the time of admission and can change his major only by a special request to the school. For such a student, the choice of college and majors is jointly determined. Since I model the choice of majors conditional on deciding to attend Northwestern University, such students are not eligible for the study. Therefore, I restrict the study to students with majors either in the Weinberg College of Arts & Sciences (WCAS) or in the School of Engineering.

The choice set for an individual is assumed to be exogenous. WCAS offers a total of 41 majors. To estimate the choice model, subjective probabilities have to be elicited for the occurrence of outcomes for each major in the respondent's choice set (i.e., for the majors that the individual is pursuing, as well as for the other majors in the individual's choice set). In order to limit the size of the choice set, similar majors are pooled together. Table 1 shows

[12]Readers may refer to Zafar (2010b) which empirically tests for various cognitive biases in subjective data, using panel data on beliefs about major-specific outcomes collected from a subset of this sample and single-major respondents. Systematic measurement error and other issues arising from the various cognitive biases are found to be minimal.

the majors divided into various categories. Categories a through g span the majors offered in WCAS, while category h spans the undergraduate majors offered in the School of Engineering. There is a trade-off between the number of categories and the length of the survey. This categorization is fairly fine and also seems reasonable. Each student is asked his subjective beliefs, $\{P_{ik}(b_r)\}_{r\in\{1,..,7\}}, P_{ik}(d_q)_{q\in\{1,..,4\}}\}$, for the 7 WCAS major categories (a through g) and category h, Engineering.

3.2 Sample Description

The data used in this study come from a survey administered to a sample of Northwestern University students. Sophomore students with at least one (intended) major in WCAS were recruited by e-mail and flyers posted around campus. Prospective participants were told that the survey was about the choice of college majors and that they would receive $10 for completing a 45-minute electronic survey. It was emphasized that students need not have declared their majors to participate in the study. The survey was conducted from November 2006 to February 2007, which corresponds to the first half of the students' sophomore year. Respondents were required to come to the Kellogg Experimental Laboratory to take the electronic survey. A total of 161 sophomores were surveyed, of whom 78 stated that they were pursuing more than one major. Of these 78 students, 69 reported to have both their majors in either WCAS or the School of Engineering. I use data on these 69 students in this paper.[13]

Analysis of the distribution of double majors in the sample shows that, with the exception of five instances, all major pairs consist of majors from different categories.[14] The most common major pair in the sample consists of Social Sciences I and Social Sciences II (eleven instances), followed by Area Studies and Social Sciences II (ten instances). These major pairs consist of majors that differ particularly in their labor market returns. As will be discussed later, these examples are consistent with specialization along certain dimensions. There is also some suggestive evidence that students stick to related fields when choosing double majors. For example, students with a major in Literature and Fine Arts or Social Sciences I are very likely to have Area Studies as their second major. This should be expected since related majors have

[13] Interested readers are referred to Zafar (2009) which uses data on all 161 respondents to investigate how students choose college majors.

[14] The distribution of majors is available from the author upon request.

several overlapping courses and, thus, the student has to complete fewer course (relative to a case with two unrelated majors) to graduate in them.

Even though this paper does not analyze the question of why students may decide to choose a double major, it is informative to analyze how double major respondents compare with single major students in terms of observable characteristics. Table 2 shows summary statistics for the two groups. While double major respondents tend to have a higher average GPA, arrive at the university with larger number of AP credits, receive more credits that count towards coursework, receive more financial support from their parents, and come from higher income families, none of these differences except for GPA are significant at conventional levels of significance. The survey also asked respondents for their preference ordering over the majors in their respective choice sets. In order to test for whether the determinants of major choice are similar for the two set of students (i.e., those who intend to double major and those with a single major), I estimate the preference parameters for single major choice for both groups using preference ranking data (see Zafar, 2009, for details of the model). Table A1 in the Appendix shows that the relative importance for the various determinants is similar. Enjoying coursework, gaining parents' approval and social status of the jobs are the three most important determinants for both groups. Moreover, importance of parents' approval is similar for the two groups.

3.3 Subjective Data

For the 7 WCAS major categories (a through g) and Engineering (category h), the survey elicited the probability of the occurrence of the binary outcomes, i.e., $P_i(b_r = 1)$ for $r = \{1, .., 7\}$, and the expected value for the continuous outcomes, i.e., $E_i(d_q)$ for $q = \{1, .., 4\}$.

Questions eliciting the subjective probabilities of major-specific outcomes are 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, more informative, and allow the respondent to express uncertainty (Juster, 1966; Manski, 2004).[15] As is standard in studies that collect subjective data, a short introduction, similar to the one in Delavande (2008a), was read and handed to the respondents

[15] Existing studies that have examined the role of non-pecuniary influences in the choice of schooling: Fiorito and Dauffenbach (1982), Daymont and Andrisani, (1984), Easterlin (1995), and Weinberger (2004) use questions that employ a Likert-scale.

at the start of the survey. Respondents had to answer two practice questions before starting the survey to make sure they understood how to answer questions based on the use of percentages. Here, I present some of the questions that elicited the subjective expectations. For example, the belief for the binary outcome a_2 was elicited as follows:

> *If you were majoring in [X], 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)?*

The question eliciting the expected number of hours per week spent on coursework (a_4) was:

> *If you were majoring in [X], how many hours per week do you think you will need to spend on the coursework?*

Social status of the available jobs (c_5) was elicited as follows:[16]

> *Look ahead to when you will be 30 years old. Rank the following fields of study according to your perception of the social status of the jobs that would be available to you and that you would accept if you graduated from that field of study.*

Wording for the question that elicited expected income (c_6) was similar to that in Dominitz and Manski (1996):

> *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 [X]. What is the average amount of money that you think you will earn per year by the time you are 30 YEARS OLD?*

The short introduction, practice questions, and questions eliciting beliefs about major-specific outcomes can be viewed in Section A.1 of the Appendix. The 15 questions that elicit beliefs about major-specific outcomes were asked for *each* of the 8 major categories. The full questionnaire (which also collected data on demographic information, formation of beliefs, etc.) is available from the author on request.

[16] This question elicits an ordinal ranking of the social status of the jobs. However, I treat these ordinal responses as cardinal in the choice model. In hindsight, this question should have been asked in terms of subjective expectations of getting a high-status job, since the ordinal ranking does not reveal the respondent's uncertainty about the outcome.

3.4 The Data

This section provides a brief description of the subjective data elicited from the students. There are three main reasons for doing this: 1) to check whether students perceive differences in the occurrence of the various outcomes across majors; 2) to highlight the heterogeneity in responses across students; and 3) to show how subjective data compare with objective measures.

Table 3 shows the mean belief for each of the eleven outcomes for the eight main major categories. This table shows substantial variation in mean belief for the same outcome across the various major categories, indicating that students 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 0.83 (on a 0-1 scale) for Engineering and Math & Computer Studies to 0.95 for Social Sciences I. Similarly, the mean belief about gaining parents' approval varies from a low of 0.59 for Literature and Fine Arts to a high of 0.87 for Natural Sciences. Since students associate Natural Sciences with a higher social status (mean value of 0.72 on a 0-0.9 scale versus 0.38 for Literature & Fine Arts) and higher returns in the labor market (mean earnings of $89,980 at the age of 30 versus $49,770 for Literature & Fine Arts), this suggests that students believe that parents are more likely to approve of majors associated with higher status and returns in the labor market. The table also shows that, on the whole, students pursuing more than one major have mean beliefs similar to those of students pursuing only one major. The only exception is the belief about expected hours per week spent on coursework; double major respondents report beliefs that are significantly lower than those of single major respondents for this outcome across all major categories.

The mean beliefs reported in Table 3 mask the heterogeneity in responses across respondents for the *same* outcome. Figures 1 through 4 present the distribution of the beliefs for some of the outcomes. The figures show that there is substantial heterogeneity in beliefs in my relatively homogenous sample. Respondents seem to be willing to use the entire scale from zero to 100. There has been some 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. Moreover, there is no evidence of anchoring, since numbers that were presented in the introductory text do not occur more often than others. The figures also indicate that the distribution of beliefs in a given major is similar for double major and single major respondents.

The figures show that there is substantial heterogeneity in beliefs for the outcomes *across* majors as well. Figure 1 presents the cumulative distribution of the belief of being able to graduate with a GPA of at least 3.5 in Engineering and Literature & Fine Arts. The belief distribution for Literature & Fine Arts first order stochastically dominates the Engineering belief distribution (both for single and double major respondents). While less than 40% of the respondents believe that there is a greater than 60% chance of graduating with a GPA of at least 3.5 in Engineering, more than 80% of the respondents believe that to be case in Literature & Fine Arts. This is consistent with data available from the Northwestern 2006 Graduate Survey, according to which the average GPA of Northwestern Engineering graduates of 2006 was 3.43, while that of Literature & Fine Arts was 3.56.

Figure 2 presents the distribution of parents' approval in Area Studies and Natural Sciences. The belief distribution in the case of Natural Sciences first order stochastically dominates that of Area Studies, which is consistent with the hypothesis that (students believe that) parents are more likely to approve majors that are associated with higher social status and with better prospects in the labor market. Figure 3 shows that students perceive higher chances of getting an acceptable job in the case of Social Sciences II relative to Social Sciences I; this is consistent with anecdotal evidence of better job prospects in Social Sciences II which includes Economics. Finally, Figure 4 shows that the belief distribution of being able to reconcile family and work at the jobs in Natural Sciences is first order stochastically dominated by the corresponding distribution in Literature and Fine Arts. This is consistent with the general perception of hectic work schedules in the pure sciences and the medical profession in which most Northwestern University Natural Sciences bachelors' graduates get jobs.

Analysis of how subjective data compare with objective measures is not an easy task since that requires *appropriate* objective measures to which the beliefs can be compared.[17] Such an

[17] Also, note that what one can learn from such a comparison exercise is limited. Such an exercise is informative in shedding light on whether respondents have well-formed expectations. However, if one were to find that subjective data don't match very well with objective measures, it doesn't weaken the case to use subjective data in explaining choices as long as the reported data are the *true* beliefs that students use when making choices.

exercise is not possible for non-pecuniary outcomes such as approval of parents or enjoying coursework since no objective measures exist for these outcomes. Even in instances where objective measures exist such as for expected income or GPA, they correspond to outcomes for students who chose 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 chose that major may not be the *correct* objective measure.[18] However, since those are the only data available, Table 4 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. Column (1a) of Table 4 shows the mean GPA by major category of bachelor graduates 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. For comparison purposes, columns (3a) and (3b) report the responses about graduating with a GPA of at least 3.5 for the single major respondents in the sample.[19] Based on mean responses, it seems that double major respondents, relative to their single major counterparts, report a higher probability of being able to graduate with a GPA of at least 3.5 in most categories (though the difference is only statistically significant in three of the cases). This result could possibly be explained by the selection into who chooses to double major: Average GPA of double major respondents is 3.49 opposed to 3.39 for single major respondents. Columns (4)-(6) 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

[18] Respondents' expectations may differ from the objective measures for other legitimate reasons too. 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.

[19] Data collected from single major respondents is not used in the estimation of the choice model in this paper but instead used in Zafar (2009). Their mean beliefs are reported here for comparison purposes.

majors by income reported by the double major 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, the responses of single major and double major respondents are not statistically different.

On the whole, analysis of the subjective data indicates that students are aware of differences across majors along the various dimensions, and that they have well-formed and sensible expectations. At the same time, as shown in Figures 1 through 4, there is substantial heterogeneity in beliefs about outcomes *across* majors as well as *within* majors. This section also highlights the advantage of eliciting beliefs as probabilistic expectations since simple binary responses would be unable to unmask this heterogeneity entirely.

4 Estimation

4.1 Baseline Model

In order to estimate the choice model of double majors described in Section 2, I assume that the random terms $\{\varepsilon_{ip}\}$ are independent for every individual i and every alternative p, and have a Type I extreme value distribution, implying that $\{\varepsilon_{ip} - \varepsilon_{is}\}$ has a standard logistic distribution.[20] Then equation (5), i.e., the probability that individual i chooses major pair p is:

$$\Pr(p|\{P_{ik}(b_r), E_{ik}(d_q)\}_{r\in\{1,..,7\},q\in\{1,..,4\};\,k\in S}) = \frac{\exp(U_{ip})}{\sum_{s\in S}\exp(U_{is})},$$

where U_{ip} is as defined in equation (4), and S is the set of all possible single and double major alternatives. For estimation, I assume that the set of alternatives, S, includes all subsets of two majors in WCAS and the School of Engineering (${}^8C_2 = 28$) and all possible single majors in WCAS (7), for a total of 35 alternatives.[21]

The elicited subjective probabilities, $\{P_{ik}(b_r = 1)\}_{r=1}^{7}$, and elicited expected values, $\{E_{ik}(d_q)\}_{q=1}^{4}$, described in Section 3.4 are used in estimation. The heterogeneity in these beliefs is crucial for identification of the model parameters. The parameters of interest are the eight major-specific

[20] This model exhibits the IIA property. Since it might be easier to choose particular double major combinations, this assumption is relaxed in Section 4.2 which allows for flexible substitution patterns between majors.

[21] The subscript on the choice set, S, is now dropped because, as mentioned earlier, the choice set is treated as exogenous and everyone has the same set of alternatives.

constants $\{\alpha_m\}_{m=1}^8$, $\{\triangle u_{r1}, \triangle u_{r2}\}_{r=\{1,3,4,5,6,7\}}$, $\triangle u_{21}$ and $\{\gamma_{q1}, \gamma_{q2}\}_{q=1}^4$, and they are identified under these parametric assumptions.

The maximum-likelihood estimates are shown in column (1) of Table 5. All of the outcomes associated with college are significant, while in the workplace, outcomes that are statistically significant are finding a job upon graduation, enjoying work at the jobs and hours per week spent at work. The estimates indicate specialization in the case of graduating in 4 years ($\triangle u_{12} > 0$, $\triangle u_{11} \approx 0$), graduating with a GPA of at least 3.5 ($\triangle u_{22} > 0$, $\triangle u_{21} \approx 0$), and finding a job upon graduation ($\triangle u_{52} > 0$, $\triangle u_{51} \approx 0$). This implies that students concentrate their chances of graduating in four years, graduating with a high GPA, and getting a job upon graduation in one of their two majors.[22] On the other hand, gaining approval of parents, enjoying coursework and enjoying work at the jobs are outcomes that are important in the choice of both majors (i.e., $\triangle u_{41} > 0$, $\triangle u_{42} \approx 0$; $\triangle u_{31} > 0$, $\triangle u_{32} \approx 0$; $\triangle u_{61} > 0$, $\triangle u_{62} \approx 0$, respectively). This suggests that the hypothesis that students major in one field to satisfy their own interests and in another to meet parents' approval does not hold in my data. The coefficient on hours per week spent on coursework, γ_{12}, is negative which supports the specialization hypothesis, i.e., individuals prefer pairs of majors that entail different hours per week in college. Column (1) of Table 6 shows the absolute ratio of $\{\triangle u_{r2}/\triangle u_{r1}\}_{r=1}^7$ and $\{\gamma_{q2}/\gamma_{q1}\}_{q=1}^4$ using the model estimates from column (1) of Table 5. This ratio is a measure of the extent to which an individual specializes his majors along the given outcome: A ratio greater than one indicates that the student desires to choose majors that differ in the likelihood of that outcome (i.e., he prefers specialization), while a ratio of less than one indicates that the student values that outcome in the choice of both majors. The table shows that there is strong evidence of specialization for graduating in 4 years, i.e., students prefer to choose majors that differ in their chances of completion. On the other hand, there is strong evidence that enjoying coursework, parents' approval, and enjoying work at the jobs are important determinants in the choice of both majors in one's major pair. There is weak evidence of specialization along the dimension of graduating with a GPA of at least 3.5, finding a job upon graduation, and reconciling work and family at available jobs; all three ratios are greater than 2 but not statistically different from one.

[22] There is evidence of the latter in the comments submitted by the respondents (see Section A.2).

Because of the non-linear nature of the model, it is hard to interpret the coefficients on the various outcomes. To get a measure of the magnitude of the estimates in column (1) of Table 5, the natural thing would be to do willingness-to-pay calculations, i.e., translate the differences in utility levels into the amount of earnings that an individual would be willing to forgo at the age of 30 in order to experience that outcome.[23] However, since expected income at age 30 is not significant, the standard errors on such calculations are huge, and the results are not very meaningful. Instead of presenting the willingness-to-pay calculations, I use a decomposition to gain insight into the relative importance of the various outcomes in the choice. For illustration, suppose that $\Pr(choice = j) = F(\mathbf{X}_j\boldsymbol{\beta})$ and that $\mathbf{X}$ includes two variables, X_1 and X_2. Given the parameter estimates, $\widehat{\beta_1}$ and $\widehat{\beta_2}$, the contribution of X_1 to the choice is defined as:

$$M_{X_1} \equiv \; || \; \overline{\Pr(choice \; = j| \; \{\widehat{\beta_1}, \widehat{\beta_2}\}} \; - \; \overline{\Pr(choice \; = j| \; \{\widehat{\beta_1} = 0, \widehat{\beta_2}\}} \; ||$$

$$= \sqrt{\sum_j \left[\; \sum_{i=1}^N \frac{\Pr(choice = j| \; \{\widehat{\beta_1}, \widehat{\beta_2}\})}{N} \; - \; \sum_{i=1}^N \frac{\Pr(choice = j| \; \{\widehat{\beta_1} = 0, \widehat{\beta_2}\})}{N} \right]^2},$$

where the first term is the average probability of majoring in major j as predicted by the model, and the second term is the average predicted probability of majoring in j if outcome X_1 were not considered. The difference in the two terms is a measure of the importance of X_1 in the choice. The *relative* contribution of X_1 to the choice is then $R_{X_1} = \frac{M_{X_1}}{M_{X_1} + M_{X_2}}$. Column (1) of Table 7 presents the results of this decomposition strategy using the estimates from column (1) of Table 5. Each cell shows the *relative* contribution (R) of the outcome to the choice. The table shows that about 55% of the choice is explained by outcomes realized in college, with enjoying coursework and gaining parents' approval, each, explaining nearly one-fifth of the choice.

4.2 Robustness Checks

4.2.1 Modified Choice Model

It could be argued that outcomes associated with the workplace come as a package; for example, one does not have the option to choose the income associated with jobs available in one major and the lifestyle associated with jobs in the second major. If that were the case, the model estimated in Section 4.1 will be biased. I modify the choice model described in Section 2, and

[23] For example, the amount that an individual would be willing to forgo in earnings at the age of 30 for a 2% change in the probability of outcome j is $\frac{0.02 \times \triangle u_j}{\gamma_{41} + \gamma_{42}}$.

apply the idea of a composite major only to outcomes associated with college. The outcomes for which the composite-major specification is used are graduating in four years, graduating with a GPA of at least 3.5, hours per week spent on coursework, enjoying the coursework, gaining approval of parents, and finding a job upon graduation.

The model estimates are shown in column (2) of Table 5. The estimates are qualitatively similar to those of the model which uses the composite-major representation for all outcomes. As before, all outcomes associated with college are significant. In the workplace, outcomes that are statistically significant are finding a job upon graduation, enjoying work at the jobs, being able to reconcile work and family at the jobs, and hours per week spent at work. The estimates indicate specialization in the case of graduating in 4 years, graduating with a GPA of at least 3.5, expected coursework hours per week, and finding a job upon graduation. As before, gaining approval of parents and enjoying coursework are important determinants in the choice of both majors. The results of the ratios shown in column (2) of Table 6, and the relative contribution of the various outcomes in the choice shown in column (2) of Table 7 are also qualitatively similar to earlier results indicating that the model with the full composite-major representation is not misspecified.

4.2.2 Flexible Substitution Patterns

The models that have been estimated so far exhibit the restrictive IIA property, which is not a very realistic assumption in this particular situation. For example, one could imagine that an individual majoring in Area Studies and in Literature & Fine Arts is more likely to choose Area Studies and Ethics & Values, rather than Natural Sciences and Ethics & Values. Similarly, it may be easier to major in certain major pairs if they have several overlapping course requirements. To allow flexible substitution patterns, I allow for a stochastic part for each major that is perhaps correlated over majors and heteroskedastic over individuals and majors (these appear as 8 random effects, one for each of the 7 alternatives in WCAS and one for Engineering) and another stochastic part that is iid over individuals and alternatives.[24] The

[24]This approach is similar to Brownstone and Train (1998), who use a "mixed logit" choice model without the IIA property and with flexible substitution patterns to forecast demand for new vehicles.

utility function of a pair of majors p is now:

$$U_{ip}(\mathbf{b},\mathbf{d},\{P_{ik}(b_r),E_{ik}(d_q)\}_{r\in\{1,..,7\},q\in\{1,..,4\};\ k\in S}) = U_{ip}\ +\ \varepsilon_{ip}\ + c_{p1}\eta_{i,1} + c_{p2}\eta_{i,2} + ... + c_{p8}\eta_{i,8},$$

where U_{ip} is as defined in equation (4), excluding the major-specific constants, and ε_{ip} is a random term with zero mean that is iid over alternatives of major pairs and is normalized to set the scale of utility. The $\eta_{i,k}$ for $k = \{1,..,8\}$ are normally distributed effects with zero mean, and $c_{px} = 1$ if major x appears in the pair of majors p.[25] For example, the utility function of a pair of majors p that includes Natural Sciences ($k = 1$) and Social Sciences II ($k = 4$) would be:

$$U_{ip}(\mathbf{b},\mathbf{d},\{P_{ik}(b_r),E_{ik}(d_q)\}_{r\in\{1,..,7\},q\in\{1,..,4\};\ k\in S}) = U_{ip} + \varepsilon_{ip} + \eta_{i,1} + \eta_{i,4}.$$

This structure allows flexible substitution patterns across alternatives. For example, the correlation between a pair of majors κ consisting of majors $\{1,2\}$, and a second pair of majors ω consisting of majors $\{2,3\}$ is $E([U_{i\kappa} + \varepsilon_{i\kappa} + \eta_{i,1} + \eta_{i,2}][U_{i\omega} + \varepsilon_{i\omega} + \eta_{i,2} + \eta_{i,3}]) = Var(\eta_{i,2})$. So utility is now correlated over alternatives. Given the vector $\boldsymbol{\eta}_i$, the *conditional* choice probability is simply logit, since the remaining error term is iid extreme value. The probability of individual i choosing the pair of majors p *conditional* on $\boldsymbol{\eta}_i$ is:

$$\begin{aligned}\Pr(p|\boldsymbol{\eta}_i) &= \Pr(p|\ \{P_{ik}(b_r),E_{ik}(d_q)\}_{r\in\{1,..,7\},q\in\{1,..,4\};\ k\in S},\boldsymbol{\eta}_i) \\ &= \frac{\exp(U_{ip}\ + c_{p1}\eta_{i,1} + c_{p2}\eta_{i,2} + ... + c_{p8}\eta_{i,8})}{\sum_{k\in S}\exp(U_{ik}\ + c_{k1}\eta_{i,1} + c_{k2}\eta_{i,2} + ... + c_{k8}\eta_{i,8})}.\end{aligned}$$

The *unconditional* probability of choosing p is the integral of this *conditional* probability over all possible values of $\boldsymbol{\eta}_i$, and depends on the density of $\boldsymbol{\eta}_i$. I denote this density $\mathbf{g}(\boldsymbol{\eta}_i|\Omega)$ where Ω are the parameters of the distribution. The *unconditional* probability for i choosing p is:

$$P_{ip}(\Omega) = \int \Pr(p|\boldsymbol{\eta}_i)\mathbf{g}(\boldsymbol{\eta}_i|\Omega)d\boldsymbol{\eta}_i.$$

Since the integral does not have a closed form in general, it is approximated through simulation.[26] The estimated parameters from maximizing the simulated log-likelihood, $\sum_i \ln(\widehat{P_i(\Omega)})$, are shown in Table 8. The first column presents the parameters of a model that allows the

[25] This is identical to allowing the major-specific constants in equation (4) to be normally distributed.

[26] I use 10,000 draws of $\boldsymbol{\eta}_i$ for a given value of the parameters Ω. For each draw, the $\Pr(p|\boldsymbol{\eta}_i)$ is calculated, and the

composite-major categorization for all outcomes (except GPA), while the second column only allows the composite-major categorization for outcomes in college. The coefficients are similar in relative magnitude, but larger in absolute terms than the corresponding fixed coefficients in Table 5. This is expected because the variance of the error term in the standard logit model (ε) is larger than in a mixed logit since some of the variance is now captured by the η's rather than the ε in the mixed logit model. Since utility is scaled so that ε has the variance of an extreme value, the variance before scaling is larger in the standard logit than in the mixed logit, and hence parameters are scaled down in a standard logit relative to the mixed logit (Revelt and Train, 1998).

In line with the previous estimates, I find that students choose a major pair such that they enjoy the coursework, gain approval of parents, and enjoy working at the jobs in *both* majors ($\triangle u_{31} > 0$, $\triangle u_{32} \approx 0$; $\triangle u_{41} > 0$, $\triangle u_{42} \approx 0$; $\triangle u_{61} > 0$, $\triangle u_{62} \approx 0$, respectively). There is evidence that individuals prefer majors that differ in their chances of graduating in four years ($\triangle u_{12}/\triangle u_{11} > 1$). The coefficient on $\min[E_{ip_1}(d_1), E_{ip_2}(d_1)]$, i.e., γ_{12} is negative, suggesting that individuals prefer pairs of majors with different time commitments at college. In general, the results are similar to earlier findings in Tables 5 and 6, but the estimates are less precise now. One possible reason for this could be that the sample size is small and the number of parameters to be estimated in the model with error components for the majors is larger.

Columns (3) and (4) of Table 7 show the relative contribution of the various outcomes using the estimates from columns (1) and (2) of Table 8, respectively. The results are qualitatively similar to those in the first two columns: Enjoying coursework, gaining approval of parents, and graduating with a GPA of at least 3.5 continue to be the more important determinants. One notable difference from the previous results is that graduating in 4 years is now the most important determinant of the choice.

average of these probabilities is taken as the approximate choice probability:

$$\widehat{P_{ip}(\Omega)} = \frac{1}{10,000} \sum_{d=1}^{10,000} \Pr(p|\boldsymbol{\eta}_i^d).$$

4.2.3 Unobserved Heterogeneity

The baseline model estimated in Section 4.1 assumes that all individuals have homogeneous preferences for various outcomes. However, any unobserved heterogeneity may bias the model estimates. As a robustness check, I specify a random parameters logit model to account for these issues (see Revelt and Train, 1998, for a discussion of mixed logit models). One could allow heterogeneity in preferences for all outcomes, but I focus on the most important outcomes because of the sample size. I consider a model in which the preference parameters for enjoying the coursework, gaining approval of parents and finding a job upon graduation are allowed to vary in the population with a specified distribution. More specifically, I assume that $\triangle u_{31i}$, $\triangle u_{32i}$, $\triangle u_{41i}$, $\triangle u_{42i}$, $\triangle u_{51i}$, and $\triangle u_{52i}$ are independently log-normally distributed.

As in Section 4.2.2, the log-likelihood function is approximated by the simulated log-likelihood function. Estimates of various outcomes (available from the author on request) are similar to those obtained in the corresponding model with no heterogeneity (column 1 of Table 5). The mean coefficient of enjoying coursework is still largest in absolute value and significant. The estimated standard deviations of the (random) coefficients are highly significant, indicating that these parameters do indeed vary in the sample. For the same reasons as in Section 4.2.2, the mean coefficients in the mixed logit model are larger than the corresponding fixed coefficients in Table 5.

On the whole, the results are similar to earlier findings, suggesting that there is no significant bias coming from ignoring unobserved heterogeneity or from the IIA assumption made in the model estimation in Section 4.1.

5 Conclusion

This paper investigates how college students choose their majors conditional on pursuing a double major. Since college majors are chosen under uncertainty (about personal tastes, academic ability, and realizations of major-specific outcomes), students use their preferences and expectations about the uncertain future aspects of the choice when making their decision of what to major in. In order to overcome the identification problem of inferring preferences from observed choices when expectations are also unknown, I elicit expectations from a sample of

69 Northwestern students pursuing double majors, and use them directly to estimate a choice model of double majors.

The subjective data reveal that there is substantial heterogeneity in expectations in my relatively homogenous sample, indicating tremendous heterogeneity in beliefs among the population of college students. This also raises concerns about the accuracy of restrictions imposed on expectations in the literature. Analysis of expectations data reveals that students seem to be aware of differences across majors along the various dimensions and that their responses seem to be meaningful. Moreover, I find that students believe that their parents are more likely to approve majors that are associated with higher social status and higher labor market returns. I use the expectations data to estimate a choice model of double majors. Contrary to stories in the popular press and anecdotal evidence (Lewin, 2002; Gomstyn, 2003), I do not find that students major in one field to satisfy their own interests and another to gain parents' approval. Instead the results show that students take into account parents' approval as well as how much they'll enjoy studying and working in that field when choosing *both* majors in their major pair. However, I do find that students act strategically by choosing majors that differ in their chances of completion, in how academically challenging they are (in terms of getting good grades in them), and in finding a job upon graduation.

The analysis in this paper is based on data from Northwestern University, and it's not clear whether the findings of this study can be generalized to other settings. This paper clearly calls for similar data collection of college major-related expectations at a larger scale (Arcidiacono et al., 2010, collect similar data from Duke University undergraduates to estimate a model of single major choice). Moreover, this paper investigates *how* students choose double majors (conditional on pursuing a double major) without attempting to explain *why* some students may choose double majors. Given that a large fraction of college students have more than one undergraduate major and the share of students pursuing more than one college major is increasing, a natural question to ask is what drives some students to choose a single major and others to pursue more than one major. The available data are insufficient to answer this question. However, the approach used in this paper of eliciting subjective expectations data and using them to understand decision-making under uncertainty could be useful in shedding light on this.

From a policy viewpoint, it is not clear whether pursuing more than one major is good or bad. College administrators have expressed concerns about possible negative impacts of loading on majors, such as neglecting extracurricular activities and elective classes essential for a balanced education. On the other hand, studies have shown that double majors are associated with higher returns in the labor market (Del Rossi and Hersch, 2008). The question of whether double majors should be encouraged or not is important, and one that remains unanswered. Data on outcomes of students (both in college and after college) are needed to address this.

References

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

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

[3] Arcidiacono, Peter, V. Hotz, and Songman Kang. 2010. "Modeling College Major Choices using Elicited Measures of Expectations and Counterfactuals." Duke University Working Paper.

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

[5] Brownstone, David, and Kenneth Train. 1998. "Forecasting New Product Penetration with Flexible Substitution Patterns." *Journal of Econometrics*, 89(1): 109-129.

[6] 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.

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

[8] Daymont, Thomas, and Paul Andrisani. 1984. "Job Preferences, College Major, and the Gender Gap in Earnings." *Journal of Human Resources*, 19(3): 408-428.

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

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

[11] Del Rosi, Alison, and Joni Hersch. 2008. "Double Your Major, Double Your Return?" *Economics of Education Review*, 27(4): 375-386.

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

[13] Easterlin, Richard. 1995. "Preferences and Prices in Choice of Career: The Switch to Business, 1972-87." *Journal of Economics Behavior and Organization*, 27: 1-34.

[14] Fiorito, Jack, and Robert Dauffenbach. 1982. "Market and Nonmarket Influences in Curriculum Choice by College Students." *Industrial and Labor Relations Review*, 36(1): 88-101.

[15] Freeman, Richard. 1971. *The Market for College-Trained Manpower:Study in the Economics of Career Choice*. Cambridge: Harvard University Press.

[16] Gomstyn, Alice. 2003. "Major Help or Major Harm?" *Chronicle of Higher Education*, 50(9): A41-A42.

[17] Juster, Thomas. 1964. *Anticipations and Purchases: An Analysis of Consumer Behavior*. Princeton: Princeton University Press.

[18] Lewin, Tamar. 2002. "For Students Seeking Edge, One Major Just Isn't Enough." *New York Times*, November 17, page 1.

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

[20] Malamud, Ofer. 2006. "Breadth vs. Depth: The Effect of Academic Specialization on Labor Market Outcomes." Harris School Working Paper Series 05:17.

[21] Manski, Charles, and Leonard Sherman. 1980. "An Empirical Analysis of Household Choice Among Motor Vehicles." *Transportation Research A*, 14A: 349-366.

[22] 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.

[23] Manski, Charles. 2002. “Identification of Decision Rules in Experiments on Simple Games of Proposal and Response.” *European Economic Review*, 46(4-5): 880-891.

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

[25] Nielsen, Helena, and Annette Vissing-Jorgensen. 2006. "The Impact of Labor Income Risk on Educational Choices: Estimates and Implied Risk Aversion." Working Paper, Kellogg School of Management, Northwestern University.

[26] Revelt, David, and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *The Review of Economics and Statistics*, 80(4): 647-657.

[27] Stinebrickner, Todd, and Ralph Stinebrickner. 2008. "Learning about Academic Ability and the College Drop-out Decision." University of Western Ontario Working Paper 20086.

[28] van der Klaauw, Wilbert. 2000. "On the Use of Expectations Data in Estimating Structural Dynamic Models: An Analysis of Career Choices", UNC-Chapel Hill Working Paper.

[29] van der Klaauw, Wilbert and Kenneth Wolpin. 2008. "Social Security and Savings and Retirement Behavior of Low Income Households." *Journal of Econometrics*, 145(1-2): 21-42.

[30] Weinberger, Catherine. 2004. "Just Ask! Why Surveyed Women Did Not Pursue Information Technology Courses or Careers?" *IEEE Technology and Society*, 23(2): 28-35.

[31] Willis, Robert, and Sherwin Rosen. 1979. "Education and Self-Selection." Journal of Political Economy, 87: S7-S36.

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

[33] Zafar, Basit. 2010a. "How Do College Students Form Expectations?" *Journal of Labor Economics*, forthcoming.

[34] Zafar, Basit. 2010b. "Can Subjective Expectations Data be Used in Choice Models? Evidence on Cognitive Biases." *Journal of Applied Econometrics*, forthcoming.

Appendix A

A.1 Survey Excerpt

A.1.1 Introduction and Practice Questions

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.

We will start with a couple of practice questions.

PRACTICE QUESTION 1: What do you think is the PERCENT CHANCE (or CHANCES OUT OF 100) that you will eat pizza for lunch next week? _ _ _ _ _ _ _ _ _ %

PRACTICE QUESTION 2: What do you think is the PERCENT CHANCE (or CHANCES OUT OF 100) that you will eat pizza for lunch on Tuesday next week? _ _ _ _ _ _ _ _ _ _ %

Once students had answered the questions, they were given the following instructions:

Since "pizza for lunch next week" INCLUDES the possibility of "pizza for lunch on Tuesday next week", your answer to PRACTICE QUESTION 2 should be SMALLER or EQUAL than your answer to PRACTICE QUESTION 1.

A.1.2 Questionnaire

The following set of questions was asked for EACH of the relevant major categories. For example, the questions below were asked for the category of 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 graduates (of 2006) 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?

A.2 Debriefing: Why Choose Two Majors?

This section presents some of the responses to the question posed to survey respondents pursuing more than one major: **"Why are you pursuing more than one major?"**

- To have more options, since I am not certain as to what career I want to follow.
- There are plenty of econ majors in the country, doubling with Math will help me stand out. Also, they complement each other well and I enjoy them both.
- My first major, MMSS, is an adjunct major. Getting a second major allows me to broaden my horizons and also specialize in a practical field. Also, I feel it looks more impressive if you have completed more than one major.
- I want to have a science major (chemistry) as well as another route (economics) for careers in life.
- One practical (MMSS) One personal interest (Linguistics). Real goal is to go to law school soon after grad. perhaps working a couple years in the consulting/finance industry.
- Because Spanish is for a career and art is for a lifetime hobby.
- Multiple personal interests, having additional options later in life, stand apart from others.

- I have a conflict between what is practical for the job prospect and what I truly would enjoy learning about, so I am pursuing one major which falls into each of the two categories.
- There is no single major at Northwestern which encompasses my interests.
- I want to have more fields open to me.
- To make it more easy to get a job and have a solid career.
- I feel that having both majors will open up a wider range of job opportunities when I graduate. I also feel that I am interested in both subjects and am taking the opportunity to further my knowledge in them.

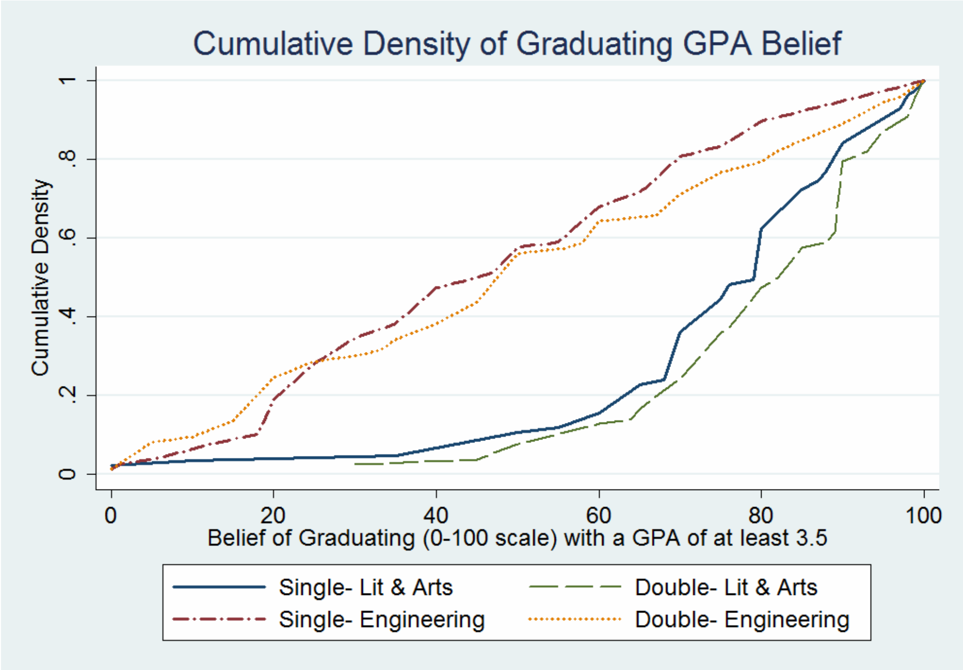

Figure 1: Cumulative density and distribution of the belief of graduating with a GPA ≥ 3.5 in Engineering and Literature & Fine Arts, for double and single major respondents.

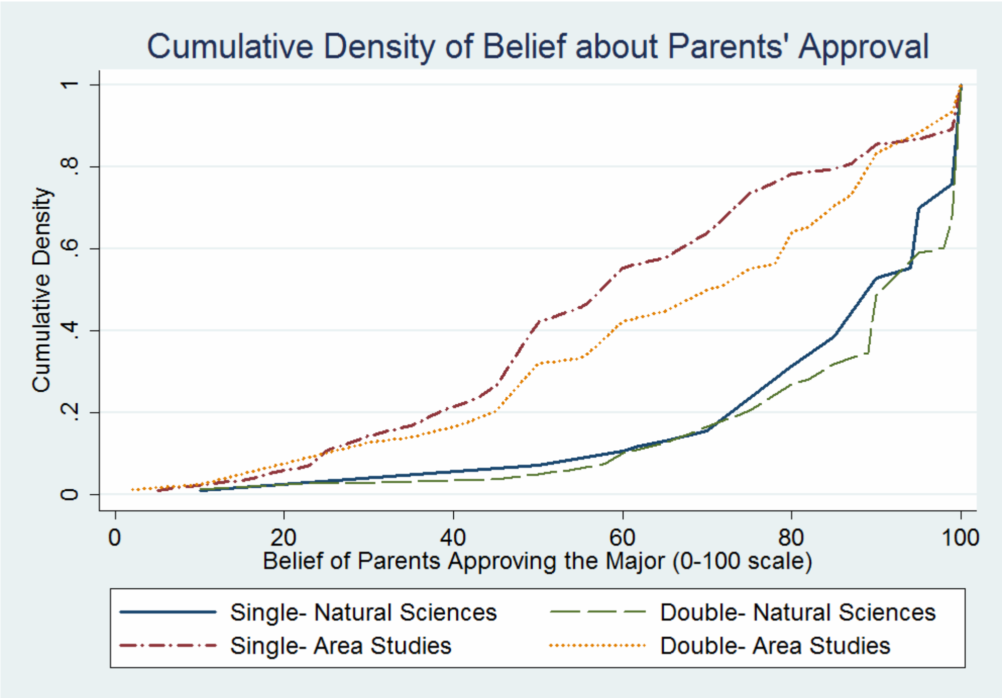

Figure 2: Cumulative density and distribution of the belief of gaining parents' approval for majoring in Area Studies and Natural Sciences, for double and single major respondents.

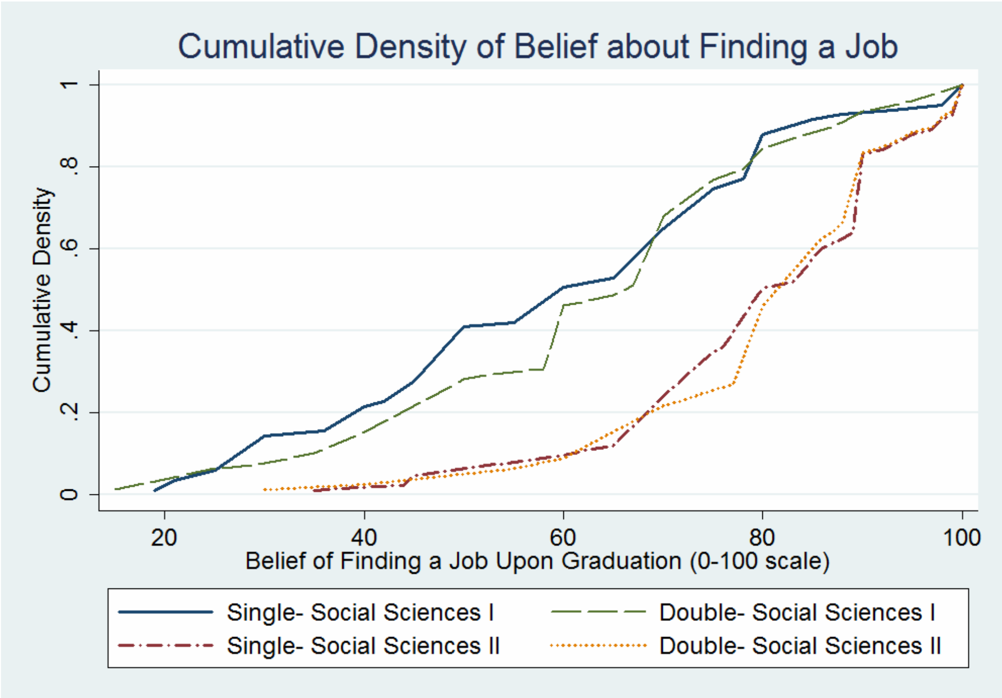

Figure 3: Cumulative density and belief distribution of finding an acceptable job upon graduation in Social Sciences I and Social Sciences II, for double and single major respondents.

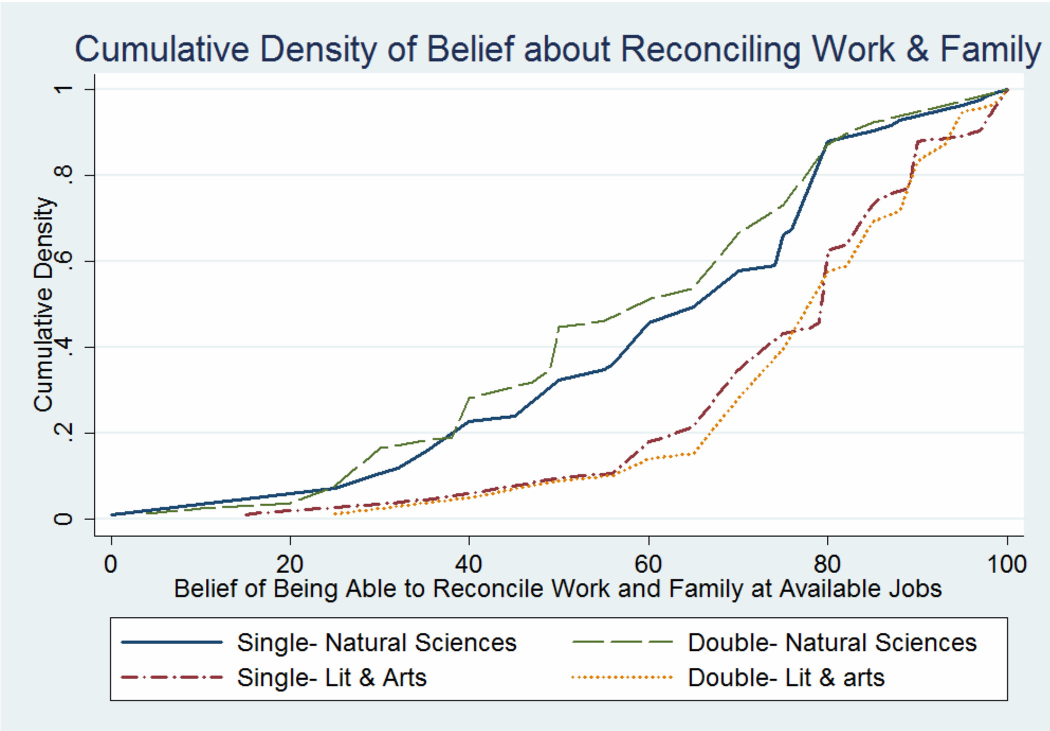

Figure 4: Cumulative density and belief distribution of being able to reconcile work and family at jobs available in Literature & Fine Arts and Natural Sciences, for double and single major respondents.

Table A1: Decomposition Analysis Using Stated Preference Data

	Double Majors	**Single Majors**
	(1)	(2)
Attributed to:		
Graduating in 4 years	1.80	4.25
Graduating with a GPA of $\geq$3.5	6.75	7.55
Enjoying the coursework	46.65	23.90
Coursework hours per week	2.20	4.75
Approval of parents and family	15.45	14.00
Finding an acceptable job upon graduation	3.15	3.95
Enjoying work at the jobs	5.50	10.60
Reconciling work and family at available jobs	2.00	4.75
Hours per week spent at work	2.70	6.60
Social status of the jobs	13.20	16.70
Expected income at the age of 30	0.60	2.95

Each cell in the table reports the relative importance of the given determinant in explaining choice of major using respondent's ranking over majors. Estimates used in this exercise come from the estimation of a single major choice model using ranking data. For details, see Zafar (2009)

Table 1: List of Majors

The following is the classification of majors into categories:

a Natural Sciences	f Area Studies
Biological Sciences	African American Studies
Chemistry	American Studies
Environmental Sciences	Asian & Middle East Languages & Civilization
Geography*	European Studies
Geological Sciences	International Studies*
Integrated Science	Slavic Languages and Literatures
Materials Science	
Physics	g Literature and Fine Arts
	Art History
b Mathematical and Computer Sciences	Art Theory and Practice
Cognitive Science	Classics
Computing and Information Systems	Comparative Literary Studies
Mathematics	Drama
Statistics	English
	French
c Social Sciences I	German
Anthropology	Italian
Gender Studies*	Spanish
History	
Linguistics	h Engineering[1]
Political Science	Applied Mathematics
Psychology	Biomedical Engineering
Sociology	Chemical Engineering
	Civil Engineering
d Social Sciences II	Computer Engineering
Economics	Computer Science
Mathematical Methods in Social Sciences*	Electrical Engineering
	Environmental Engineering
e Ethics and Values	Industrial Engineering
Legal Studies*	Manufacturing and Design Engineering
Philosophy	Materials Science& Engineering
Religion	Mechanical Engineering
Science in Human Culture*	

* Adjunct majors. These do not stand alone.

[1] Majors in the McCormick School of Engineering.

Table 2: Sample Characteristics

Characteristics	Double Majors[a] Freq./Mean	(%/Std .dev)	Single Majors[b] Freq.	(%/Std .dev)	Population[c] Freq.	(%)
	(1)		(2)		(3)	
Gender:						
Male	34	(49)	33	(40)	465	(46)
Female	35	(51)	50	(60)	546	(54)
Total	69		83		1011	
Ethnicity						
Caucasian	35	(51)	40	(48)	546	(54)
African American	4	(6)	7	(8.5)	71	(7)
Asian	25	(36)	27	(32.5)	232	(23)
Hispanic	3	(4)	2	(2.5)	61	(6)
Other	2	(3)	7	(8)	101	(10)
Average GPA:	3.50*	(0.04)	3.39	(0.05)	-	
Average AP Credit:[d]	4.33	(3.40)	3.90	(3.22)	-	
NU Credit[e]	3.96	(3.42)	3.69	(3.58)	-	
Average SAT:						
Math	719.23	(60.30)	718.97	(61.70)	-	
Verbal	701.28	(59.61)	704.52	(61.50)	-	
Parents' Support[f]	$23,076	(13,579)	$20,843	(13,882)	-	
Parents' Income[g]	$248,313	(284,601)	$217,493	(273,150)	-	

a Individuals pursuing double majors in Engineering and WCAS.
b Individuals who reported that they were pursuing a single major only.
c Population statistics for the sophomore class. (Source: Northwestern Office of the Registrar).
d Number of AP credits the respondent had when enrolled at Northwestern University.
e Number of course credits respondent received that counted towards degree requirements.
f Average support (for tuition and board) that respondent received from parents in the last academic year.
g Average annual parents' income in the last academic year.
∗ Difference in GPAs between single and double major respondents is significant at 10% (2-tailed t-test).

Table 3: Summary Statistics: Mean and Standard Deviation of Elicited Beliefs

	Graduate in 4 years	GPA ≥ 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	0.86	0.55	0.54*	24.94***	0.87	0.73	0.64	0.58	49.72	0.72**	89.98
	(0.19)	(0.30)	(0.27)	(9.30)	(0.17)	(0.20)	(0.26)	(0.23)	(12.02)	(0.20)	(98.43)
Math & Computer Sci	0.83	0.57	0.47	24.20***	0.75	0.73	0.51	0.65	43.65	0.53**	73.85
	(0.23)	(0.28)	(0.27)	(10.39)	(0.22)	(0.20)	(0.25)	(0.21)	(9.70)	(0.20)	(59.58)
Social Sciences I	0.95	0.83*	0.82*	18.52***	0.77*	0.64	0.74	0.73	43.62	0.59	64.34
	(0.09)	(0.15)	(0.15)	(6.48)	(0.19)	(0.20)	(0.17)	(0.17)	(8.92)	(0.18)	(31.08)
Social Sciences II	0.90	0.68	0.66	23.43***	0.85	0.82	0.64	0.63	50.68	0.70**	108.14
	(0.14)	(0.24)	(0.24)	(9.37)	(0.15)	(0.14)	(0.23)	(0.21)	(12.01)	(0.18)	(117.95)
Ethics and Values	0.91	0.81	0.73	19.43***	0.66	0.56	0.64	0.71	43.52	0.45	63.16
	(0.13)	(0.16)	(0.20)	(7.59)	(0.26)	(0.20)	(0.21)	(0.20)	(9.24)	(0.19)	(28.37)
Area Studies	0.93	0.83**	0.76**	19.14***	0.67	0.59	0.67	0.70	42.14	0.36	56.93
	(0.09)	(0.16)	(0.19)	(6.98)	(0.25)	(0.21)	(0.20)	(0.17)	(8.62)	(0.17)	(21.79)
Literature & Fine Arts	0.93**	0.80*	0.69	20.44***	0.59	0.51	0.63	0.78	40.52	0.38	49.77
	(0.08)	(0.16)	(0.25)	(7.43)	(0.29)	(0.25)	(0.25)	(0.16)	(7.26)	(0.20)	(18.37)
Engineering	0.83	0.52	0.41	23.42***	0.83	0.84	0.61	0.66	46.19**	0.67	87.40
	(0.22)	(0.30)	(0.30)	(8.95)	(0.21)	(0.17)	(0.24)	(0.18)	(8.15)	(0.19)	(63.15)

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.
There are 69 observations for each cell (except for outcomes related to Engineering for which there are 66 observations).
*,**,***Mean is statistically significant from the corresponding mean for single major respondents at the 10%, 5%, and 1% levels, respectively.

Table 4: Comparing GPA and Income Beliefs with Objective Measures

	Average GPA[a]		Belief GPA $\geq 3.5^b$		Single Major GPA $\geq 3.5^c$		Average Salary[d]		Belief of Income at 30[e]		Single Major Income at 30[f]	
	Mean	**Rank[g]**	**Mean**	**Rank**	**Mean**	**Rank**	**Mean**	**Rank**	**Mean**	**Rank**	**Mean**	**Rank**
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
Natural Sciences	3.22	5	0.55	7	0.59	6	75.86	3	89.98	2	80.43	2
			(0.30)		(0.26)				(98.43)		(51.20)	
Math & Computer Sci	3.21	6	0.57	6	0.58	7	73.32	4	73.85	4	65.30	5
			(0.28)		(0.27)				(59.58)		(35.68)	
Social Sciences I	3.29	1	0.83	1	0.78**	1	72.73	5	64.34	5	58.68	6
			(0.15)		(0.19)				(31.08)		(25.88)	
Social Sciences II	3.09	8	0.68	5	0.63	5	78.10	2	108.14	1	118.9	1
			(0.24)		(0.26)				(117.95)		(152.8)	
Ethics and Values	3.29	1	0.81	3	0.76	2	68.23	6	63.16	6	67.40	4
			(0.16)		(0.20)				(28.37)		(49.78)	
Area Studies	3.29	1	0.83	1	0.76**	2	68.23	6	56.93	7	55.48	7
			(0.16)		(0.22)				(21.79)		(26.44)	
Literature & Fine Arts	3.29	1	0.80	4	0.75*	4	62.87	8	49.77	8	53.56	8
			(0.16)		(0.20)				(18.37)		(30.09)	
Engineering	3.11	7	0.52	8	0.48	8	89.26	1	87.40	3	80.03	3
			(0.30)		(0.26)				(63.15)		(46.81)	

Mean response of double and single major respondents is statistically significant at 1% (***), 5% (**), 10% (*) using a 2-tailed T-test.

[a]Mean GPA of bachelor graduates of Doctoral/Research Universities in 2001 (Source: 2001 Baccalaureate and Beyond Longitudinal Study).

[b]Belief of survey respondents with DOUBLE MAJORS about graduating with a GPA ≥ 3.5 (on a scale of 0-100) and divided by 100. (N= 69 except for Engineering which has 66 observations).

[c]Same belief as in *b* but of SINGLE MAJOR respondents (N= 83 except for Engineering for which there are 78 observations)

[d]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).

[e]Expected salary at the age of 30 (in 1000s) elicited from survey respondents with DOUBLE MAJORS.

[f]Same belief as in *e* but of SINGLE MAJOR respondents.

[g]Majors are ranked from highest GPA or expected salary (rank 1) to lowest (rank 8).

Table 5: Double Major Choice Model - Estimation Using Choice Data

Maximum Likelihood Estimates	(1)	(2)
Δu_{11} for graduating within 4 years	-2.38	-2.67
	(2.58)	(2.54)
Δu_{12} for maximum of graduating in 4 years	27.28***	30.36***
	(3.49)	(3.46)
Δu_{21} for graduating with a GPA of ≥3.5	0.73	0.14
	(2.38)	(2.14)
Δu_{22} for maximum of graduating with a GPA of ≥3.5	6.22**	6.64**
	(3.04)	(2.85)
Δu_{31} for enjoying the coursework	12.27***	12.53***
	(3.21)	(3.05)
Δu_{32} for maximum of enjoying coursework	1.25	1.75
	(3.68)	(3.42)
γ_{11} for hours/week spent on coursework	0.13***	0.13***
	(0.05)	(0.05)
γ_{12} for min. of hours/week on coursework	-0.15***	-0.17***
	(0.05)	(0.05)
Δu_{41} for approval of parents and family	12.85***	12.29***
	(2.42)	(2.27)
Δu_{42} for maximum of approval of parents	-2.52	-1.69
	(2.91)	(2.72)
Δu_{51} for finding a job upon graduation	-1.67	-2.20
	(2.01)	(1.93)
Δu_{52} for maximum of finding a job	3.74*	4.33**
	(2.12)	(2.19)
Δu_{61} for enjoying work at the jobs	7.72***	6.56***
	(2.86)	(1.25)
Δu_{62} for maximum of enjoying work at jobs	-0.45	-
	(2.91)	
Δu_{71} for reconciling family & work at jobs	-1.19	1.79*
	(2.93)	(1.07)
Δu_{72} for max of reconciling family & work	4.17	-
	(3.25)	
γ_{21} for hours/week spent at work	0.11**	0.05***
	(0.05)	(0.02)
γ_{22} for minimum of hrs/wk spent at work	-0.07	-
	(0.05)	
γ_{31} for the social status of the jobs	-0.77	0.25
	(2.42)	(1.16)
γ_{32} for maximum of social status of jobs	0.98	-
	(2.15)	
γ_{41} for expected income at the age of 30	5.190E-06	-1.16E-06
	(1.530E-05)	(3.20E-06)
γ_{42} for max of expected income at 30	-4.50E-06	-
	(9.90E-06)	
Log Likelihood	-132.77	-154.85
Number of Respondents	69	69

* significant at 10%; ** significant at 5%; *** significant at 1%

Status is on a scale of 0.1-0.9; job hrs/wk and coursework hrs/wk are on a scale of 0-100; income is in dollars; all other outcomes are on a normalized scale of 0-1.

Table 6: Extent of Specialization in Choice of Double Majors

Extent of Specialization	(1)	(2)
Graduating within 4 years	11.47***	11.35***
Graduating with a GPA of $\geq$3.5	8.46	47.81
Enjoying the coursework	0.10*	0.14*
Hours/week spent on coursework	1.13	1.28*
Approval of parents and family	0.20***	0.14***
Finding a job upon graduation	2.24	1.96*
Enjoying work at the jobs	0.06***	-
Reconciling family & work at jobs	3.51	-
Hours/week spent at work	0.64*	-
Social status of the jobs	1.28	-
Expected income at the age of 30	0.87	-

Each cell is $\left|\frac{\text{estimated parameter on max. of variable}}{\text{estimated parameter on variable}}\right|$. For example, the first cell is $\left|\frac{\Delta u_{12}}{\Delta u_{11}}\right|$.

Ratio statistically diff from 1 (* 10%; ** 5%; *** 1%) using a Wald test.

Table 7: Decomposition Analysis

	(1)	(2)	(3)	(4)
Attributed to:				
Outcomes in College				
Graduating in 4 Years	8.54%	11.13%	16.45%	14.74%
Graduating with a GPA of $\geq$3.5	8.42%	10.29%	12.11%	4.74%
Enjoying Coursework	19.27%	20.97%	12.97%	12.58%
Hours per Week spent on Coursework	9.38%	11.71%	7.16%	10.06%
Approval of Parents	19.39%	20.48%	11.63%	10.02%
Total	55.30%	71.82%	51.23%	43.16%
Outcomes in the Workplace				
Finding a Job Upon Graduation	7.89%	11.04%	7.79%	8.75%
Enjoying Work at Jobs	9.13%	7.38%	8.63%	10.60%
Reconciling Work and Family	5.97%	1.86%	7.41%	9.25%
Hours per Week Spent at Work	6.77%	3.82%	8.21%	8.46%
Social Status of Jobs	3.25%	0.79%	4.46%	7.64%
Expected Income at Age 30	1.99%	0.54%	3.17%	3.16%
Total	44.70%	28.18%	48.77%	56.84%

Each cell in the table reports the relative importance of the given determinant in explaining the choice. Refer to the text for the decomposition strategy that is used for this purpose.

Table 8: Double Major Choice Model with Error Components

Variables	(1)	(2)
Δu_{11} for graduating within 4 years	-3.31	-3.85
	(5.20)	(5.59)
Δu_{12} for maximum of graduating within 4 years	54.45**	55.78**
	(23.44)	(23.79)
Δu_{21} for graduating with a GPA of at least 3.5	0.23	-0.98
	(4.65)	(5.00)
Δu_{22} for maximum of graduating with a GPA of at least 3.5	10.61	12.79
	(7.71)	(8.75)
Δu_{31} for enjoying the coursework	21.86**	24.84**
	(10.26)	(10.87)
Δu_{32} for maximum of enjoying the coursework	3.94	4.07
	(7.30)	(7.14)
γ_{11} for hours/week spent on coursework	0.19	0.18
	(0.15)	(0.16)
γ_{12} for minimum of hours/week on coursework	-0.30*	-0.32**
	(0.17)	(0.15)
Δu_{41} for approval of parents and family	22.65**	23.81**
	(10.79)	(10.27)
Δu_{42} for maximum of approval of parents	-5.51	-5.86
	(5.52)	(5.72)
Δu_{51} for finding a job upon graduation	-3.91	-5.35
	(5.45)	(5.42)
Δu_{52} for maximum of finding a job	6.51	8.15*
	(4.92)	(4.84)
Δu_{61} for enjoying work at the available jobs	12.01*	10.08**
	(6.40)	(4.68)
Δu_{62} for maximum of enjoying work at jobs	-1.97	-
	(4.94)	
Δu_{71} for reconciling family and work at jobs	-1.05	3.08
	(5.02)	(4.33)
Δu_{72} for max of reconciling family & work	6.20	-
	(5.15)	
γ_{21} for hours/week spent at work	0.13	0.10
	(0.11)	(0.08)
γ_{22} for minimum of hours/week spent at work	-0.06	-
	(0.08)	
γ_{31} for the social status of the available jobs	1.49	1.36
	(4.22)	(2.48)
γ_{32} for maximum of social status of jobs	-0.46	-
	(3.16)	
γ_{41} for expected income at the age of 30	-1.22E-06	-4.30E-06
	(2.24E-05)	(1.00E-05)
γ_{42} for max of expected income at 30	-1.63E-06	-
	(1.31E-05)	
Log Likelihood	-131.39	-133.25
Number of Respondents	69	69

Estimation includes Error Components for each major category (estimates not shown).
* significant at 10%; ** significant at 5%; *** significant at 1%