

Wiswall, Matthew; Zafar, Basit

Working Paper

Preference for the workplace, investment in human capital, and gender

Staff Report, No. 767

Provided in Cooperation with:

Federal Reserve Bank of New York

Suggested Citation: Wiswall, Matthew; Zafar, Basit (2016) : Preference for the workplace, investment in human capital, and gender, Staff Report, No. 767, Federal Reserve Bank of New York, New York, NY

This Version is available at:

<http://hdl.handle.net/10419/146666>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Federal Reserve Bank of New York
Staff Reports

Preference for the Workplace, Investment in Human Capital, and Gender

Matthew Wiswall
Basit Zafar

Staff Report No. 767
March 2016



This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

Preference for the Workplace, Investment in Human Capital, and Gender

Matthew Wiswall and Basit Zafar

Federal Reserve Bank of New York Staff Reports, no. 767

March 2016

JEL classification: J24, J16

Abstract

In this paper, we use a hypothetical choice methodology to robustly estimate preferences for workplace attributes. Undergraduate students are presented with sets of jobs that vary in their attributes (such as earnings and job hours flexibility) and asked to state their probabilistic choices. We show that this method robustly identifies preferences for various job attributes, free from omitted variable bias and free from considering the equilibrium matching of workers to jobs. While there is substantial heterogeneity in preferences, we find that women, on average, have a higher willingness to pay for jobs with greater work flexibility (lower hours, and part-time option availability) and job stability (lower risk of job loss), while men have a higher willingness to pay for jobs with higher earnings growth. In the second part of the paper, using data on students' perceptions about the types of jobs that would be offered to them conditional on their college major choices, we relate these job attribute preferences to major choice. We find that students perceive jobs offered to humanities majors to have fewer hours, more worktime flexibility, and higher stability than jobs offered to economics/business majors. These job attributes are found to play a role in major choice, with women exhibiting greater sensitivity to nonpecuniary job attributes in major choice.

Key words: workplace preferences, compensating differentials, human capital, college majors, gender

Wiswall: Arizona State University (e-mail: matt.wiswall@gmail.com). Zafar: Federal Reserve Bank of New York (e-mail: basit.zafar@ny.frb.org). Ellen Fu provided excellent research assistance. The authors thank Joseph Altonji for feedback on the survey design, and participants at Harvard University, the Federal Reserve Bank of New York Micro Working Lunch, Princeton University, and the 2015 Econometric Society World Congress for valuable comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

Economists have long recognized that job and occupational choices are not solely determined by expected earnings.¹ While simple models based on earnings maximization abound (see for example the classic Roy (1951) model), and are quite useful in some applications, it is also clear that individuals have a rich set of preferences for various aspects of jobs beyond expected earnings, including earnings and dismissal risk, and various non-pecuniary aspects such as work hours flexibility and enjoyment of workplace activities. These preferences for various job attributes affect not only job choices, as individuals choose jobs not solely to maximize expected earnings, but also human capital investments, as individuals alter their human capital investment in anticipation of particular future job choices. And, by altering job choices and human capital accumulation, these workplace preferences ultimately shape the distribution of earnings and welfare in the economy.

However, empirically isolating the role of worker-side preferences for job attributes is difficult. One reason is that the equilibrium allocation of workers to jobs reflects not only the workers' preferences but the structure of the labor market and firm demands for workers. If one assumes the labor market is perfectly competitive with jobs of all types offered to all workers, then the equilibrium job choices and wages observed directly identify individual preferences—this is the classical model of compensating differentials (Rosen, 1984). If however firms have preferences for some types of workers and offer jobs only to a subset of workers with preferred characteristics, then the observed job choices do not reflect worker preferences only. Various kinds of labor market frictions which prevent workers from matching with their most preferred job types also break the direct connection between observed job choices and worker preferences. Even when the labor market is perfectly competitive, a second empirical challenge is that because jobs likely vary in many unobserved (to the researcher) characteristics, there exists a familiar omitted variable or selection bias problem in identifying worker preferences from realized job choices. If the observed characteristics in realized choice data are correlated with the unobserved characteristics, then our estimates of the importance of the observed characteristics to worker job choices are biased.

To address these empirical challenges, this paper presents a new methodology for esti-

¹Adam Smith writes “The five following are the principal circumstances which, so far as I have been able to observe, make up for a small pecuniary gain in some employments, and counterbalance a great one in others: first, the agreeableness or disagreeableness of the employments themselves; secondly, the easiness and cheapness, or the difficulty and expense of learning them; thirdly, the constancy or inconstancy of employment in them; fourthly, the small or great trust which must be reposed in those who exercise them; and, fifthly, the probability or improbability of success in them.” (Wealth of Nations, 1776, Book 1, Chapter 10).

inating individual preferences for workplace attributes and the importance of these preferences for human capital investments. We collect data on job attribute preferences through a survey in which we present undergraduate students with a series of hypothetical job choice scenarios and elicit their expected future choices across the jobs. The hypothetical job scenarios were constructed to offer students a realistic menu of potential jobs varying in expected earnings and other characteristics such as future earnings growth, dismissal probability, and work hours flexibility. The students' stated preferences for these jobs allows us to construct a "pure" measure of individual preferences for various job characteristics and estimate, in a simple and robust way, the distribution of their preferences for job attributes. In this way, our data isolates the preference for workplace attributes, free from considering the equilibrium job allocation mechanism, the preferences of employers, and whether an omitted variable bias problem exists in which unobserved job attributes are related to observed attributes.^{2,3} And, because our data collection in essence conducts a kind of "experiment" at the individual student level, the panel data generated by our design allows us to estimate the *distribution* of preferences allowing for unrestricted forms of preference heterogeneity. We combine this data on job attribute preferences with rich data on students' educational choices and expectations, including data on how students believe potential college majors relate to the jobs which would be available to them, and test whether the job preferences young adults hold in college in fact affect their human capital investments during college.

In our sample of high ability students, we estimate considerable heterogeneity in preferences for workplace attributes: students have preferences reflecting, on average, a dis-taste

²Our hypothetical choice methodology is a kind of "stated choice" analysis which shares some relationship to "conjoint analysis" and "contingent valuation" methods, used in fields including marketing and environmental and natural resource economics. Typically these methods are used to identify preferences for new, as yet unavailable, consumer products or for public goods like environmental quality, for which realized choices and markets do not exist. For examples in marketing and consumer choice, see Green and Srinivasan (1978), Beggs, Cardell, and Hausman (1981), Louviere and Woodworth (1983), Manski and Salomon (1987), and Ben-Akiva and Morikawa (1990). For examples, in environmental and resource economics see Smith (2004), Kling, Phaneuf, Zhao (2012), Carson (2012), and Hausman (2012). More recently Blass, Lach, and Manski (2010), Delavande and Manski (2015), Ameriks et al. (2015), and Fuster and Zafar (2015) have used this hypothetical choice methodology to analyze preferences for electricity reliability, political candidates, life saving, and housing demand, respectively. Like these previous studies, we exploit the possibilities that hypothetical data can provide rich variation in product characteristics, allowing for identification of preferences under weak assumptions about the form of preference heterogeneity. However, our primary motivation for hypothetical choice data is not because markets and realized choices do not exist, as is the issue in identifying preferences for new products or public goods, but to resolve problems of endogeneity of realized job choices.

³A natural question is whether preferences recovered from hypothetical choices measured in a survey match those recovered from actual real-world choices. Answering this would require both revealed choice data that is free of any confounds and stated choice data— data that are usually not available. However, the little evidence that exists shows a close correspondence between preferences recovered from the two approaches (see Hainmuller, Hangartner, and Yamamoto, 2015).

for higher job dismissal potential, and a taste for workplace hours flexibility. We estimate that on average students are willing to give up 2.8% of annual earnings for a job with a percentage point lower probability of job dismissal. The largest average WTP estimate is for the availability of a part-time hours option. Individuals, on average, are willing to give up 5.1% of their salary to have a job which offers the option of working part-time hours rather than one that does not offer this option. When dividing our sample by gender, we find that, on average, women have a higher preference for workplace hours flexibility, with an implied willingness-to-pay of 7.3% compared to 1% for men. Likewise, women have a higher WTP for more secure jobs- they are willing to give up 4% of their salary for a percentage point lower probability of job dismissal (versus a 0.6% WTP for males). On the other hand, men have a higher WTP for jobs with higher earnings growth: they are willing to give up 3.4% of annual earnings for a job with a percentage point higher earnings growth (the corresponding estimate for women is a statistically insignificant 0.6%). There is, however, substantial heterogeneity in preferences for workplace attributes, even within gender. In addition, the skewness of the estimated preference distributions is significantly different from zero for most attributes, suggesting that parametric distributional assumptions that impose symmetry in preferences are not supported in the data.

The substantial differences in willingness-to-pay for job amenities between men and women is consistent with prior work noting that the gender segregation of occupations and jobs is such that women are more likely to be found in jobs offering greater workplace flexibility (Goldin and Katz, 2011; Flabbi and Moro, 2012; Goldin, 2014). However, the observation that women tend to work in certain job types may not reveal women's preferences alone, but may also be affected by firm side demands for specific workers and discrimination or be driven by some other job attributes that are unobserved in our datasets (Blau and Kahn, 2006). Our finding of a substantial and gender-specific preference for job attributes such as work hours flexibility, even among college students, indicates that these types of preferences form even before entry into the labor market. These distinct preferences by gender suggest that part of the gender gap in earnings we observe is a compensating differential in which women are willing to give up higher earnings to obtain other job attributes.⁴

In the second part of the paper, we test whether the job preferences young adults hold in college in fact affect their human capital investments during college.⁵ We collect beliefs

⁴Gender differences in preferences may arise from innate differences (Baron-Cohen, 2003), and/or be a consequence of social factors including anticipated discrimination (Altonji and Blank, 1999). The question of understanding the sources of gender differences in preferences is an important one, but is beyond the scope of this paper.

⁵It is important to recognize that the preferences we identify are those for the students at the point in time at

about *expected* attributes of jobs students anticipate being offered if they were to complete particular majors. Importantly, these beliefs are elicited for not only the student's chosen major, but also the counterfactual majors. Following previous work on using expectations data to understand choices, we argue that perceptions or beliefs at the time of choice, whether accurate or biased in some way, are important to use as these beliefs are the basis of decision making.

We find that students believe that completing a major in the humanities would lead to being offered jobs with fewer work hours, greater work hours flexibility, lower job dismissal probability, and lower earnings uncertainty than if they were to complete majors in Economics/Business. Students believe that the jobs they would be offered if they completed a major in Economics/Business would, on the other hand, offer higher salaries and greater earnings growth. "Weighting" the perceived differences in non-pecuniary job attributes across fields using the estimated preferences over these attributes reveals that student's college major choices are, on the margin, affected by perceived differences in job attributes. When considering only the non-pecuniary job dimensions, we find that women need to be compensated in order to major in Economics/Business because it is perceived to lead to jobs with less desirable non-pecuniary attributes.

In order to quantify the importance of job attributes to major choice, we estimate a simple model of major choice where students receive utility from major-specific characteristics (such as perceived ability), job attributes conditional on major, and tastes.⁶ We compute "marginal effects" to gauge the importance of job attributes in major choice, whereby we vary each job attribute keeping the other job- and major- specific attributes fixed at their average values. We find that job attributes have a sizable impact on major choice. For example, increasing the job firing probability from 1% to 10% reduces the probability of majoring in the associated major by between 2% and 5%. To put this change in perspective, an increase of 10% in average earnings leads to a 6%-18% increase in the likelihood of majoring in that field for females (and of 15%-33% for males). Thus, for females, this

which we conduct the experiment. Preferences for jobs may change over time, and the preference we identify may not be those at the time of graduation and when students make job choices. They are the preferences at the time at which students are making important human capital decisions, one of the most important of which is the choice of college major.

⁶Prior research on college major choice examines the role of earnings expectations, ability perceptions, college costs, and tastes, but does not examine other job attributes. For examples of recent work, see Arcidiacono, 2004; Beffy, Fougere, and Maurel, 2012; Arcidiacono, Hotz, and Kang, 2012; Zafar, 2013; Stinebrickner and Stinebrickner, 2014a; Gemici and Wiswall, 2014; Wiswall and Zafar, 2015; Altonji, Arcidiacono, and Maurel, 2015. Arcidiacono et al. (2015), using expectations data from male undergraduate students about earnings in different major-occupation pairs, find evidence for complementarities in preferences between different majors and occupations, where occupations here include Health, Business, Government, and Education.

change is equivalent to nearly a third of the effect on major choice of increasing earnings by 10%. We find qualitatively similar impacts for the importance of other job attributes, such as work hours. In general, we find that females' major choices are more responsive to changes in non-pecuniary job attributes (relative to changes in earnings) than males, suggestive of their greater sensitivity to non-monetary aspects of the choice in human capital investment decisions.

Our results are consistent with burgeoning research on the substantial gender differences in job choices and human capital. For example, it has been noted that women tend to work in jobs with a higher degree of workplace hours flexibility, which may partially explain why women's average earnings are lower than men's, as women trade-off earnings for other workplace attributes (Goldin and Katz, 2011 and forthcoming; Blau and Kahn, 2012). Our innovation is to quantify the willingness-to-pay for job attributes using a flexible and robust methodology. Our findings also shed light on the drivers of occupational choice. While much of this literature traditionally has focused on earnings and earnings uncertainty, our results underscore the importance of other (non-pecuniary) workplace aspects in the choice. Finally, by linking these preferences directly to human capital investments, we contribute to our limited understanding of how career/workplace preferences shape educational choices.⁷

The paper is organized as follows. In the next section, we briefly motivate our analysis by using nationally representative surveys for the US to describe the distribution of occupation and college major choices and their associated job characteristics. Section 3 describes our data collection. Section 4 details the model of job choice, and shows how hypothetical data can solve important identification issues with realized choice data. Section 5 provides the empirical estimates of job preferences. Section 6, using data on students' beliefs about the types of jobs which would be offered to them conditional on their major choices, quantifies the importance of job attributes for college major choice. Finally, Section 7 concludes.

2 Background: Job Choices and Human Capital Investments in the United States

To set the stage for the analysis of our hypothetical choice scenario data, we first briefly describe the distribution of college majors, jobs, and associated job characteristics. To do so, we use two large sample representative datasets for the United States, the January 2010- De-

⁷Our findings are consistent with the finding in Arcidiancono et al. (2015) that non-monetary considerations related to future occupations are an important consideration in choosing a college major.

ember 2012 monthly Current Population Survey (CPS) and the 2013 American Community Survey (ACS).

Table 1 shows that the gender distribution across work sectors differs (Appendix A provides details on how variables in this table were constructed). While nearly two-thirds of women workers are in Health or Education, less than 40% of male workers are employed in these sectors. Business, the field most likely to be chosen by males, ranks third likeliest for females. These sectors differ substantially in their labor market returns: column (3) of Table 1 shows that average annual earnings of full-time workers are the lowest for Education, and highest for Science and Business. But these sectors differ along other dimensions as well: more than a quarter of the workers in Health and Education are employed part-time, possibly suggesting the amiability of these sectors to work hours flexibility. Job stability, as measured by the likelihood of being fired, is lowest in Government and Education, and highest in Business. Jobs in these sectors also differ in the skills that they demand of their workers. So what explains the propensity of men and women to work in different sectors—is it differences in preferences for workplace attributes, differences in tastes for occupations/industries, or differences in skills? Or is it a result of the labor market structure, firm labor demand, or discrimination by employers? This is something extremely challenging to answer with observational data.

We next turn to Table 2 to document the link between field of study and associated job characteristics.⁸ The table is based on the 2013 ACS, restricting the sample to 25-40 year olds with more than a high school education (at least some college). The first two columns show that while nearly a third of women have a Bachelor's degree in Humanities, only about a fifth of men do. And, while 12% of males have a Bachelor's in Engineering, the corresponding proportion for females is only 3%. We can reject the equality of the distribution of major choices by gender.

Column (3) of Table 2 shows that these majors differ significantly in their average earnings (as reflected by the F-test for the joint equality of means in the last row of the table). Engineering – the field which females are least likely to be present in – has the highest average earnings, while Humanities – the most popular Bachelor's field for females – has the lowest average earnings among the four Bachelor's fields. These majors, however, differ along other dimensions too. Columns (4) and (5) show that work hours flexibility is the highest for jobs associated with Humanities: 38.6% of all Humanities graduates are part-time workers, versus 21.7% of Engineering bachelor's graduates. Among the four Bachelor's fields,

⁸Altonji, Kahn, and Speer (2015) provide a more detailed discussion of the relationships between college majors and labor market outcomes.

average hours per week for full-time workers are also the lowest in Humanities. The last three columns of the table show that job stability and earnings growth also vary significantly across the fields of study.

With observational data, the challenge in identifying the determinants of field of study should then be clear. Females, for example, may be more likely to choose Humanities majors for any number of reasons: because they have a lower preference for earnings, they value the associated job attributes more, they have perceived or actual differences in abilities to complete coursework in this subject, or they have differences in “tastes” for studying humanities subjects, relative to other fields. In addition, the observed patterns of jobs we see in the data are equilibrium outcomes, and we cannot ascertain from this data alone whether these outcomes are due to worker demand or due to the supply of certain jobs—for example, part-time work may either be a voluntary or involuntary decision.

Our experimental approach, which we describe next, attempts to overcome these identification challenges.

3 Data

This section describes the administration of the data collection, the form of the hypothetical choice scenarios, and the sample we use for the estimation.

3.1 Administration

Our data is from an original survey instrument administered to New York University (NYU) undergraduate students over a 2-week period, during May 2012. NYU is a large, selective, private university located in New York City. The students were recruited from the email list used by the Center for Experimental Social Sciences (CESS) at NYU. Students were informed that the study consisted of some simple economic experiments and a survey about educational and career choices. Upon agreeing to participate, students could sign up for a 90-minute session, which was held in the CESS Computer Lab located on the main NYU campus.⁹

The data for this paper was collected through a computer-based survey (constructed using the SurveyMonkey software). The survey took approximately 30 minutes to complete, and

⁹During the same session, and immediately prior to completing the survey, students also took part in some economic experiments. Students also earned additional income through participation in the experiments. See Ernesto, Wiswall, and Zafar (forthcoming) for information on this data collection.

consisted of several parts. Many of the questions had built-in logical checks (e.g., percent chances of an exhaustive set of events such as majors had to sum to 100). Students were compensated \$10 as a show-up fee, and \$20 for successfully completing the survey.

3.2 Data Collection Instrument

In addition to questions about demographics, family background, and educational experiences, the main survey instrument consisted of two parts. The first part collected data on students' preferences for job attributes using hypothetical job choices, while the second collected data on consequential life activities that would plausibly be key determinants of college major choice, such as attributes of jobs associated with each major and measures of the student's perception of her ability to complete the coursework for each major. We describe the hypothetical job choice data in detail next, and leave the description of major-specific data to the second part of the paper where we relate the job attribute preferences to college major choices.

Our hypothetical job choice data were collected by presenting students with a total of 16 job scenarios. Each scenario consisted of 3 different potential jobs. We exogenously varied different aspects of the job with the intention of creating realistic variation in job attributes. The first 8 hypothetical job scenarios were introduced as follows:

In each of the 8 scenarios below, you will be shown hypothetical jobs offers.

Each job offer is characterized by:

Annual earnings when working full-time

Annual percentage increase in earnings from age 30 onwards until retirement

Full-time work hours per week

Work flexibility (whether part time work is an option); part time work is work where you only work at most half as many hours as full-time work and for half of the full-time salary

These jobs are otherwise identical in all other aspects.

Look forward to when you are 30 years old. You have been offered each of these jobs, and now have to decide which one to choose.

In each scenario, you will be asked for the percent chance (or chances out of 100) of choosing each of the alternatives. The chance of each alternative should

be a number between 0 and 100 and the chances given to the three alternatives should add up to 100.

Each scenario consisted of three jobs, with each job being characterized by four attributes. The notable point that was highlighted was that these jobs were identical in all other aspects. The last 8 scenarios were introduced in a similar way, except that the job offer was now characterized by a different set of attributes: annual earnings when working full-time; probability of being fired over a one year period; amount of additional annual bonus pay based on relative performance the respondent may qualify for (in addition to base pay); proportion of males in the firm in similar job positions.

Following the approach of Blass et al. (2010), we asked respondents to provide a choice probability instead of a discrete choice (that is, a zero or 1). This allows respondents to express uncertainty about their future behavior. It also allows individuals to rank their choices, providing more information than if we asked only about the most preferred job.

Besides earnings, the scenarios focus on six different job attributes. We chose to not vary these six dimensions all at once since the cognitive load to process such information could have been overwhelming. We focus on these dimensions based on findings from prior literature, and the fact that there is considerable variation along these dimensions across occupations as well as majors (Tables 1 and 2). Earnings and earnings growth were included since they have been found to be a factor in career/education choice (see Wiswall and Zafar, 2015, and references therein). Work hours and work flexibility are included since they tend to be associated with the remuneration structure in jobs and the associated gender gap in earnings (Flabbi and Moro, 2012; Goldin, 2014). We recognize that workplace flexibility is a multidimensional concept: for example, the number of hours to be worked matters but perhaps also the particular hours (Goldin, 2014). We varied two hours-related attributes: number of hours and the availability of a part-time option, since these are easy to vary in a meaningful fashion. Job stability, as proxied by the likelihood of being fired from the job, is included because of the importance of risk and uncertainty to job choices (Dillon, 2015) and gender differences in risk preferences (Croson and Gneezy, 2009). Finally, relative performance compensation and proportion of males are meant to capture the competitiveness of the job environment, preferences for which have been found to differ by gender (Niederle and Vesterlund, 2007; Reuben et al., 2015).¹⁰

To keep the scenarios realistic, the job attributes shown to respondents in the scenarios

¹⁰Lordan and Pischke (2015) find a strong relationship between female job satisfaction and the proportion of males in that occupation.

were based on the actual joint distribution of job characteristics in the Current Population Survey (except for the bonus pay variable, since data were not available for that dimension). In addition, no scenario included a job that was clearly dominant or dominated along all dimensions. We also made a conscious effort to keep the variation in job attributes *within* each scenario relatively "local", so that the claim that the jobs were otherwise identical was credible; for example, two jobs offering \$50,000 and \$90,000, respectively, with little variation along the specified dimensions are unlikely to be identical. At the same time, we had substantial variation in the job attributes *across* the scenarios. This ensures that we are not recovering preferences in a local region only.

While the job characteristics we provide are certainly not exhaustive of all possible job characteristics, and are purposely kept limited so as not to "overload" the respondents with too many job features, the key feature of the hypothetical experimental setting is that we instruct respondents that the jobs differ only in the finite number of job characteristics we provide, and are otherwise identical. There is no additional information provided that the respondent could use to believe otherwise.¹¹

3.3 Sample Description

A total of 257 students participated in the study. We drop 10 respondents for whom we have missing data for the relevant section of the survey. Sample characteristics are shown in Table 3. 35 percent of the sample (86 respondents) is male, 29 percent is white and 51 percent is Asian. The mean age of the respondents is 21.5, with 11 percent of respondents freshmen, 11 percent sophomores, 37 percent juniors, and the remaining seniors or higher. The average grade point average of our sample is 3.5 (on a 4.0 scale), and students have an average Scholastic Aptitude Test (SAT) math score of 696, and a verbal score of 674 (with a maximum score of 800). These correspond to the 93rd percentile of the US national population score distributions. Therefore, as expected, our sample represents a high ability group of college students. Parents' characteristics of the students also suggest that they are over-represented among high socioeconomic groups. The last panel of the table shows that 48 percent of the students have a major in the Humanities and Social Sciences category, 31 per-

¹¹This distinguishes our design from "audit" based studies in which employers are presented resumes which are otherwise identical except for the one chosen attribute (say the gender of applicant). The criticism of audit studies is that even if you make two groups (say men and women) identical on observables, employers might have very different distributions in mind about unobservable for the two groups, biasing the inference (for an analysis of this issue, see Neumark, Burn, and Button, 2015). In our case, students are instructed that the hypothetical jobs are identical in all other ways. Students may have different preferences for these job attributes, but we can identify this heterogeneity flexibly using our rich panel data.

cent have a major in Business and Economics, while the remaining have a major in Natural Sciences and Math (16%), and Engineering (5%). The gender composition of our sample compares favorably with that of the NYU undergraduate population: males constituted 33% of the graduating class of 2010 at NYU, and 34.8 percent of our sample.

Columns (2) and (3) of Table 3 report the characteristics by gender. That last column of the table reports the p-value of tests of equality of the statistics by gender. We see that male and female respondents are similar in all dimensions, except two. One, male students in our sample have a significantly higher average SAT Math score than females, of about 33 SAT points. Second, the two genders choose very different college majors. Nearly half – 49 percent – of males report majoring in Business/Economics, with 30 percent majoring in Humanities and Social Sciences, and 12 percent in Natural Sciences/Math. On the other hand, 57 percent of the females report majoring in Humanities and Social Sciences, followed by about 22 percent majoring in Business/Economics, and 18 percent majoring in Natural Sciences/Math. That is, female students are almost twice as likely as males to major in the Humanities (the field, which we show below, is perceived to have the lowest average earnings among college graduates), and only half as likely as males to major in the highest earnings major category, Economics/Business. The gender-specific major distributions are statistically different (p-value ≤ 0.001 , using a Chi-square test for equality of distributions). These substantial gender gaps in major choice mirror the national patterns from the ACS data (Table 2).

4 Model and Identification Analysis

In this section, we present a simple attribute-based job choice model and discuss identification of the model using two types of data: i) standard realized job choices (as observed after job choices are made), and ii) stated probabilistic job choices (as observed in our job hypothetical experimental data). We show that under weak conditions the job hypotheticals data identifies the distribution of job preferences, while standard realized job choice data does not. Later in the paper we present the model of college major choice and our framework for analyzing the relationship between these choices and job preferences.

4.1 A Canonical Random Utility Model of Job Choice

Jobs are indexed by j , and there is a finite set of jobs $j = 1, \dots, J$. Each job is characterized by a vector of K attributes $X_j = [X_{j1}, \dots, X_{jK}]$. These job attributes include earnings as

well as various non-pecuniary attributes, such as job dismissal probabilities and work hours flexibility. We explicitly allow for the possibility that individuals are not necessarily pure income or consumption maximizers, and may value many other outcomes associated with their job choice.

Let $U_{ij} \in R$ be individual i 's utility from job j . Following common practice, we assume the utility from job j has the additively separable form:

$$U_{ij} = u_i(X_j) + \epsilon_{ij}. \quad (1)$$

$u_i(X) \in R$ is the preferences of individual i over the vector of characteristics X . $\epsilon_{ij} \in R$ is the additional job-specific preference component for job j reflecting all remaining attributes of the job which affect utility, if any. Let ϵ_i be the vector of these components for individual i , $\epsilon_i = \epsilon_{i1}, \dots, \epsilon_{iJ}$. After observing the attributes X_1, \dots, X_J for all jobs and ϵ_i , individual i chooses the one job with the highest utility: i chooses job j if $U_{ij} > U_{ij'}$ for all $j' \neq j$.

Population preferences for jobs is the collection of u_i preferences over the job attributes X and the job-specific components ϵ_i . The joint distribution of preferences in the population is given by $F(u_i, \epsilon_i)$. This distribution determines the fraction of individuals choosing each job, $q_j \in [0, 1]$:

$$q_j = \int 1\{U_{ij} > U_{ij'} \text{ for all } j' \neq j\} dF(u_i, \epsilon_i), \quad (2)$$

4.2 Identification using Realized Choice Data

Almost without exception, empirical research on job choice consists of analyzing data on actual or realized job choices, which provides the one best job chosen by each individual.¹² Using this realized job choice data, we compute the fraction of the sample choosing each job q_j . In order to analyze the potential advantages of hypothetical data, we first detail the identification using realized choice data.

A common model of realized choice data assumes $\epsilon_{i1}, \dots, \epsilon_{iJ}$ are i.i.d. Type I extreme value, and independent of all of the $u_i(X_1), \dots, u_i(X_J)$ terms. Under these assumptions, we can write the population fractions as

¹²We confine attention to cross-sectional data. Panel data on repeated job choices over an individual's life-cycle may provide more identifying power but at the cost of requiring additional assumptions about the evolution of model features (e.g., preferences) as individuals age.

$$q_j = \int \frac{\exp(u_i(X_j))}{\sum_{j'=1}^J \exp(u_i(X_{j'}))} dG(u_i). \quad (3)$$

$G(u_i)$ is the distribution of preferences over attributes u_i in the population. (3) is the mixed multinomial logit model of McFadden and Train (2000). They show that the distributional assumption on the ϵ_i terms that yield the logit form is without loss of generality as this model can arbitrarily closely approximate a broad class of random utility models. For ease of exposition, we consider a linear model of utility given by $u_i(X) = X'\beta_i$. We next discuss the two main concerns in using realized choice data.

4.2.1 Omitted Variable Bias

One key concern in using realized job choices is that the dataset of job characteristics which the researcher has at hand is not complete in the sense that there are omitted unobserved job characteristics which are potentially correlated with the included observed characteristics. Divide the vector of job characteristics X into observed $X(obsv)$ and unobserved characteristics $X(unob)$, $X = [X(obsv), X(unob)]$. Similarly divide the vector of preference parameters $\beta_i = [\beta_i(obsv), \beta_i(unob)]$. The log odds of job j relative to job j' , using (3), is then:

$$\ln\left(\frac{q_j}{q_{j'}}\right) = (X_j(obsv) - X_{j'}(obsv))\beta_i(obsv) + (X_j(unob) - X_{j'}(unob))\beta_i(unob)$$

$$\ln\left(\frac{q_j}{q_{j'}}\right) = (X_j(obsv) - X_{j'}(obsv))\beta_i(obsv) + \eta_j.$$

The omitted variable bias problem is the generic one found in a variety of contexts. For example, if the researcher's dataset includes only current salaries, but not any of the non-pecuniary benefits of the job, we would expect that the estimate for preferences for salaries will be biased. The theory of compensating differentials (Rosen, 1984) predicts a close connection among various job characteristics – a tradeoff between salary and non-pecuniary benefits – and therefore would suggest important omitted variable bias in estimates of job preferences using realized data. The omitted variable bias issue could also arise more subtly from the selection/matching mechanism to jobs, reflecting employer preferences over potential job candidates. If the labor market equilibrium is such that employers only offer a limited set of jobs to candidates, then the realized jobs they hold therefore do not reflect

their preferences only. Taste discrimination by employers, by which employers prefer not to hire workers of certain groups (women, minorities. e.g.), is one example (Becker, 1957). In the presence of an important demand side aspect to job choice, one would not want to interpret the equilibrium allocation of jobs as reflecting only worker preferences.¹³ As we detail below, our hypothetical data avoids this because it experimentally manipulates the characteristics offered to individuals, thereby allowing a “pure” measure of preferences, free from considering the equilibrium job allocation mechanism, preferences of employers, or any omitted unobserved job characteristics.

4.2.2 Limited Distributional Flexibility

In addition to the omitted variable bias issue, the nature of the realized choice data limits the flexibility one can allow in the distribution of population preferences $G(u_i)$. The nature of realized choice data is that the data only represents what the individuals consider their most preferred job, not their preference ranking over all alternatives. Because of this limitation, a restricted parametric form for population preferences is required for point identification. For example, one might restrict $G(u_i)$ to assume all individuals have the same preferences across jobs: $u_i = u$ for all i and $G(u_i)$ is degenerate. Another possibility is to assume $G(u_i)$ is distributed Normal and then identification is of the mean and covariance elements. These type of parametric assumptions might bias the inference about the population distribution of preferences and distort predictions about the distribution of the demand for certain job characteristics. As we detail below, the hypothetical job choice data provides panel data for each individual in the sample, thereby allowing flexible and robust identification of population preferences.

4.3 Model of Hypothetical Job Choices

We next consider a framework for analyzing hypothetical job choice data, connecting the canonical model of realized job choice specified above in (1) with the hypothetical job choice data we collect. Our hypothetical data is asked prior to a job choice (while students are in school). We observe each individual’s beliefs about the *probability* they would take each future job offered within the scenario (and not simply the individual’s one chosen (realized) job). To analyze this type of data, we require a model of hypothetical future jobs. Our model

¹³We can represent demand side restrictions in jobs offered in the omitted variable framework by considering some unobservable job characteristic $X(\text{unobserv})$, such that $X(\text{unobserv}) \rightarrow -\infty$ if a job is not offered.

of hypothetical job choices presumes individuals are rational decision makers who anticipate the job choice structure as laid out in the canonical model of job choice (1). To allow for the possibility of uncertainty about future job choices, we assume that the realizations of $\epsilon_{i1}, \dots, \epsilon_{iJ}$ job-specific utility terms are not known at the time we elicit individual beliefs. Individual i then faces a choice among J hypothetical jobs with characteristics X_1, \dots, X_J . Each individual expresses their probability of taking a given job j as

$$p_{ij} = \int 1\{U_{ij} > U_{ij'} \text{ for all } j' \neq j\} dH_i(\epsilon_i), \quad (4)$$

where $H_i(\epsilon_i)$ is individual i 's *belief* about the distribution of $\epsilon_{i1}, \dots, \epsilon_{iJ}$ elements. As in Blass et al. (2010), the ϵ_{ij} has an interpretation as *resolvable uncertainty*, uncertainty at the time of our data collection but uncertainty that the individual knows will be resolved (i.e., known or realized) prior to making the job choice.¹⁴

It should be noted that the preferences for workplace attributes elicited in our data collection are potentially specific to the time at which the survey is collected (during the college years in our case). Preferences for job attributes may change as individuals age, and may be different when the students in our sample were younger (say prior to college) and different still when they actually enter the labor market and make job choices. With this caveat in mind, we can still use our research strategy to understand job preferences at a point in time and study how these preferences relate to important human capital investments, which are being made contemporaneously.¹⁵

4.4 Identification using Hypothetical Choice Data

We previously analyzed identification of preferences using realized job choice data and discussed two key shortcomings: realized choice data potentially suffers from omitted variable bias and limits the flexibility one can allow in the distribution of population preferences. First, with regard to the omitted variable bias, because we can experimentally manipulate the hypothetical choice scenarios we provide individuals, we avoid bias from the correla-

¹⁴An alternative model is that agents have uncertainty about preferences over attributes, that is the utility function $u_i(\cdot)$ is uncertain. For example, an individual may be uncertain about the number of children she may have at a future date, and the number of young children at home may affect her preference for workplace flexibility (an element of the X_j vector). We explore this later by relating preferences for job characteristics as revealed in our hypothetical data with a rich set of beliefs about future outcomes (e.g., individual beliefs about future own fertility and marriage).

¹⁵See Stinebrickner and Stinebrickner (2014a; 2014b) for evidence on the dynamics in beliefs formation among college students.

tion of observed and unobserved job characteristics from which realized choice data suffers. Rather than use naturally occurring variation in realized job choices—which are the result of many unobserved job characteristics and an unknown labor market equilibrium mechanism, as discussed above—we present individuals with an artificial set of job choices. While the job characteristics we provide are certainly not exhaustive of all possible job characteristics, and are purposely kept limited so as not to “overload” the respondents with too many job features, the key feature of the hypothetical experimental setting is that we instruct respondents that the jobs differ only in the finite number of job characteristics we provide, and are otherwise identical. There is no additional information provided that the respondent could use to believe otherwise. Under this particular design then, our hypothetical data is free from the possibility of omitted variable bias.

The second advantage of the hypothetical data is that hypothetical data provides a kind of panel data on preferences which, under fairly weak assumptions, identifies the full preference rankings over job attributes. Notice the key distinction between (4) and (2). Observed job choices q_j provide limited information on the job preferences and allow only parametric identification of the population distribution of preferences. With realized choices some parametric assumption on the joint distribution of preferences for job attributes must be imposed. With job hypotheticals data, on the other hand, we observe for each individual i multiple subjective job probabilities p_{i1}, \dots, p_{iJ} .¹⁶ The job hypotheticals provide a type of panel data allowing less restricted forms of identification, allowing identification of the $u_i(X)$ preferences without a parametric restriction on the population distribution of preferences.

The limited assumption for non-parametric identification preferences is that the $\epsilon_{i1}, \dots, \epsilon_{iJ}$ job specific terms are i.i.d. and independent of the job attributes X_1, \dots, X_J . This is guaranteed by the experimental design: respondents are instructed that the jobs vary only in the listed characteristics, and are otherwise identical. Under this assumption then the hypothetical data p_{i1}, \dots, p_{iJ} identifies the preference ranking for individual i over all jobs J in the choice set: for any two jobs, the characteristics vector X_j is preferred to that of $X_{j'}$ if the probability of choosing that job is higher than that for job j' : $p_{ij} > p_{ij'}$.¹⁷ Non-parametric identification of preferences over all X characteristics then follows from a sufficient number

¹⁶In order not to “overload” respondents with a very large set of choices, we break the different jobs into multiple scenarios, in which for each there are only 3 jobs to choose from. See the data description section.

¹⁷Proof: For jobs $j = 1, 2$, we can write the probabilities of choosing each job as

$$p_{i1} = \frac{1}{J} \mathbb{1}\{u_i(X_1) - u_i(X_{j'}) > \epsilon_{ij'} - \epsilon_{i1} \text{ for all } j' \neq 1\} dH_i(\epsilon_i)$$

$$p_{i2} = \frac{1}{J} \mathbb{1}\{u_i(X_2) - u_i(X_{j'}) > \epsilon_{ij'} - \epsilon_{i2} \text{ for all } j' \neq 2\} dH_i(\epsilon_i)$$

of choice scenarios so that the number of hypothetical job probabilities for each individual is sufficient given the dimensionality of the preference function u_i . That is, we need to have enough job scenarios to allow us to separately identify preferences for each element of the X characteristic vector. In the estimation, we use this identification result constructively and estimate preferences for each sample observation one-by-one. We then use the sample distribution of preferences as the sample estimator of the population distribution of preferences. Details on estimation are provided below.

5 Estimates of Preferences for Job Characteristics

5.1 Variation in Choice Probabilities

Identification relies, in part, on variation in probabilities that respondents assign to the various jobs in the hypothetical scenarios. We next present some evidence on this, which should allow the reader to become familiar with the sources of identifying variation. The top panel of Table 4 shows two examples from the data sample using the first set of hypothetical scenarios. Recall that each of these 8 scenarios included 3 different job offers, which differed according to the characteristics shown in the table. The last two columns of the table show the mean probability assigned by each gender to the jobs.

Turning to the first example, we see that, for males, Job 3 is the most preferred job in our sample (receiving the highest average probability of choosing this job), where Job 3 is the job without part-time availability and the highest earnings growth. For females, on the other hand, this job received the lowest average probability. Women assigned the highest probability, on average, to Job 2, the job with a part-time option and an intermediate number of work hours per week and intermediate earnings. In this example, the distribution of choices differs significantly by gender. The gender-specific distributions of average probabilities do not differ in the second example.

Panel B of Table 4 shows two examples from the second set of hypothetical scenarios, which vary a different set of attributes. In the first example, the distribution of average probabilities again differs by gender. For females, Job 1 receives the highest probability on average (37 percent). Job 1 is the job with the lowest probability of being fired, and the lowest proportion of men as colleagues. Male respondents, on the other hand, assign the

If $\epsilon_{i1}, \dots, \epsilon_{iJ}$ are i.i.d., then we can exchange $\epsilon_{ij'} - \epsilon_{i1}$ for $\epsilon_{ij'} - \epsilon_{i2}$. $p_{i1} > p_{i2}$ then indicates that $u_i(X_1) > u_i(X_2)$.

highest average probability to Job 3, the job with the highest earnings and proportion of men, but with a high likelihood of being fired.

Another notable aspect of Table 4 is the large standard deviation in elicited choice probabilities, reflective of substantial heterogeneity in choices, even within gender. Figure 1 shows the histogram of elicited percent chance responses for Job 1, pooled across the 16 hypothetical scenarios. Several things are of note. First, responses tend to be multiples of 10 or 5, a common feature of probabilistic belief data (Manski, 2004), reflecting a likely rounding bias; this is something that we return to below. Second, while there is pooling at multiples of 5, there is little evidence of excessive heaping at the standard focal responses of 0, 50, and 100. The most prevalent response is 20 percent, but even that receives a response frequency of only 0.11. Third, most respondents (87.5 percent) report values in the interior (that is, not zero or 1), reflecting a belief that there is some chance they might choose each of the jobs. This underscores the importance of eliciting probabilistic data, rather than simply the most preferred option, as respondents are able to provide meaningful probabilistic preferences for the full set of choices, revealing rankings of choices.

5.2 Empirical Model of Job Preferences

Next, we discuss our empirical model of job preferences, which we estimate using our hypothetical data. Our estimator here follows the identification analysis we laid out above. For the job preferences over attributes, we use the form $u_i(X) = X'\beta_i$, where $\beta_i = [\beta_{i1}, \dots, \beta_{iK}]$ is a K dimensional vector which reflects the individual i 's preferences for each of the K job characteristics. The X vector of job characteristics is described below and we consider several different functional forms. We assume beliefs about future job utility $H_i(\cdot)$ is i.i.d. Type I extreme value for all individuals. The probability of choosing each job is then

$$p_{ij} = \frac{\exp(X'_j\beta_i)}{\sum_{j'=1}^J \exp(X'_{j'}\beta_i)}, \quad (5)$$

where it is important to note that the probabilities assigned to each job j are individual i specific. While we maintain a particular assumption about the distribution of probabilistic beliefs, we place no parametric restrictions on the distribution of preferences, represented by the vector β_i . Our goal is to estimate the population distribution of preferences β_i . We maintain a maximum degree of flexibility by estimating the preference vector β_i separately for each sample member, and do not impose any "global" distributional assumptions about the population distribution of preferences (e.g. that preferences $\beta_i \sim N(\mu, \Sigma)$).

Applying the log-odds transformation to equation (5) yields the linear model:

$$\ln\left(\frac{p_{ij}}{p_{ij'}}\right) = (X_j - X_{j'})'\beta_i.$$

β_i has the interpretation of the marginal change in the log odds for some level difference in the X characteristics of the job. Given the difficulty of interpreting the β_i preference parameters directly, we also present results in which we compute individual level willingness-to-pay statistics.

5.3 Measurement Error

One potential issue in using hypothetical data for estimating preferences is that individuals may report their preferences with error. Given that these preferences have no objective counterpart (we cannot ascertain the “accuracy” of a self reported preference), we cannot point to definitive evidence on the extent of measurement error. The most apparent potential measurement issue is that individuals report rounded versions of their underlying preferences (rounded to units of 5 or 10 percent). To guard against the potential of rounding bias or other sources of measurement error, we follow Blass et al. (2010) in introducing measurement error to the model and in flexibly estimating the model using a least absolute deviations (LAD) estimator.

We assume that the actual reports of job choice probabilities in our data, denoted \tilde{p}_{ij} , measure the “true” probabilities p_{ij} with error. The measurement error takes a linear in logs form such that the reported log-odds take the following form:

$$\ln\left(\frac{\tilde{p}_{ij}}{\tilde{p}_{ij'}}\right) = (X_j - X_{j'})\beta_i + \omega_{ij}, \quad (6)$$

where ω_{ij} is the measurement error. We assume that the $\omega_{i1}, \dots, \omega_{iJ}$ are i.i.d. and have median zero, conditional on the X_1, \dots, X_J observed job characteristics. Given these measurement error assumptions, we have the following median restriction:

$$M\left[\ln\left(\frac{\tilde{p}_{ij}}{\tilde{p}_{ij'}}\right) \mid X_j, X_{j'}\right] = (X_j - X_{j'})\beta_i, \quad (7)$$

where $M[\cdot]$ is the median operator. This median restriction forms the basis for our estimator. Our measurement error assumptions are limited compared to commonly imposed fully parametric models which assume a full distribution for the measurement error process. In contrast, our assumption is that the measurement errors are only median unbiased.

5.4 Estimation

We estimate the K dimensional vector β_i by Least Absolute Deviation (LAD) for each student i separately. In our data, each student makes choices across 16 scenarios, assigning probabilities to 3 possible jobs in each scenario. Equation (7) therefore is estimated for each respondent using $16 \times 2 = 32$ unique observations. Variation in the job attributes ($X's$), which is manipulated exogenously by us, and variation in respondents' choice probabilities allows us to identify the parameter vector β_i . From the full set of estimates of β_1, \dots, β_N for our size N sample we estimate population statistics, such as mean preferences, $E(\beta_i)$. We conduct inference on the population statistics using block or cluster bootstrap by re-sampling (with replacement) the entire set of job hypothetical probabilities for each student. The block bootstrap preserves the dependence structure within each respondent's block of responses.

As discussed in the study design section, we varied 4 job attributes at a time in each scenario. For estimation, we combine all of these scenarios and assume the dimensions that were not varied in a given scenario were assumed by the respondent to be held constant, as we instructed. As mentioned earlier, we instruct respondents that the jobs differ only in the finite number of job characteristics we provide, and are otherwise identical. There is no additional information here that the respondent could use to believe otherwise. The vector of job attributes is as follows: $X = \{\log \text{ age 30 earnings; probability of being fired; bonus as a proportion of earnings; proportion of males in similar positions; annual increase in earnings; hours per week of work; availability of part-time}\}$.¹⁸ We also include job number dummies in equation (7) to allow for the possibility that the ordering of the jobs presented could affect job preferences, although there is no prior reason to suspect this given our experimental design.¹⁹

5.5 Job Preference Estimates

We first discuss the sign and statistical significance level of the β_i estimates. Because of the difficulty of interpreting the magnitude of these estimates, below we also present results in which we convert the parameter estimates into an individual-level willingness-to-pay measure.

The first column of Table 5 shows the average estimate for each job characteristic (averaged across all individual-level estimates). The standard errors in parentheses are derived

¹⁸We also estimate the model with the utility specified as linear in earnings (instead of log earnings). Results are qualitatively similar.

¹⁹This is related to the possibility of "session effects" in laboratory experiments. See Frechette (2012).

from a block bootstrap procedure. We see that the average estimates have the expected signs. With the exception of the probability of being fired and work hours per week, all other estimates are positive, as expected. The positive estimates indicate that individuals prefer jobs with these characteristics: individuals prefer higher salaries and work-time flexibility, and dislike jobs with a higher probability of being fired and high numbers of work hours. The only estimate that is not statistically or economically significant is the proportion of males at the job, indicating that we cannot reject that, on average, individuals are indifferent to the gender composition of the workplace. Turning to the average estimates by gender, reported in columns (2) and (3) of Table 5, we see similar qualitative patterns in terms of magnitudes. We return to the differences in magnitudes of the preferences by gender below, and also provide a willingness-to-pay interpretation.

An advantage of our approach is that we can identify the β_i vector without a parametric restriction on the population distribution of preferences. This allows us to flexibly estimate the distribution of population preferences. Table 6 shows various statistics of the estimated distribution of preferences. For brevity, the table does not present the bootstrap standard errors, but the precision of the estimates, derived from the bootstrap procedure, is denoted by asterisks.

There is substantial heterogeneity in preferences within gender. Take, for example, the probability of being fired. The median and 25th percentile of the individual-specific estimates is negative, indicating a distaste for a higher likelihood of job dismissal, and is statistically different from zero for both males and females. The 75th percentile is negative for both genders, but only statistically significant for females, indicative of a greater distaste for job instability among female respondents. Underscoring the heterogeneity in preferences, we see that the standard deviation of the estimates of preferences, for all job characteristics, are sizable and statistically different from zero. Notably, we also see that the skewness of the estimates is sizable and statistically different from zero in all cases (except for proportion of males at jobs). This indicates that we can reject that the individual parameter estimates are symmetric around the mean, and that the assumption that the parameter estimates are distributed Normal, as is commonly assumed when estimating heterogeneous preferences using standard revealed choice data, is not supported in this sample.

5.6 Willingness-to-Pay

The parameter estimates in Table 5 are difficult to interpret given the necessarily non-linear nature of the model. To ease interpretation, we next present willingness-to-pay (WTP) es-

timates, by translating the differences of utility levels into earnings that would make the student indifferent between giving up earnings and experiencing the outcome considered.

Willingness-to-pay (WTP) to experience job attribute X_k is constructed as follows. Consider a change in the level of attribute X_k from value $X_k = x_k$ to $X_k = x_k + \Delta$, with $\Delta > 0$. X_k is “bad” attribute (e.g., probability of job dismissal). Given our linear utility function, we can write an indifference condition in terms of earnings Y as

$$x_k \beta_{ik} + \beta_{i1} \ln(Y) = \beta_{ik}(x_k + \Delta) + \beta_{i1} \ln(Y + \text{WTP}_{ik}(\Delta))$$

where Y is the level of earnings, one of the job attributes included in every job scenario. $\text{WTP}_{ik}(\Delta) > 0$ is individual i ’s willingness to pay to avoid increasing the “bad” attribute k by Δ . Solving, WTP is given by

$$\text{WTP}_{ik}(\Delta) = \left[\exp\left(\frac{-\beta_{ik}}{\beta_{i1}} \Delta\right) - 1 \right] \times Y. \quad (8)$$

WTP for individual i depends on her preference for the attribute β_{ik} versus her preference for earnings β_{i1} . Given that we allow for a log form to utility in earnings (allowing for diminishing marginal utility in earnings and implicitly consumption), willingness-to-pay for an individual also depends on the level of earnings.

Table 7 shows the average WTP estimates for changing each of the job characteristics by one unit (for the probabilistic outcomes, this is increasing the likelihood by 1 percentage point; for hours per week, increasing it by an hour; for part-time availability, this is going from a job with no part-time option to one which does).²⁰ The first three columns of the table present the estimates in dollars, using the average annual earnings across all scenarios, which in this case is \$75,854.17 and does not vary across respondents. The last three columns show the estimates as a proportion of the average earnings. We focus on the latter here.

We see, for example, that increasing the likelihood of being fired by 1 percentage point, that is, $X_k = x_k + 1$, would yield an average WTP of 2.8% for the full sample. That is, students on average would have to be compensated by 2.8% of annual earnings if job stability were to be decreased, for them to remain indifferent. The gender-specific averages, reported in the last two columns of Table 7, paint a very interesting picture. Females, on average, have to be compensated by 4% of average earnings for a unit increase in the likelihood of being fired (with the estimate being statistically significant at the 1% level), and statistically different from the male average of 0.60% (which is indistinguishable from zero).

²⁰The WTP is computed for each individual, using the individual-specific β_i estimates. The table reports the mean WTP across respondents, with bootstrap standard errors in parentheses.

The average WTP estimate for the availability of the part-time option is sizable. Individuals, on average, would have to be compensated by -5.1% of their annual salary (that is, they are willing to give up 5.1%) when going from a job with no part-time option to one that does. The estimate is driven by the female respondents in the sample, for whom the average WTP is -7.3%, versus -1.0% for males (with the male estimate not being statistically different from zero). The much higher average preference among women for the part-time option is statistically significant from zero and statistically different from the male average, at the 5 percent level.

Looking at the other estimates, we see that the average WTP for annual earnings growth is statistically precise for males, who are willing to give up 3.4% of average annual earnings for a 1 percentage point increase in earnings growth; the female coefficient is also negative but indistinguishable from zero (though not statistically different from the male estimate). We see that females have a stronger distaste for the number of hours of work, with the average WTP indicating that they need to be compensated by 1.3% of annual earnings for an increase of 1 hour in the work week; the male estimate is not precise (but we cannot reject the two gender-specific averages being equal). Both genders are willing to give up 0.8-1.7% of annual earnings for a percentage point increase in bonus compensation (in addition to base salary).²¹ Finally, the average WTP for proportion of men at jobs is economically and statistically insignificant.²²

5.6.1 Heterogeneity in Willingness-to-Pay

Table 7 reports the average WTPs only. As seen in Table 6, there is substantial heterogeneity in preferences for workplace characteristics, and overlap in the male and female distributions of preferences.

The job scenarios provided to respondents are incomplete scenarios, that is, respondents are given only a subset of the information they would have in actual choice settings. Differ-

²¹That the WTP for a percentage point increase in bonus is greater than 1 in magnitude for females is surprising, since it implies that women are on average willing to give up more in base salary to gain a smaller increase in bonus compensation. This is driven by a few outliers. In fact, we cannot reject that the mean WTP for women is different from either -1 (that is, a one-to-one substitution between base pay and bonus pay), or from the mean of -0.8 for male respondents.

²²The utility from jobs, specified in equation (1), is linear and separable in outcomes. We also estimate a variant of this model which allows for interactions between certain job attributes. The value of part-time flexibility to an individual may depend on the number of work hours at the job. Likewise, the desirability of performance-based bonus may depend on the gender composition of the workplace or job stability. To allow for this possibility, we include interactions of these terms in equation (1). The WTP estimates that we obtain (which are evaluated at the average value of these attributes) are qualitatively similar. These estimates are available from the authors upon request.

ences in estimated preferences then may reflect true underlying heterogeneity in preferences, or may be driven in part by differences in how students think about the unspecified dimensions when answering these questions. For example, our finding that male respondents on average have a higher WTP for earnings growth may be a result of males assigning a higher value to earnings growth over the lifecycle, or females expecting a shorter tenure at the jobs than men. Likewise, an individual may be uncertain about the number of children she may have at a future date, and the number of young children at home may affect her preference for workplace flexibility. Similarly cognitive biases such as the exponential growth bias (Stango and Zinman, 2009) may lead certain respondents to underestimate the implications of a given earnings growth rate for earnings over the lifecycle. It is important to note that these cross-sectional differences in processing of information or in perceptions regarding the unspecified dimensions have implications for understanding the heterogeneity in the estimated preferences; the process of preference estimation itself is not biased by this heterogeneity since the estimation uses the panel of choices to produce estimates for each individual separately. We next investigate how the WTPs are associated with various individual-level characteristics.²³

Table 8 shows the average WTP for these attributes (expressed as a percentage of average earnings) for various sub-samples. Parents' income and race seem to be the only demographic variables that are systematically related to the heterogeneity in WTP. We see that students from households with below sample median income (in our sample, \$87,500) are more sensitive to workplace characteristics: their average WTP for job stability (probability of being fired), work hours, and part-time availability is significantly larger than that of their counterparts.²⁴ Nonwhite students are also more sensitive to certain workplace characteristics, in particular those related to work flexibility (work hours per week and part-time availability).

Notably, student ability (as measured by SAT scores) is not systematically related to the average WTPs, as would have been the case if differences in preferences were driven by differences in respondent ability to comprehend the scenarios. We see some evidence of sorting into majors based on job preferences: students who assign a likelihood of 50 percent

²³An alternative is to make the scenario more complete by specifying the job conditional on a number of characteristics, such as tenure at the job and the individual's household structure (number of children, etc.). There is a trade-off—by making the situation more specific and stylized, the scenario may become unrealistic from the individual's perspective.

²⁴To the extent that the student's wealth is increasing in their parent's income, one would expect a lower willingness to pay for non-pecuniary aspects of a job for lower-income students (as the marginal utility of labor earnings is higher for low-wealth households). We see the opposite pattern here. One hypothesis consistent with our empirical finding is that students with lower-income parents expect less parental support (to provide childcare, for example) and therefore have a higher willingness-to-pay for job hours flexibility.

or less to majoring in Economics/Business (71% of the sample), on average, value job flexibility (part-time availability and work hours) and job stability more than their counterparts; however, only the difference in WTP for part-time availability is statistically different.

We also examine whether preferences for workplace characteristics differ by the student's expected future household composition (whether they expect to be married and their expected number of children) and labor supply. Notably, the likelihood of being married by age 30 and the expected number of children by age 30 – data that we collect directly from respondents – are not significant correlates of the willingness to pay; that is the case even when we look within gender. Respondents who assign a higher likelihood (80 percent or more) to working full-time at age 30, on average, value earnings growth at the job more (and part-time availability less) than their counterparts: for example, they are willing to give up 2.7% of age 30 earnings for a percentage point increase in earnings growth versus an average of 1.4% for their counterparts (these differences are not statistically significant).

To understand the extent to which these individual-level correlates drive the underlying heterogeneity in WTP, we conduct multivariate linear regressions of the WTP onto these covariates (results available from the authors upon request). The R-squared of these regressions indicate that at most 10% of the variation in the WTP can be explained by these individual correlates. Even after including these controls, the gender difference in the WTP for part-time availability and probability of being fired continues to be significant. WTP for non-pecuniary job attributes, while quantitatively meaningful and displaying a distinct gender difference, is not well explained by standard demographic variables (other than gender). Overall, this suggests that these types of preferences are difficult to “control for” by simply conditioning on these types of variables. In addition, the fact that observables explain a small part of the variation in the WTPs indicates that the heterogeneity in preferences is largely a result of true underlying variation in preferences, and not driven by differences across individuals in how they answer and perceive these hypothetical scenarios.

6 Job Preferences and Major Choice

The preceding section used a robust hypothetical choice methodology to estimate individual level preferences for various job attributes. The estimates reveal important heterogeneity in preferences, with a substantial mass of individuals having a large willingness-to-pay in foregone earnings for non-pecuniary characteristics such as a low probability of job dismissal and work-time flexibility. This next section relates these preferences for job attributes to

human capital investments by using these preferences to understand the importance of job characteristics to college major choices.

6.1 Major Choice Probabilities

We first discuss the data we have available on expected future major choices. The first column of Table 9 shows the response to the question: "*What do you believe is the percent chance (or chances out of 100) that you would either graduate from NYU with a PRIMARY major in the following major categories or that you would never graduate/dropout (i.e., you will never receive a Bachelor's degree from NYU or any other university)?*" Because of time constraints, we aggregated the various college majors to 5 groups: 1) Business and Economics, 2) Engineering and Computer Science, 3) Humanities and Other Social Sciences, 4) Natural Sciences and Math, and 5) Never Graduate/Drop Out.²⁵ The most likely major for males is Economics/Business (43 percent), followed by Humanities/Social Sciences (29 percent). For females, on the other hand, the most likely major is Humanities/Social Sciences (53 percent), followed by Economics (23 percent). The probability of not graduating is less than 3 percent for both genders. The average probabilities assigned to the majors differ significantly by gender for all majors except Engineering and Natural Sciences.

6.2 Perceptions of Major-specific Job Attributes

Our goal is to understand the extent to which students' choice of college major is driven by the characteristics of jobs available to them conditional on graduating with a given major. Therefore, we also collected data from respondents on their perceptions of characteristics of the jobs that would likely be offered to them *if* they were to complete each type of major. An important characteristic of our dataset is that we gather students' beliefs about workplace characteristics (such as likelihood of being fired, earnings) for all major categories (Business and Economics, Humanities and other Social Sciences, Natural Sciences and Math, and Never graduate/Drop out), and not just for the one major they have chosen. Descriptive statistics for these questions are shown in Table 9. The top and bottom panels show the statistics for the male and female respondents, respectively. The questions we asked were

²⁵We provided the respondents a link where they could see a detailed listing of college majors (taken from various NYU sources), which described how each of the NYU college majors maps into our aggregate major categories. Before the official survey began, survey respondents were first required to answer a few simple practice questions in order to familiarize themselves with the format of the questions.

prefaced to think ahead about the future labor market when they are 30 years old and take into account any advanced degrees (beyond their undergraduate major) they might complete.

6.2.1 Earnings Beliefs

We start with student beliefs about the future earnings they would receive after completing each major. Age 30 earnings beliefs were elicited as follows: "*If you received a Bachelor's degree in each of the following major categories and you were working FULL TIME when you are 30 years old, what do you believe is the average amount that you would earn per year?*". The second column of Table 9 shows that both genders expect average earnings to be the highest in Economics/Business, followed by Engineering, and then Natural Sciences. Humanities are expected to have the lowest average earnings among the graduating majors, with the average earnings expected to be less than two-thirds of the average conditional on graduating in Economics. The mean beliefs reported by males are significantly higher than those reported for females for each of the five fields. The large standard deviations, however, indicate there is considerable heterogeneity in beliefs. Column (3) shows the perceived earnings growth at the jobs.²⁶ While there is no clear trend, we reject the null that the average growth is the same across the five majors (as shown by the p-value of the F-test in the last row of each panel).

6.2.2 Other Job Attributes Beliefs

Columns (4)-(6) of Table 9 describe various measures of a major's perceived competitiveness. Respondents were asked: (1) the probability of being fired, (2) bonus pay based on relative performance, as percent of annual base pay, and (3) the fraction of male employees, for jobs offers they expect to receive at age 30 conditional on college major.²⁷ We see that, among graduating majors, both male and female students expect jobs in Economics/Business to generally be the most competitive- it is the major category with the highest perceived av-

²⁶Perceived earnings growth is backed out from the age 30 and age 45 full-time expected earnings, assuming a constant growth rate.

²⁷The precise wording of the questions is: (1) "What do you believe would be the percent chance of being fired or laid off in the next year from positions similar to those from which you would receive job offers at age 30 if you received a Bachelor's degree in each of the following major categories?", (2) "What do you believe would be the average amount of bonus pay based on relative performance (as a percent of your annual base pay) among the job offers you receive at age 30 if you received a Bachelor's degree in each of the following major categories?", and (3) "What do you believe would be the proportion of men in positions similar to those from which you would receive job offers at age 30 if you received a Bachelor's degree in each of the following major categories?".

erage ratio of bonus pay and probability of being fired. Engineering jobs are, on average, perceived to be closer to those in Economics/Business, along the dimension of bonus pay and proportion of males. Jobs in Humanities are perceived to be the least competitive according to the ratio of bonus pay and the fraction of male employees, while jobs in Natural Sciences are considered the safest.

There is substantial heterogeneity in these beliefs as reflected by the large standard deviations. The top panel of Figure 2 shows the distribution of female respondents' beliefs regarding the likelihood of part-time availability in jobs conditional on graduating in Humanities and in Economics/Business. We see there is substantial variation in beliefs within a given major, as well as across majors. In addition, the belief distribution conditional on Humanities first order stochastically dominates the distribution conditional on Economics. For example, nearly 60 percent of the female students assign a probability of more than 40% to the likelihood of part-time availability at jobs available conditional on graduating in Humanities. On the other hand, less than a quarter of students assign a probability of more than 40% to this outcome in the case of jobs in Economics.

Columns (7)-(8) of Table 9 show that, among graduating majors, jobs in Humanities are on average perceived to have the lowest hours and highest hours flexibility, by both male and female respondents. Economics/Business is perceived to be the most demanding in terms of work hours and lack of work flexibility, followed by Engineering. Jobs in Natural Science are perceived to be in between.

6.2.3 Gender Differences

The patterns in Table 9 show clearly that both males and females perceive very different workplace characteristics conditional on college major. For each attribute in the table, we reject the equality of mean beliefs being the same across majors, for the two genders. Comparing the two panels, we see that male and female respondents generally have similar relative beliefs regarding the attributes conditional on major, but the levels are quite different in several cases. For example, females assign a higher probability of being fired and a higher likelihood of part-time availability for all majors, compared to their male counterparts. This gender difference in beliefs is highlighted in the lower panel of Figure 2, which shows the gender-specific belief distributions for part-time availability for jobs in Humanities. While 60 percent of females assign a probability of at least 40% to the likelihood of part-time availability in Humanities, less than 40 percent of males do so. This could be a result of either male and female respondents applying to different sectors, and/or them expecting to receive

offers from different industries, even conditional on the same broad major category.

6.2.4 "Accuracy" of Perceptions

One might wonder about the accuracy of these expectations. We cannot speak to that because, by construction, these outcomes would be realized only in the future and, importantly, would be observed only for the chosen major. It is important to note that it is the *expected* job attributes, as perceived by the respondent at the time of choosing a college major, that matter in the choice decision. Whether these perceptions are biased is then an irrelevant question from the perspective of understanding the decision. Systematic biases in expectations do, however, imply a policy case for information interventions, something that is not the subject of the current study (interested readers should look at Wiswall and Zafar, 2015, and references therein).

A comparison of students' expected job attributes conditional on college major with those of current college-graduate workers (reported in Table 2) is not very informative. The two may differ for many reasons: (1) our sample consists of high ability students at a selective private university, whose graduates disproportionately enter the financial sector, so it is not clear that the ACS sample is the correct reference group; (2) students may have private information about themselves that may justify having perceptions that differ from current realizations; or (3) the distribution of realizations may not be stationary, and so future outcomes may differ from past realizations. With these caveats in mind, a comparison of students' perceptions in Table 9 with realizations of current workers (in Table 2) indicates that, on average, the students' relative ranking of majors in terms of perceived workplace characteristics, particularly job hours, earnings, and work flexibility are consistent with the realized job data from the ACS, suggesting that these beliefs are in some sense reasonable.

6.2.5 Sources of Different Workplace Attributes Perceptions

A natural question that arises is where do the differences in perceptions of workplace characteristics conditional on major come from. One plausible explanation is that there is some mapping of majors to industries/occupations. Given that industries/occupations seem to differ significantly in their workplace characteristics (as shown in Table 1), this would then result in different perceptions of job attributes conditional on major. In a separate set of questions, we also asked our respondents: "*What do you believe is the percent chance that you would be working in the following [Science/Technology; Health; Business; Government/Non-profit; Education] at age 30 if you received a Bachelor's degree in each of the following?*"

Appendix Table A1 shows the mean belief of working in each sector conditional on major, by gender. Several things are of note. One, the perceived probability of entering the different sectors varies substantially by college major (the p-values reported in the last row of the table reject the equality of the average likelihood of being in a given sector across the majors). Second, certain majors seem to be more closely tied to certain sectors: for example, the perceived probability of working in Business (Science) exceeds 50 percent, conditional on graduating in Economics (Engineering). Third, it is certainly not the case that there is a one-to-one mapping of majors to sectors. None of the majors are concentrated in one or two sectors. The mean probability exceeds 5 percent for all sectors, for each of the majors. Fourth, there is little systematic difference by gender in the perceived mapping of majors to sectors. Arcidiacono et al. (2015), in their survey of Duke undergraduate students, find similar patterns regarding the perceived mapping of majors to occupations.

And, finally, there is substantial dispersion across students in the perceived mapping, as reflected by the large standard deviations. Thus, it seems that the different mapping of majors to workplace characteristics that we documented in Table 9 is driven, at least in part, by students' perceived link between majors and sector of work.

6.3 Willingness-to-Pay for Majors

How important are the perceived differences in job attributes by major to the choice of college major? Given the individual-specific WTP estimates for each job attribute and the perceived mapping of majors to workplace characteristics, we can calculate each student's willingness-to-pay for each major relative to a baseline major (in our case, Humanities):

$$WTP_{im} = \sum_{k=1}^K WTP_{i,k} * (X_{i,m,k} - X_{i,humanities,k}), \quad (9)$$

where $WTP_{i,k}$ is i 's WTP for job attribute k in dollars as estimated in section 5.6. $X_{i,m,k}$ is i 's *perceived* level of job attribute X_k in major m (e.g., likelihood of being offered part-time work hours if the student were to complete major m).²⁸ A positive WTP value would indicate that i needs to be compensated to major in m (relative to Humanities); a negative value would imply that the student is willing to pay to major in m instead of Humanities. The WTP in (9) is in dollars. We instead present them as a percent of age 30 earnings that

²⁸With hopefully minimal confusion, we use the same notation X to refer to job attributes as in our hypothetical analysis and to perceptions about job attributes associated with each major. Note that here the $X_{i,m,k}$ are indexed by i as these attributes are each student's *perception* of the job attributes rather than the exogenously determined attributes in the hypotheticals we created.

the individual expects to make (age 30 earnings is the weighted average of the individual's age 30 major-specific expected earnings, with the weights being the subjective probability of majoring in each of the majors).

We first compute the WTP in equation (9) restricting the vector of job attributes to non-pecuniary dimensions (that is, part-time availability, job hours, probability of being fired, and proportion of males). Estimates are presented in the first 2 columns of Table 10. The mean estimate is positive and large for the dropout/not graduate category. Male and female respondents would, on average, have to be compensated by 16% and 32% of age 30 earnings to choose the dropout option instead of Humanities. For males, the mean WTP is negative for both Economics/Business and Engineering, suggesting that on average they are willing to pay to major in them instead of Humanities. For females, the mean (median) estimate for Economics/Business indicates that they need to be compensated by 3% (7%) of age 30 earnings to major in Economics instead of Humanities, reflecting a disutility from the kinds of jobs available in Economics/Business in terms of non-pecuniary dimensions. The female mean WTP estimates for Engineering and Natural Sciences indicate that females are willing to give up 16-20% of age 30 earnings to major in these fields, indicative of the attractiveness of jobs in these fields along these dimensions. The large standard deviations indicate there is substantial heterogeneity (in preferences and perceptions) for these dimensions within gender. Another notable aspect is that the magnitude of the mean WTPs tends to be larger for females than for males (except in Economics), suggestive of females being more sensitive to this dimensions. Appendix Table A2 shows the WTPs for each job attribute, instead of pooling them together as done here. Columns (1)-(4) of the table show that the mean WTPs for work flexibility (part-time availability and work hours) are positive for both genders for the three graduating majors. That is, individuals, on average, need to be compensated to major in these fields if these were the only job dimensions that mattered. This is intuitive since these job attributes are desirable, and are perceived to be worse in the non-Humanities fields. On the other hand, the mean WTP for the probability of being fired for the non-Humanities majors is negative for both genders (columns 1a and 1b).

The last two columns Table 10 show the WTP estimates considering all the job-specific attributes. In addition to the job attributes in the first two columns, we add in bonus and earnings growth. Now, with the exception of the dropout option, the mean WTP is negative for all majors for both genders. That is, the perceived job attributes and associated WTP is such that, on average, males and females are willing to give up money to not major in Humanities. For example, the average (median) female is willing to give up 41% (5%) of

age 30 earnings to major in Economics instead of Humanities. Appendix Table A2, which disaggregates the WTP by each attribute, shows that these negative estimates are primarily driven by two attributes: probability of being fired, and bonus pay as a fraction of base salary.

6.4 Model of Major Choice

We provide a second test for the importance of job attributes to major choice by directly estimating a model of major choice incorporating our flexible estimates of preferences for job attributes, and using this estimated model to quantify the importance of each job attribute. We start with a simple framework in which we suppose that utility for student i from major m is given by:

$$V_{im} = X'_{im}\alpha_i + Z'_{im}\gamma + \eta_{im}, \quad (10)$$

where Z_{im} is a vector of other major-specific characteristics perceived by student i , including a major-specific constant.²⁹ η_{im} captures the remaining unobservable attributes of each major, unobservable to the econometrician but observable to the student.³⁰

The student-specific preference for each job attribute is given by the vector $\alpha_i = [\alpha_{i1}, \dots, \alpha_{iK}]$. α_i , the preference for job characteristics as it relates to the utility from each major, is potentially distinct from the preferences for job characteristics in the job choice problem, given by β_i (in equation 5). Job characteristics, such as earnings at the job, may be quite important when choosing among different job offers, but might have a more limited value to choosing majors, relative to other major characteristics given by Z_{im} and η_{im} . To allow for this possibility, for each job characteristic k , we specify that each α_{ik} is proportional to the β_{ik} up to some free parameter δ :

$$\alpha_{ik} = \beta_{ik}\delta.$$

δ indicates the importance of job attributes to major choice, relative to other determinants of college major as given by Z_{im} and η_{im} . We expect $\delta \geq 0$, indicating that job characteristics have weakly the same direction of relationship with major choice as with utility specifically about jobs. Given that there are two choice problems here, one for job choice directly and

²⁹This vector consists of student i 's perceived major-specific ability (on a 1-100 scale, where 1 is the highest ability) and hours of study required to attain a GPA of 4.0. According to the two measures, both males and females consider Engineering the most difficult, and Humanities the least difficult major category. Economics/Business and Natural Sciences fare somewhere between the two. These two measures are summarized in the last two columns of Table 9.

³⁰We restrict the choice set to graduating majors, that is, $m = \{\text{Economics/Business; Engineering; Humanities: Natural Sciences/Math}\}$, since beliefs about study hours do not apply to the not graduate major.

one for college major, δ plays the role of providing the mapping between the two levels of utility, which in general need not have the same scale. In particular, δ can reflect standard discounting given that the utility from working at jobs occurs later in life than utility derived from taking courses while in school.

Under the assumption that the random terms $\eta_{i1}, \dots, \eta_{iM}$ in equation (10) are independent and identically distributed Type 1 extreme value across individuals and majors m , the probability that student i chooses major m is:

$$q_{im} = \frac{\exp(X'_{im}\alpha_i + Z'_{im}\gamma)}{\sum_{m'=\{1,\dots,M\}} \exp(X'_{im'}\alpha_i + Z'_{im'}\gamma)}. \quad (11)$$

As before, once we apply the log-odds transformation, we have:

$$\ln\left(\frac{q_{im}}{q_{im'}}\right) = (X_{im} - X_{im'})\alpha_i + (Z_{im} - Z_{im'})\gamma + \omega_{im}, \quad (12)$$

where ω_{im} is the error due to rounding. We continue to assume that the $\omega_{i1}, \dots, \omega_{iM}$ are i.i.d. and have median zero, conditional on X and Z .

6.4.1 Estimation

Estimation of the major preferences proceeds in two steps. In the first step, using equation (7), we estimate the job characteristic preference vector β_i for each individual. Call the estimate $\hat{\beta}_i$. We then create an individual- and major- specific scalar of weighted job characteristics for each major m :

$$B_{im} = X'_{im}\hat{\beta}_i. \quad (13)$$

The second step of the estimator is as before, where we use the LAD estimator. However, in this case, we use a pooled estimator over the whole sample and estimate δ and the vector γ .

$$M \left[\ln\left(\frac{q_{im}}{q_{im'}}\right) \mid X \right] = (B_{im} - B_{im'})\delta + (Z_{im} - Z_{im'})'\gamma. \quad (14)$$

As with the job preferences estimation, we use a cluster bootstrap for inference.

Table 11 presents the LAD estimates of equation (14). The estimate of δ is positive and precise, indicating that the preferences of students over job attributes and the major-specific beliefs about the distribution of job attributes has a statistically significant relationship with

major choices. This is consistent with the non-zero WTP for major estimates we presented previously. Estimates on the ability measures are negative, as one would expect (note that higher rank is a lower ability rank with our measures). The major-specific dummy terms are all negative, indicative of negative median tastes for the non-Humanities majors.

6.4.2 Importance of Job Attributes in Major Choice

Given the non-linear nature of the model, it is difficult to assess the importance of job attributes in major choice. To quantify the effects, we use standard methods to evaluate "marginal effects" in non-linear models. The marginal effect of job attribute x_j in major choice is computed by varying the value of that job attribute, while keeping the other job and major-specific attributes and preferences fixed at their sample average values. The likelihood of majoring in m , for a given value of $X_k = x_k$ and evaluated at the sample mean for the other attributes and preferences, is given as:

$$\overline{q}_m(x_k) = \frac{\exp(x_k \overline{\alpha}_k + \overline{X}'_{-km} \overline{\alpha} + \overline{Z}'_m \gamma)}{\exp(x_k \overline{\alpha}_k + \overline{X}'_{-km} \overline{\alpha} + \overline{Z}'_m \gamma) + \sum_{m' \neq m} \exp(\overline{X}'_{m'} \overline{\alpha} + \overline{Z}'_{m'} \gamma)}, \quad (15)$$

where $\overline{\alpha}_k$ is the sample average preference for attribute k , \overline{X}'_{-km} is the vector of job attributes excluding k (at the sample mean for each of the attributes in major m), and \overline{Z}'_m is the average major-specific beliefs for m . The effect for a given attribute is then obtained by computing (15) at two distinct values of that attribute. For example, we estimate the "marginal effect" for job firing probability by varying it from 1% to 10% (keeping the preference parameters and other job and major-specific beliefs at the sample average), and computing the percent change in the predicted probability of majoring in that major. Note that the marginal effect for a given variable may vary by major since the averages for perceived job attributes and ability measures are major-specific.

Table 12 presents the marginal effects for specific changes in job attributes, for each major, and separately by gender (in the two panels of the table).³¹ Column (1), for example, shows that increasing the probability from 1% to 10% of being fired from jobs associated with a major decreases the likelihood of majoring in that major by between 2.2%-4.9%

³¹Marginal effects are computed for: a change from 1% to 10% in the probability of being fired; no part-time availability to part-time availability; change in work hours from 30 hrs/week to 50 hrs/week; a change from 1% to 10% in bonus as percentage of base pay; a change from 1% to 10% in earnings growth; a change from 30% male colleagues to 70%, and; a change of 10% in age 30 earnings. Gender-specific sample averages are used for the other beliefs/preferences.

for males, and by between 1.9 and 5.7% for females. Part-time availability increases the probability of completing a major from between 0.5 to 1.5 percent, on average. Column (3) shows that increasing weekly work hours from 30 to 50 reduces the likelihood of majoring in the associated major by between 3.5%-7.7% for males, and by 1.5-4.2% for females. Bonus pay and earnings growth both have sizable marginal effects. The last column of Table 12 shows the percent change in the major probability for a 10% increase in age 30 earnings.

A few things are of note in the table. First, the effects are typically smaller (larger) for Humanities (Engineering). This is partly a result of Humanities (Engineering) having a high (low) baseline probability for both genders in the sample. Second, a comparison of the effects in the first three columns with those in the last column for earnings gives a sense of the relative importance of non-pecuniary job attributes in major choice. We see that, for females, the effects for job hours and probability of being fired are nearly a third of the effects for earnings. For males, the relative impacts are smaller (though still sizable). Overall, this indicates that, at the margin, job attributes do matter in major choice, and that they are more relevant for females' choices.

7 Conclusion

The contribution of this paper is two-fold. First, using a novel hypothetical job choice framework that experimentally varies different dimensions of the workplace, we are able to robustly estimate individual preferences for workplace attributes. Second, these workplace preferences combined with unique data on students' perceived mapping of college majors to workplace characteristics allow us to investigate the role of anticipated future job characteristics – particularly non-pecuniary aspects of these jobs – in major choice, a key human capital investment decision.

We document substantial heterogeneity in willingness-to-pay for job amenities both within and across genders. On average, females have a stronger taste for workplace flexibility (as measured by hours and part-time availability) and job stability. Males, on average, have a greater willingness to pay for jobs with higher earnings growth. Our results suggest that part of the gender gap in earnings we observe is a compensating differential in which women are willing to give up higher earnings to obtain other job attributes. Because students perceive systematic differences in attributes of job offers conditional on college major, and because job attributes other than earnings are also valued, we find that job preferences matter in college major choice. In addition, females are much more sensitive to non-pecuniary job aspects

in major choice.

On the methodological front, we are able to recover the *true* individual preferences for various job characteristics, since our methodology is free from considering the equilibrium job allocation mechanism, the preference of employers, and the concern that jobs may differ along unobservable dimensions (see Blau and Kahn, 2006). In addition, given that we have multiple revealed choices for each participating student, we can use the panel structure of the data to construct a preference vector over the attributes student-by-student. The richness of the experimental data allows us to estimate the distribution of preferences while imposing minimal assumptions on the parametric distribution of preferences (such as imposing joint Normality). These types of parametric distributional assumptions would be required in general for identifying preferences from discrete realized job choices. Importantly, our analysis reveals substantial skewness in the distribution of workplace preferences, indicating that workplace preferences are not symmetric in the sample. While we find substantial variation in workplace preferences in our sample of high ability students enrolled at a selective private US college, it is not clear how this heterogeneity compares to that in the broader population. It would clearly be useful to follow our design and collect similar data in other settings.

Given that prior literature on educational choice finds that the residual unobserved taste component is the dominant factor (Arcidiacono, 2004; Beffy et al., 2001; Gemici and Wiswall, 2014; Wiswall and Zafar, 2015), our approach can be viewed as trying to get into the black box of tastes by directly incorporating certain non-pecuniary dimensions into choice models. We believe the approach used in this paper illustrates the potential of using such methods to understand the determinants of human capital and occupational choice and the underlying causes of the gender gap.

References

- [1] Altonji, Joseph, and Rebecca Blank. 1999. Race and Gender in the Labor Market. In *Handbook of Labor Economics*, Volume 3c, eds. Orley Ashenfelter and David Card, 3144-3259. Elsevier Science.
- [2] Altonji, Joseph, Peter Arcidiacono, and Arnaud Maurel. 2015. The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects. *Handbook of the Economics of Education* Volume 5, Chapter 7.
- [3] Altonji, Joseph, Lisa Kahn, and Jamin Speer. 2015. Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success. *Journal of Labor Economics*, Forthcoming.
- [4] Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew Shapiro, and Christopher Tonetti. 2015. Long-term Care Utility and Late in Life Saving. Working Paper.
- [5] Arcidiacono, Peter, Joseph Hotz, and Songman Kang. 2012. Modeling College Major Choices using Elicited Measures of Expectations and Counterfactuals. *Journal of Econometrics*, 166(1): 3-16.
- [6] Arcidiacono, Peter, Joseph Hotz, Arnaud Maurel, and Teresa Romano. 2015. Recovering Ex Ante Returns and Preferences for Occupations Using Subjective Expectations Data. Working Paper.
- [7] Baron-Cohen, Simon. 2003. *The Essential Difference: Men, Women, and the Extreme Male Brain*. London: Allen Lane.
- [8] Becker, Gary. 1957 [1971]. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- [9] Beffy, M., D. Fougere, and A. Maurel. 2012. "Choosing the Field of Study in Post-Secondary Education: Do Expected Earnings Matter?" *The Review of Economics and Statistics*, 94(1): 334–347.
- [10] Beggs, S., S. Cardell, and J. Hausman. 1981. Assessing the Potential Demand for Electric Cars. *Journal of Econometrics*, 16(1): 1-19.
- [11] Ben-Akiva, M. and T. Morikawa. 1990. Estimation of Switching Models from Revealed Preferences and Stated Intentions. *Transportation Research A*, 24A, 485-495.

- [12] Bertrand, Marianne, Claudia Goldin, and Lawrence Katz. 2010. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2: 228-255.
- [13] Blass, Asher, Saul Lach, and Charles Manski. 2010. Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability. *International Economic Review*, 51(2): 421-440.
- [14] Blau, Francine, and Lawrence Kahn. 2006. The US Gender Pay Gap in the 1990s: Slowing Convergence. *Industrial and Labor Relations Review*. 60(1):45-66.
- [15] Blau, Francine, and Lawrence Kahn. 2012. The sources of the gender pay gap. In Grusky, David and Tamar Kricheli-Katz, editors, *The New Gilded Age: The Critical Inequality Debates of our Time*, pages 843-860. Stanford University Press, Stanford, CA.
- [16] Carson, Richard T. 2012. "Contingent Valuation: A Practical Alternative When Prices Aren't Available." *Journal of Economic Perspectives*, 26(4): 27-42.
- [17] Croson, Rachel, and Uri Gneezy. 2009. Gender Differences in Preferences. *Journal of Economic Literature*, 47(2): 448-74.
- [18] Delavande, Adeline, and Charles Manski. 2015. Using Elicited Choice Probabilities in Hypothetical Elections to Study Decisions to Vote. *Electoral Studies*, 38, 28-37.
- [19] Dillon, Eleanor. 2015. Risk and Return Tradeoffs in Lifetime Earnings. Working Paper.
- [20] Flabbi, Luca, and Andrea Moro. 2012. The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model. *Journal of Econometrics*, 168: 81-95.
- [21] Fréchette, Guillaume. 2012. Session-Effects in the Laboratory. *Experimental Economics*, 15(3): 485-498.
- [22] Fuster, Andreas, and Basit Zafar. 2015. The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey. Federal Reserve Bank of New York Staff Report no. 702.

- [23] Gemici, Ahu, and Matthew Wiswall. 2014. "Evolution of Gender Differences in Post-Secondary Human Capital Investments: College Majors at the Intensive Margin." *International Economic Review*, 55 (1): 23-56.
- [24] Goldin, Claudia. 2014. A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4): 1091-1119.
- [25] Goldin, Claudia, and Lawrence Katz. 2011. The Cost of Workplace Flexibility for High-Powered Professionals. *The Annals of the American Academy of Political and Social Science*, 638(1):45-67.
- [26] Goldin Claudia, and Lawrence Katz. Forthcoming. "A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation," *Journal of Labor Economics*, forthcoming.
- [27] Green, Paule E., and V. Srinivasan. 1978. Conjoint Analysis in Consumer Research: Issues and Outlook. *Journal of Consumer Research*, 5 (2) (Sep., 1978): 103-123
- [28] Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112(8): 2395-2400.
- [29] Hausman, Jerry. 2012. Contingent Valuation: From Dubious to Hopeless. *Journal of Economic Perspectives*, 26(4): 43-56.
- [30] Kerry Smith, V. 2006. Fifty Years of Contingent Valuation. In A. Alberini and J.R. Kahn (eds), *Handbook on Contingent Valuation*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, pp. 7-65.
- [31] Kling, Catherine L., Daniel J. Phaneuf, and Jinhua Zhao. 2012. From Exxon to BP: Has Some Number Become Better Than No Number? *Journal of Economic Perspectives*, 26(4): 3-26.
- [32] Lordan, Grace, and Jorn-Steffen Pischke. 2015. Does Rosie Like Riveting? Male and Female Occupational Choices. Working Paper.
- [33] Louviere, Jordan, and George Woodworth. 1983. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. *Journal of Marketing Research*, 20(4): 350-367.

- [34] Manski, Charles, and Ilan Salomon. 1987. The demand for teleshopping. An application of discrete choice models. *Regional Science and Urban Economics*, 17(1):109-121.
- [35] Manski, Charles. 2004. Measuring Expectations. *Econometrica*, 72(5): 1329-1376.
- [36] McFadden, Daniel, and Kenneth Train. 2000. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, Vol. 15(5): 447-470.
- [37] Neumark, David, Ian Burn, and Patrick Button. 2015. Is It Harder for Older Workers to Find Jobs? New and Improved Evidence from a Field Experiment. NBER Working Paper.
- [38] Niederle, Muriel, and Lise Vesterlund. 2007. Do Women Shy away from Competition? Do Men Compete too Much? *Quarterly Journal of Economics*, 122(3): 1067-1101.
- [39] Reuben, Ernesto, Matthew Wiswall, and Basit Zafar. Forthcoming. Preferences and Biases in Educational Choices and Labor Market Expectations: Shrinking the Black Box of Gender. *Economic Journal*, forthcoming.
- [40] Rosen, Sherwin. 1986. The theory of equalizing differences," *Handbook of Labor Economics*, Vol. 1, pp. 641-692.
- [41] Roy, A. D. 1951. Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, New Series, 3(2): 135-146.
- [42] Siow, Aloysius. 1984. Occupational Choice under Uncertainty. *Quarterly Journal of Economics*, 52(3), 631-45.
- [43] Stango, Victor, and Jon Zinman. 2009. Exponential Growth Bias and Household Finance. *Journal of Finance* 64(6), 2807-2849.
- [44] Stinebrickner, Ralph and Todd Stinebrickner. 2014a. A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *The Review of Economic Studies*. 81(1): 426-472.
- [45] Stinebrickner, Todd R. and Ralph Stinebrickner. 2014b. Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model. *Journal of Labor Economic*, 32(3), 601-644.

- [46] Wiswall, Matthew, and Basit Zafar. 2015a. Determinants of College Major Choice: Identification Using an Information Experiment. *Review of Economic Studies*. 82(2): 791-824.
- [47] Wiswall, Matthew, and Basit Zafar. 2015b. How do College Students Respond to Public Information about Earnings? *Journal of Human Capital*, 9(2): 117-169.
- [48] Zafar, Basit. 2013. College Major Choice and the Gender Gap. *Journal of Human Resources*, 48(3): 545-595.

Appendix A: Summary for Table 1 variables

Table 1 was generated using data from the Current Population Survey (CPS). Statistics regarding job attributes of each sector were computed using pooled monthly samples from January 2010 to December 2012.

We classified job occupations in the CPS into our five groups of focus, Science, Health, Business, Government, and Education. Our sector category definitions are the following:

- *Science*: Architectural, engineering, and related services; Computer systems design and related services; Scientific research and development services; Administration of economic programs and space research.

- *Health*: Offices of physicians; Offices of dentists; Offices of chiropractors; Offices of optometrists; Offices of other health practitioners; Outpatient care centers; Home health care services; Other health care services; Hospitals; Nursing care facilities; Residential care facilities, without nursing. These occupations are grouped in the CPS as “Health Care and Social Assistance”.

- *Business*: Banking and related activities; Savings institutions, including credit unions; Non-depository credit and related activities; Securities, commodities, funds, trusts, and other financial investments; Insurance carriers and related activities; Real estate; Accounting, tax preparation, bookkeeping, and payroll services; Management, scientific, and technical consulting services; Advertising and related services; Management of companies and enterprises.

- *Government*: Executive offices and legislative bodies; Other general government and support; Justice, public order, and safety activities; Administration of human resource programs; Administration of environmental quality and housing programs; National security and international affairs. These occupations are grouped in the CPS as “Public Administration”

- *Education*: Elementary and secondary schools; Colleges and universities, including junior colleges; Business, technical, and trade schools and training; Other schools, instruction, and educational services. These occupations are grouped in the CPS as “Educational Services”.

To construct our job attributes of interest, we restricted the 2010-2012 CPS sample to those with at least a Bachelor’s degree, and between the ages of 25 and 60. The sample consists of 385,201 individuals. When computing the statistics, we used the CPS sampling weights.

- **Prop. of Part-time workers**: We define a worker to be full-time if he/she was

“employed full time”, “full time for economic reason”, or “full time for non-economic reason”. The percent of people working part time within each industry is obtained by dividing the total number of part time people by the total number of employed workers in that industry. We calculate the percent of part-time workers for each month, and report the average over the 36 months.

- **Yearly firing rate:** To construct the yearly probability of being fired in each sector, we computed the monthly probability of being fired. We first flagged all workers who are laid off in a given month but have been unemployed for less than one month. Their sum is divided by the total number of (part and full time) employed workers at the beginning of that month, giving us an estimate of the monthly probability of being fired. The yearly probability was then computed by multiplying the monthly probability by 12. We compute this yearly probability for each of the 36 monthly surveys and report the average.

When constructing the following variables, we further restricted the CPS sample to those employed and working full-time.

- **Hours/week worked:** We used the number of hours actually worked in the last week (as opposed to the number of hours usually worked) – this was available for all workers who were employed in the last week. We chose to drop the top and bottom 1%. We calculate this for each survey month, and report the average over the 36 months.

- **Annual earnings:** The nominal annual earnings were calculated from the weekly earnings variable. This variable included overtime pay, tips and commissions, and before taxes or other deductions. We dropped the bottom 1% of weekly earnings. We also dropped observations where the reported weekly earnings were less than the number of hours they actually worked in the past week times the federal minimum wage (\$7.25/hour). Weekly earnings are multiplied by 52 to get the annual earnings. We compute this for each survey month, and report the average over the 36 months.

- **Annual % raise in earnings:** This is constructed by computing the percent increase in annual average earnings in a given industry from 2010 and 2011, as well as from 2011 and 2012. The raise of each sector is the average for the two years.

- **Proportion of male workers:** For each survey month, we calculated the proportion of full-time workers employed in each sector who are male. We compute this for each survey month, and report the average over the 36 months.

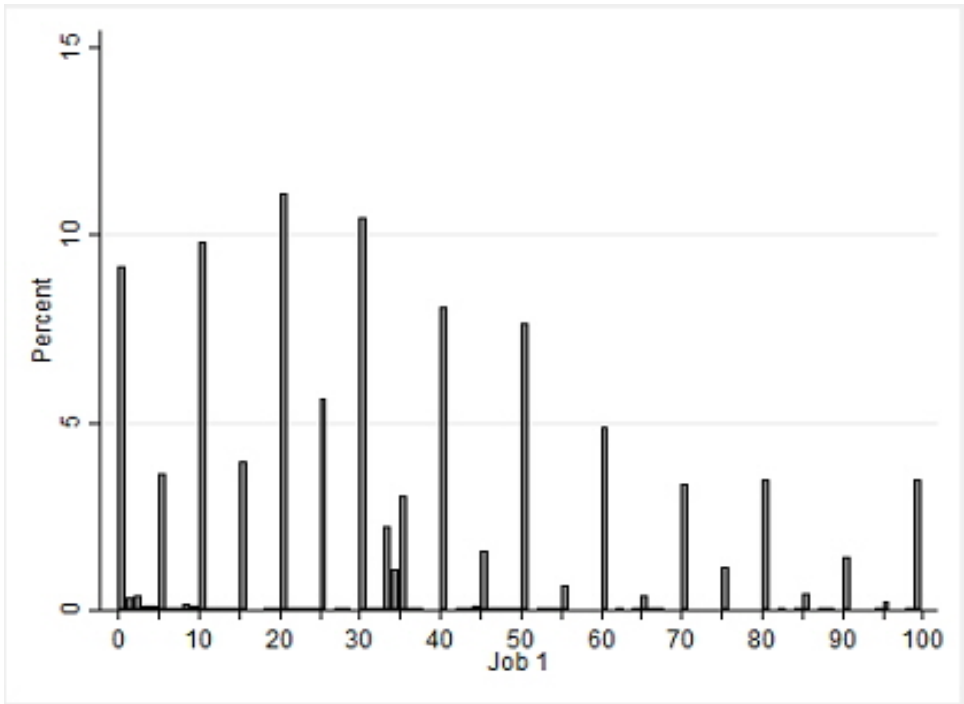
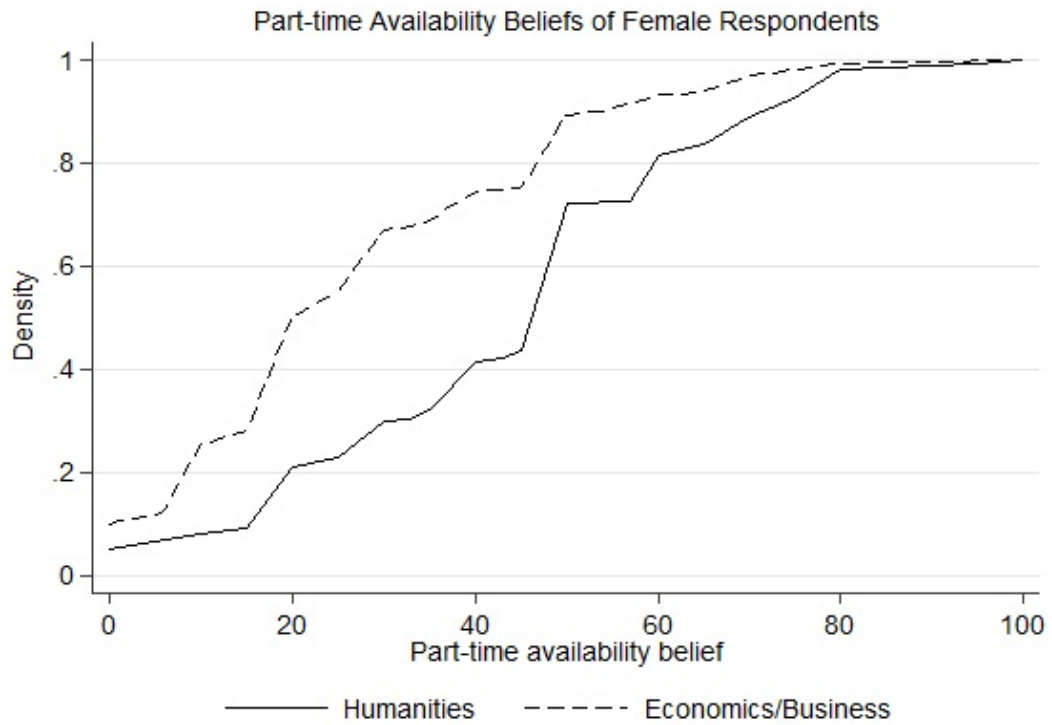
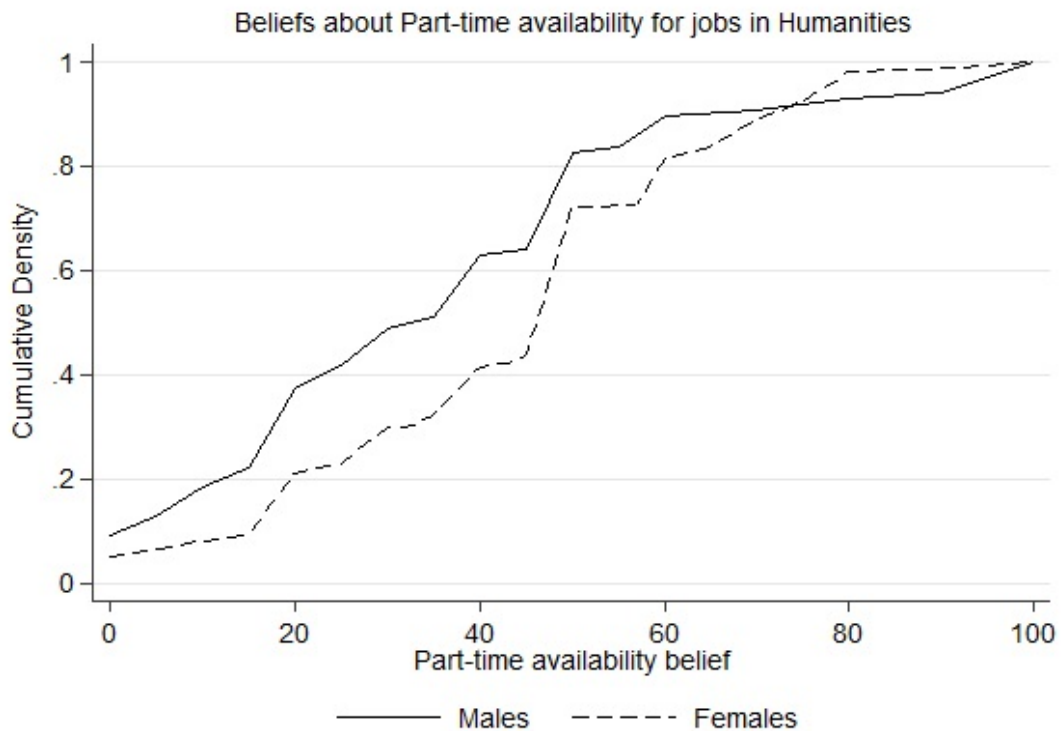


Figure 1: Choice probabilities for Job 1 (pooled across hypothetical scenarios)



Pane A: Female respondents' beliefs about part-time availability in jobs



Panel B: Part-time availability beliefs for jobs in Humanities

Table 1: Job Attributes by Sector, based on the CPS

	% of males working in ^c (1)	% of females working in (2)	Annual earnings for full-time workers (3)	Hrs/wk for full-time workers (4)	Prop. of part-time workers (5)	Yearly firing rate (6)	Prop. male workers ^b (7)	Annual % raise in earnings (8)
Science	19.4	6.0	82743 (35993)	44.21 (7.17)	0.16 (0.03)	3.67% (1.68)	70.1% (1.08)	2.06
Health	15.9	28.9	67876 (35751)	43.76 (7.82)	0.27 (0.01)	3.85% (0.82)	28.7% (0.53)	1.17
Business	28.7	17.4	77071 (39020)	44.97 (7.78)	0.20 (0.02)	4.06% (1.34)	54.6% (0.79)	0.25
Government	13.2	8.4	67530 (32299)	43.31 (6.86)	0.16 (0.05)	1.46% (0.94)	53.5% (0.65)	1.85
Education	22.8	39.3	60588 (29156)	43.94 (7.42)	0.30 (7.42)	1.87% (3.00)	29.8% (0.36)	0.53
p-value ^c	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Tables reports means. Standard deviations reported in parentheses.

Statistics are based on the 2010-2012 CPS monthly data. Sample restricted to those with at least a Bachelor's degree, between ages 25-40. See Appendix A for details on variable construction.

Variables in columns 3, 4, 7 and 8 are based on full-time workers.

^a Proportion of all male workers who are employed in each industry (columns sum to 100).

^b Males as a proportion of all full-time workers in that sector.

^c F-test of equality of means/proportions across the industry categories.

Table 2: Workplace Attributes by Major

	Shares		Annual Earnings for full-time workers ^a	Hrs/wk for full-time workers ^a	% Part time workers ^a	UE Rate ^b	% Not Employed	Ann % Raise ^c
	(1)	(2)						
Bachelor's (or more) in:								
Business	13.0	10.3	77377 (68733)	45.5 (8.0)	27.0	4.3	8.4	3.9
Engineering	11.7	3.2	83142 (54295)	45.0 (7.8)	21.7	3.5	6.0	4.5
Humanities	19.6	30.5	59306 (49355)	44.4 (7.8)	38.6	3.9	10.1	4.0
Natural Science	6.5	9.5	72889 (59494)	45.0 (9.4)	34.3	2.3	8.6	5.2
Some College ^d	49.2	46.5	42565 (26863)	43.5 (7.9)	43.8	7.3	13.8	3.3
F-test ^e	0	0	0	0	0	0	0	0

Table shows statistics from the 2013 American Community Survey (ACS), restricting the sample to 25-40 year olds with at least some college education. Sample size is 376,113 respondents.

^a Means (standard devs.) shown for annual earnings and hrs/week for full-time workers.

^b Proportion of Part-time workers from pool of those currently employed.

^c Unemployment rate is number not employed and currently looking for a job, divided by sum of unemployed and employed respondents.

^d Calculated by linearly regressing log earnings on age for that major group.

^e Individuals with 1-4 years of college (and no bachelor's degree), or some college.

^e p-value of F-test of equality of means across majors (rows).

Table 3: Sample Statistics

	All	Males	Females	p-value
	(1)	(2)	(3)	(4)
Number of respondents	247	86	161	
School Year:				
Freshmen	10.9%	9.3%	11.8%	0.549
Sophomore	10.9%	11.6%	10.6%	0.798
Junior	36.4%	32.6%	38.5%	0.355
Senior or more	41.7%	46.5%	39.1%	0.262
Age	21.49 (1.5)	21.69 (1.8)	21.37 (1.2)	0.103
Race:				
White	29.2%	33.7%	26.7%	0.248
Asian	50.6%	51.1%	50.3%	0.898
Non-Asian Minority	17.8%	14.0%	19.9%	0.247
Parent's Characteristics:				
Parents' Income (\$1000s)	137 (121)	141 (126)	135 (118)	0.731
Mother B.A. or More	67.6%	74.4%	64.0%	0.095
Father B.A. or More	69.6%	72.1%	68.3%	0.539
Ability Measures:				
SAT Math Score	696.0 (88)	717.7 (72)	684.3 (94)	0.006
SAT Verbal Score	674.0 (84)	677.0 (78)	672.5 (88)	0.704
GPA	3.5 (0.32)	3.5 (0.33)	3.5 (0.32)	0.938
Intended/Current Major				
Economics/Business	31.2%	48.8%	21.7%	0.000
Engineering	4.9%	8.1%	3.1%	0.080
Humanities and Soc Sciences	47.8%	30.2%	57.1%	0.000
Natural Sciences/Math	16.2%	12.8%	18.0%	0.289

For the continuous outcomes, means are reported in the first cell, and standard deviations are reported in parentheses.

P-value reported for a pairwise test of equality of means (proportions) between males and females, based on a Wilcoxon rank-sum (Chi square) test.

Table 4: Example Choice Scenarios

Panel A	Earnings per year at age 30 if working full time	Annual percentage increase in earnings from age 30 on	Average work hours per week for full-time	Work flexibility: Part-time work available?	Probability assigned by:	
					Males	Females
Example 1						
Job 1	\$96,000	3%	52	Yes	31.93 [30] (22.48)	31.46 [30] (21.36)
Job 2	\$95,000	2%	45	Yes	31.16 [30] (23.71)	39.34*** [40] (22.71)
Job 3	\$89,000	4%	42	No	36.91 [30] (24.71)	29.20** [25] (22.57)
Example 2						
Job 1	\$76,000	4%	50	Yes	19.38 [20] (19.34)	20.65 [20] (15.23)
Job 2	\$81,000	3%	44	Yes	49.47 [50] (26.63)	49.45 [50] (22.08)
Job 3	\$88,000	2%	49	No	31.15 [25] (25.36)	29.91 [25] (21.98)
Panel B						
	Earnings per year at age 30 if working full time	Probability of being fired from the job in the next year	Amount of bonus based on relative performance (% of full time earnings)	Proportion of men in the firm in similar positions	Males	Females
Example 1						
Job 1	\$87,000	1%	\$4,350 (5%)	49%	30.34 [30] (22.48)	36.68* [30] (24.33)
Job 2	\$84,000	6%	\$10,920 (13%)	67%	26.86 [30] (23.71)	30.27 [30] (21.36)
Job 3	\$95,000	5%	\$4,750 (5%)	69%	42.80 [31.5] (24.71)	33.05*** [30] (20.83)
Example 2						
Job 1	\$61,000	1%	\$6,710 (11%)	41%	25.48 [20] (26.57)	26.80 [20] (23.20)
Job 2	\$65,000	5%	\$7,800 (12%)	71%	12.14 [9.5] (12.98)	15.53** [10] (11.81)
Job 3	\$67,000	2%	\$10,050 (15%)	60%	62.38 [60] (31.55)	57.67 [60] (27.19)

Means [median] (standard deviations) reported in the last two columns. Pairwise t-tests conducted for equality of means by gender. Significance denoted on the female column by asterisks: *p<0.10, **p<0.05, ***p<0.01.

Table 5: Estimates of Job Choice Model

	Overall ^a	Males	Females
	(1)	(2)	(3)
Age 30 log earnings	15.40*** (1.65)	22.86*** (3.88)	11.42*** (1.43)
Probability of being fired	-0.38*** (0.04)	-0.39*** (0.10)	-0.37*** (0.04)
Bonus, as a prop. of earnings	0.28*** (0.03)	0.38*** (0.05)	0.22*** (0.03)
Prop of males in similar positions	0.00 (0.00)	-0.01 (0.01)	0.005 (0.01)
% increase in annual earnings	0.55*** (0.10)	1.09*** (0.22)	0.27** (0.10)
Hours per week of work	-0.15*** (0.02)	-0.21*** (0.05)	-0.12*** (0.02)
Part-time option available	0.79*** (0.11)	0.86*** (0.22)	0.76*** (0.12)
Observations	247	86	161

Table reports the average of the parameter estimates across the relevant sample. Asterisks denote estimates are statistically different from zero based on bootstrap standard errors. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 6: Heterogeneity in Job Preferences

		Overall ^a	Males	Females
		(1)	(2)	(3)
Age 30 log earnings	Median	7.37***	9.33***	6.72***
	25th pct.	2.68***	2.91*	2.58***
	75th pct.	16.18***	29.65***	13.61***+
	Std. dev	26.67***	36.42***	18.20***++
	Skewness	2.12***	1.37***	2.35***+
Probability of being fired	Median	-0.17***	-0.15***	-0.20***
	25th pct.	-0.42***	-0.39***	-0.44***
	75th pct.	-0.04**	-0.01	-0.05***
	Std. dev	0.68***	0.85***	0.56***+
	Skewness	-2.02***	-1.49***	-2.59***+
Bonus, as a prop. of earnings	Median	0.12***	0.16***	0.11***
	25th pct.	0.03***	0.04**	0.03**
	75th pct.	0.31***	0.71***	0.24***+++
	Std. dev	0.42***	0.50***	0.35***+++
	Skewness	1.88***	1.02***	2.71***+++
Prop of males in similar positions	Median	-0.003***	-0.009	-0.002**
	25th pct.	-0.024***	-0.037***	-0.016***++
	75th pct.	0.017***	0.011**	0.02***
	Std. dev	0.066***	0.07***	0.062***
	Skewness	0.591	0.382	0.837
% increase in annual earnings	Median	0.19***	0.39***	0.13***++
	25th pct.	-0.07*	0.001	-0.11***+
	75th pct.	0.69***	1.75***	0.55***+
	Std. dev	1.61***	2.05***	1.21***+++
	Skewness	1.58***	1.20***	1.12
Hours per week of work	Median	-0.07***	-0.09***	-0.06***
	25th pct.	-0.19***	-0.26***	-0.16***+
	75th pct.	-0.02*	-0.01	-0.02**
	Std. dev	0.33***	0.43***	0.26***++
	Skewness	-1.89***	-1.50***	-1.76
Part-time option available	Median	0.47***	0.43***	0.48***
	25th pct.	-0.01	-0.14	0.02
	75th pct.	1.12***	1.39***	1.02***
	Std. dev	1.70***	2.01***	1.49***
	Skewness	1.66***	1.41***	1.73***
Observations		247	86	161

Asterisks denote estimates are statistically different from zero based on bootstrap standard errors. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

+++, ++, + on the female column denote estimates for males and females are statistically different from each other at the 1%, 5%, and 10% levels, respectively.

Table 7: Willingness to Pay (WTP) Estimates

	WTP(dollars)			WTP (as % of average earnings)		
	Overall (1)	Male (2)	Female (3)	Overall (4)	Male (5)	Female (6)
Prob. of being fired	2147.40*** (525.46)	467.79 (670.29)	3044.58***+++ (715.28)	2.83%*** (0.69%)	0.62% (0.88%)	4.01%***+++ (0.94%)
Bonus as percent of earnings	-1069.78*** (258.47)	-645.92* (368.77)	-1296.19*** (345.77)	-1.41%*** (0.34%)	-0.85%* (0.49%)	-1.71%*** (0.46%)
Prop. of men at jobs	43.20 (38.31)	63.74 (46.81)	32.24 (53.93)	0.06% (0.05%)	0.08% (0.06%)	0.04% (0.07%)
Annual % raise in earnings	-1186.28 (773.44)	-2564.93** (1226.42)	-449.86 (957.85)	-1.56% (1.02%)	-3.38%** (1.62%)	-0.59% (1.26%)
Hours per week of work	854.65*** (235.25)	594.70 (416.17)	993.50*** (267.19)	1.13%*** (0.31%)	0.78% (0.55%)	1.31%*** (0.35%)
Part-time option available ^a	-3892.01*** (1024.91)	-829.94 (1822.82)	-5527.65***+++ (1221.72)	-5.13%*** (1.35%)	-1.09% (2.40%)	-7.29%***+++ (1.61%)

Table reports mean WTP (amount of earnings an individual needs to be compensated for a unit change in the job attribute).

^a WTP for moving from a job without a part-time option to one that has it.

Bootstrap standard errors in parentheses. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

T-tests conducted for differences in means by gender. +++, ++, + denote estimates differ at 1%, 5%, and 10% levels, respectively.

Table 8: Correlates of Heterogeneity in WTP

Characteristics		Prob of being Fired	Bonus as % of Salary	Prop. of Males	Annual % Raise	Hrs/wk of work	Part-time available
Gender	Male	0.60 (0.88)	-0.85* (0.49)	0.08 (0.06)	-3.41** (1.62)	0.76 (0.55)	-1.01 (2.40)
	Female	4.04***+++	-1.72*** (0.46)	0.04 (0.07)	-0.57 (1.26)	1.33*** (0.35)	-7.33***+
SAT Score ^a	Above Median	2.86***	-1.61*** (0.26)	-0.02 (0.05)	-1.00 (1.03)	1.58*** (0.37)	-4.56**
	Below Median	2.81***	-1.22* (0.65)	0.14 (0.09)	-2.15 (1.74)	0.66 (0.47)	-5.72***
Race	White	1.638	-0.96*** (0.36)	0.17** (0.09)	-4.47** (2.02)	-0.01 (0.65)	-2.68 (2.37)
	Nonwhite	3.32***	-1.60***	0.01+	-0.38++	1.60***+++	-6.13***
	Above Median	(0.88)	(0.46)	(0.06)	(1.15)	(0.33)	(1.60)
Parents' income ^b	Above Median	1.33	-1.56*** (0.35)	0.04 (0.06)	-2.59* (1.49)	0.64 (0.47)	-2.69 (1.79)
	Below Median	4.27***+++	-1.28**	0.08	-0.59	1.60***+	-7.45***+
Likelihood of marriage ^c	Above Median	1.05	-1.86*** (0.49)	0.00 (0.08)	1.43 (1.29)	0.39 (0.51)	(2.02)
	Below Median	2.82***	-1.03** (0.50)	0.00 (0.07)	-0.37* (1.54)	1.29*** (0.36)	-4.70***
Expected fertility ^d	Above Median	2.85***	-0.97* (0.56)	0.04 (0.09)	-2.61* (1.73)	0.99*** (0.55)	-5.50***
	Below Median	1.75	-1.87*** (0.40)	0.07 (0.05)	-2.31** (1.13)	1.17*** (0.29)	-5.50***
School year	Senior or older	1.09	-1.68*** (0.49)	0.20** (0.09)	-2.13 (1.62)	0.81 (0.66)	-5.22**
	Junior or younger	3.62***	-1.23**	-0.05++	-1.16	1.36***	-5.06***
Prob. of full-time work ^e	>80%	(0.91)	(0.48)	(0.05)	(1.34)	(0.25)	(1.80)
	<=80%	2.47**	-1.86*** (0.43)	0.048 (0.06)	-2.72 (1.98)	1.06* (0.64)	-3.26**
	>50%	2.90***	-1.33*** (0.39)	0.058 (0.06)	-1.35 (1.14)	1.14*** (0.35)	-5.48***
Economics Probability ^f	>50%	2.35***	-1.77***	0.03	-1.98	0.687**	-1.82
	<=50%	(0.89)	(0.64)	(0.06)	(1.53)	(0.31)	(2.61)
		3.04***	-1.27*** (0.41)	0.067 (0.07)	-1.39 (1.31)	1.31*** (0.42)	-6.48***++

Table reports the mean WTP for each group. Bootstrap standard errors are in parentheses. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

A t-test is conducted for whether means differ within each group (males versus females; below median SAT score versus above median, etc.) Significance denoted by plus signs shown on the means of the second row within each sub-group. +, ++, +++, + denote significance at the 1%, 5%, and 10% levels, respectively.

^a SAT score is the combined Math and Verbal sections. The sample median is 1400.

^b The sample median for parents' annual income is \$87500.

^c Students report the probability of being married by age 30, conditional on major. This is the probability weighted by the major choice probabilities. The sample median is 60%.

^d Students report the expected number of children by age 30, conditional on major. This is the probability weighted by the major choice probabilities. The sample median is 0.91 children.

^e Students report the likelihood of working full-time at age 30, conditional on major. This is the probability weighted by the major choice probabilities. It exceeds 80% for 32 percent of the sample.

^f Students' reported likelihood of majoring in economics (on a 0-100 scale). It exceeds 50 for 29 percent of the sample.

Table 9: Perceived Job Attributes, Conditional on Major

	Prob of majoring (1)	Annual age 30 earnings (2)	Annual % inc in earnings (3)	Prob of being fired (4)	Bonus as % of Salary (5)	Prop of males at jobs (6)	Hrs/ week work (7)	Likelihood of Part-time availability (8)	Ability rank ^a (9)	Hrs/wk of study time ^b (10)
Male respondents										
Econ/Business	43.4 (45.2)	135267 (77501)	1.5 (4.2)	13.1 (16.4)	46.2 (54.1)	60.8 (15.8)	55.5 (13.2)	24.0 (23.3)	32.0 (28.2)	22.7 (16.1)
Engineering	8.4 (21.3)	102174 (43046)	1.4 (3.5)	7.0 (7.2)	19.2 (22.0)	60.7 (19.3)	48.5 (9.5)	26.7 (22.9)	42.5 (29.7)	27.8 (17.5)
Humanities	28.9 (39.5)	67802 (20290)	1.4 (3.2)	8.9 (9.1)	13.0 (26.7)	42.5 (14.2)	44.0 (9.9)	36.5 (25.8)	25.4 (23.7)	19.2 (16.3)
Natural Sciences	16.5 (31.1)	85790 (34264)	1.5 (3.6)	6.4 (6.8)	14.2 (19.3)	53.7 (17.5)	47.6 (10.2)	29.0 (24.1)	39.5 (29.1)	27.7 (17.9)
No Degree	2.8 (9.1)	42518 (17739)	2.5 (4.3)	18.9 (18.8)	9.9 (23.9)	50.5 (16.0)	46.0 (11.5)	48.4 (32.5)	31.5 (40.3)	-
F-test ^c	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Female respondents										
Economics	23.3*** (36.6)	100469*** (53496)	1.2 (3.0)	19.1** (17.9)	39.4 (50.7)	61.4 (16.1)	52.5 (17.0)	28.6 (20.7)	46.7*** (26.8)	24.8 (12.4)
Engineering	5.7 (17.0)	92247* (37364)	1.1 (3.8)	14.8*** (16.1)	26.3 (38.9)	65.6*** (18.4)	47.8 (12.3)	32.6*** (21.5)	53.6*** (29.7)	29.3 (16.5)
Humanities	52.6*** (43.0)	60595*** (18767)	1.4 (2.4)	18.2*** (16.8)	15.4 (24.1)	40.8 (14.1)	42.8 (11.0)	44.8*** (21.7)	31.1* (25.7)	19.3 (10.2)
Natural Sciences	17.2 (32.3)	75710** (29787)	1.5 (2.8)	14.0*** (14.8)	19.4 (30.5)	55.0 (17.7)	46.4 (12.5)	35.8** (21.9)	44.7 (28.3)	28.1 (14.0)
No Degree	1.3* (3.6)	33362*** (12746)	2.1 (2.8)	34.8*** (27.9)	7.2 (13.9)	47.5 (18.5)	45.9 (11.9)	55.3 (31.0)	38.3 (40.6)	-
F-test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table reports the mean belief regarding the characteristic of the job conditional on the major (row), as perceived by male (top panel) and female respondents. Standard deviations in parentheses.

T-test conducted for whether means differ by gender. Significance denoted by asterisks shown on the female means. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

^aAnswer to: Think about the other individuals (at NYU and other universities) who will graduate in each of these categories or never graduate/drop out. On a ranking scale of 1-100 (where 1 is the HIGHEST rank), where do you think you would rank in terms of ability when compared to all individuals in that category?

^bAnswer to: How many hours per week do you think you would need to spend studying in each of the following majors in order to achieve an average GPA in that major of 4.0?

^cF-test of equality of means across majors.

Table 10: WTP (relative to Humanities)

		Non-Pec Vars ^a		All Vars	
		Males	Females	Males	Females
		(1)	(2)	(3)	(4)
Econ	Mean	-5%	3%	-29%	-41%
	Med	4%	7%	-11%	-5%
	St Dv	74%	80%	79%	129%
Eng	Mean	-6%	-20%*	-21%	-30%
	Med	0%	-1%	0%	-7%
	St Dv	46%	73%	88%	89%
Nat Sci	Mean	1%	-16%***	-11%	-12%
	Med	0%	-1%	0%	-6%
	St Dv	27%	52%	60%	67%
Dropout	Mean	16%	32%	10%	57%***
	Med	11%	11%	10%	21%
	St Dv	64%	110%	84%	157%

The table shows $WTP_{i,m} = \sum_{k=1}^K WTP_{i,k} * (X_{i,m} - X_{i,humanities})$.

WTPs are expressed as a percentage of own age 30 earnings.

Positive mean/median indicates respondent needs to be compensated to major in that field.

T-test conducted for whether means differ by gender. ***, **, * on female columns denote significance at the 1%, 5%, and 10% levels, respectively.

^a Non-pec variables include: prob fired; part-time availability; prop males; hours.

Table 11: LAD Estimates of Major Choice

LAD estimates	
Job attributes	0.018** (0.007)
Ability rank	-0.064*** (0.006)
Study time	-0.009 (0.025)
Economics Dummy	-0.583 (0.435)
Engineering Dummy	-1.155*** (0.363)
Natural Sci Dummy	-0.816** (0.381)
Total Observations	741
Number of Individuals	247

Bootstrap standard errors in parentheses.

***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 12: Marginal Contribution of Job Attributes in Major Choice

	Fired Prob. ^a	Part-time Available ^b	Hours ^c	Bonus ^a	Earnings Growth ^a	Prop Males ^d	Earnings ^e
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Males							
Economics	-3.7%	1.2%	-5.8%	4.5%	11.5%	-0.9%	25.6%
Engineering	-4.9%	1.5%	-7.7%	5.4%	15.0%	-1.1%	33.4%
Humanities	-2.2%	0.7%	-3.5%	2.6%	7.0%	-0.5%	15.5%
Natural Sci	-4.7%	1.5%	-7.4%	5.2%	14.5%	-1.1%	32.2%
Panel B: Females							
Economics	-5.0%	1.2%	-3.7%	2.9%	3.0%	0.5%	15.7%
Engineering	-5.7%	1.3%	-4.2%	3.2%	3.4%	0.6%	17.7%
Humanities	-1.9%	0.5%	-1.5%	1.2%	1.2%	0.2%	6.3%
Natural Sci	-5.3%	1.2%	-3.8%	2.9%	3.2%	0.5%	16.4%

Table shows the percent change in the probability of majoring in a given major ("marginal effect") for a given change in the job attribute (column variable).

^a Marginal effect for a change from 1% to 10% in that variable (probability of being fired; bonus as percent of base salary; earnings growth).

^b Marginal effect for a change from no part-time availability to part-time availability.

^c Marginal effect for a change from 30 work hrs/week to 50 work hrs/week.

^d Marginal effect for a change from 30% male peers to 70% male peers.

^e Marginal effect for a 10% change in age 30 earnings.

Table A1: Perceived Mapping of Majors to Work Sector

	Science		Health		Business		Government		Education	
	Males (1a)	Females (1b)	Males (2a)	Females (2b)	Males (3a)	Females (3b)	Males (4a)	Females (4b)	Males (5a)	Females (5b)
Economics	14.30 [10] (14.95)	12.76 [10] (14.82)	9.00 [5] (12.32)	11.67 [5] (16.08)	59.14 [60] (24.46)	54.45 [60] (23.85)	11.03 [10] (11.38)	13.72* [10] (10.73)	6.52 [5] (8.43)	7.40 [5] (7.29)
Engineering	55.27 [60] (24.99)	56.37 [60] (22.58)	9.31 [10] (9.12)	12.15* [10] (13.80)	18.79 [20] (15.80)	13.06*** [10] (12.38)	8.64 [5] (9.61)	10.28 [10] (9.39)	7.99 [5] (10.28)	8.15 [5] (8.44)
Humanities	13.69 [10] (14.06)	12.94 [10] (15.03)	18.21 [15] (19.10)	17.83 [15] (17.48)	20.33 [20] (19.48)	14.92** [10] (14.96)	23.78 [20] (16.43)	26.49 [20] (18.99)	24.00 [20] (19.47)	27.82 [25] (19.88)
Natural Sciences	38.76 [35] (18.85)	38.84 [40] (21.96)	22.44 [20] (18.26)	25.66 [20] (20.45)	16.17 [10] (16.07)	11.16*** [10] (11.53)	8.85 [7] (9.23)	9.24 [5] (9.79)	13.78 [10] (11.73)	15.11 [10] (14.63)
No Degree	12.97 [10] (13.48)	11.98 [10] (11.34)	13.73 [12.5] (12.28)	18.26** [20] (16.01)	29.71 [20] (22.67)	23.01** [20] (20.29)	30.10 [20] (22.65)	28.62 [20] (22.50)	13.49 [10] (14.62)	18.13** [20] (18.01)
F-test ^a	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table reports the mean belief of working in the given sector (column) conditional on the major (row), as perceived by male and female respondents. Standard deviations in parentheses.

T-test conducted for whether means differ by gender. Significance denoted by asterisks shown on the female means. ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively.

^a F-test of equality of means across majors, in each industry.

Table A2: Attribute-specific WTP (relative to Humanities)

	Prob. Fired		Work Hours		Part-time Availability		Prop. of Males		Earnings Raise		Bonus Pay	
	Males (1a)	Females (1b)	M (2a)	F (2b)	M (3a)	F (3b)	M (4a)	F (4b)	M (5a)	F (5b)	M (6a)	F (6b)
Econ	Mean	-11%	3%	13%***	1%	1%	1%	-1%	-1%	0%	-32%	-26%
	Med	0%	1%	4%	0%	0%	1%	1%	0%	0%	-16%	-13%
	St Dv	54%	70%	22%	31%	3%	3%	10%	11%	9%	51%	45%
Eng	Mean	-7%	0%	4%**	0%	1%***	1%	-2%*	1%	1%	-8%	-11%
	Med	0%	0%	0%	0%	0%	0%	0%	0%	0%	-2%	-3%
	St Dv	32%	62%	12%	14%	2%	3%	12%	13%	7%	21%	20%
Nat Sci	Mean	-4%	-22%***	2%	5%*	0%	1%***	0%	-1%	1%	-2%	-4%
	Med	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	St Dv	32%	52%	12%	14%	1%	2%	8%	9%	7%	9%	12%
Dropout	Mean	18%	28%	2%	4%	-1%	-1%	0%	0%	0%	3%	7%**
	Med	2%	5%	0%	0%	0%	0%	0%	0%	0%	0%	2%
	St Dv	60%	99%	13%	21%	3%	3%	7%	7%	10%	11%	13%

The table shows $WTP_{i,m} = \sum_{k=1}^K WTP_{i,k} * (X_{i,m} - X_{i,humanities})$.

WTPs are expressed as a percentage of own age 30 earnings.

Positive mean/median indicates respondent needs to be compensated to major in that field.

T-test conducted for whether means differ by gender. ***, **, * on female columns denote significance at the 1%, 5%, and 10% levels, respectively.