The Specificity of General Human Capital: Evidence from College Major Choice

Josh Kinsler
Ronni Pavan
University of Rochester

October 3, 2012

Abstract

College graduates with a science or business related degree earn up to 25% larger wages than other college graduates. However, individuals do not always pursue careers related to their major, generating within-major gaps in wages that are similar in size to the across major gaps. As an example, science majors who work in jobs related to their field of study earn approximately 30% higher wages than those working in non-related jobs. In this paper, we aim to estimate the true returns to college major accounting for the specificity of skill and to assess whether students know their labor market returns at the time they choose their field of study. We develop a structural model of human capital that allows for both skill uncertainty and differential accumulation of human capital across major. Our findings indicate that the returns to obtaining a business or science degree, although quite large, are smaller than the raw gaps would indicate, as is the return to working in a job related to the individual's field of study. We also find that individuals are uncertain about their future productivity at the time of the college major decision. The combination of skill uncertainty and the specificity of the return make majoring in a science related field a relatively risky human capital investment.

1 Introduction

Workers with a college degree earn significantly higher wages than high school graduates. This empirical fact has generated a vast literature that seeks to estimate the true extent of the return to schooling and to understand the factors that influence the choice of schooling quantity. An interesting feature of the data that has been largely neglected by this literature is the wide dispersion in the return to a college education across field of study. For example, college graduates with a science or business related degree earn up to 25% larger wages than other college graduates. A natural question to ask is whether these wage differences originate from true differences in the returns to college major or from ability bias, i.e. ex-ante more productive workers cluster in the same majors. Moreover, if sorting alone cannot explain the observed wage gaps, why do these differences persist and what drives students towards less remunerative majors?

Identifying the true returns to college major and the factors that influence student choices can help inform policy makers about decisions regarding higher education and field of study. Recently developed programs in the U.S., such as the National Science and Mathematics Access to Retain Talent Grant (SMART) and the Science, Technology, Engineering, and Mathematics Expansion Program (STEP), indicate a strong desire among policy makers to increase the share of college students entering technical fields.[1] The success of these and future programs will hinge on whether the policies target the key mechanisms driving major choices and whether the estimated benefits of such interventions are based on realistic returns.

The purpose of this paper is two-fold. First, we aim to estimate the true returns to college major, accounting for sorting into major based on both observable and unobservable student attributes. As part of this effort, we investigate how the returns to college major relate to student ability and future labor market choices. Second, we assess whether students know their major specific labor market returns at the time they choose their field of study. The presence, or absence, of information regarding future returns could be an important factor driving college major choices. In our data analysis, we show that individuals working in a job related to their major field of study earn significantly higher wages than those working in jobs unrelated to their major. This is particularly true for science majors, where there is

[1]The rationale for these programs is that policy makers believe that graduates who obtain a degree in a technical field generate a positive externality in the broader economy. Murphy et al. (1991) provides support for this belief. The focus on math and science skills is not limited to the higher educations sector. The National Math and Science Initiative is another recently developed program that focuses strictly on primary and secondary school students.

approximately a 30 point gap in log wages between workers in related and non-related jobs.[2] As a result, the expected return to majoring in a science related field depends critically on whether students can anticipate their occupational choices once in the labor market.

Although wage differences related to the applicability of one's major could be driven by selection, the patterns we observe in the data suggest that differences in skill prices across job types play an important role. The idea that a portion of a worker's accumulated human capital is not universally applicable across all jobs has been incorporated into career and occupational choice models for decades.[3] Yet, in most of these models, the only channel through which workers accumulate specific human capital is through work experience and the amount of schooling.[4] In this paper we instead allow for the human capital accumulation process to differ across college major. Further, since students graduating with different majors tend to work in distinct occupations, we allow the price of human capital in the labor market to vary according to the chosen major and whether the job is related to said major.

Explicitly modeling major choice in terms of human capital accumulation separates us from much of the previous literature and allows us, for example, to isolate wage variation related to skill price heterogeneity from variation stemming from the differential accumulation of human capital. Previous works estimating the returns to college major, such as Arcidiacono (2004), have allowed the returns to observable skill measures like SAT scores and cumulative grade point average to vary with college major. However, this variation could be driven by either differences in prices or unobserved human capital. An additional benefit of our approach is that we can treat schooling related human capital separate from labor market productivity. In other words, the model is flexible enough to allow for the fact that some individuals may perform quite well in school and on tests, but be relatively unproductive in the labor market. The distinction between schooling productivity and labor market productivity is especially

[2]In comparison, the gap in wages across related and non-related jobs for business and all other majors is only about 10 log points. See data Section 2 for details.

[3]Becker (1962) and Oi (1962) originally developed the notion of firm-specific human capital. Over time, the idea was expanded to include occupation, industry, and location specific human capital. Prominent examples include McCall (1990), Parent (2000), Neal (1995), Pavan (2011), and Kennan and Walker (2011). Recent papers, such as Poletaev and Robinson (2008), Gathmann and Schnberg (2010), and Yamaguchi (2011) have focused more on task specific human capital, stressing that what matters in not the job's label, but the actual tasks a job employs.

[4]Examples include Keane and Wolpin (1997) and Sullivan (2010).

important when considering how much information is available to students at the time of the college major decision.

To investigate the specificity of human capital through schooling we turn to the Baccalaureate and Beyond Longitudinal Study (B&B). The B&B consists of a representative sample of 11,192 graduating seniors drawn from the 1993 National Postsecondary Student Aid Study (NPSAS). This group of college graduates is subsequently surveyed in 1994, 1997, and 2003, providing detailed information on schooling and labor market outcomes. The B&B has two key features that make it ideal for estimating the true returns to college major and understanding the specificity of human capital through schooling. First, we are able to construct major specific grade point averages even for fields which students decided not to pursue. This gives us a window into what a student's outcomes might have looked like had they chosen a different major. Second, individuals are asked whether their job is related to their field of study. This allows us to to identify the specificity of human capital within major by looking at wage variation across job types. Of course, the raw gaps in wages across majors and job types could simply be the result of sorting. However, we show through a series of reduced-from regressions that the observed wage gaps cannot be explained by observable ability measures.

Using our reduced-form findings as a guide, we then develop a structural human capital model in order to better understand the human capital accumulation process and the true returns to college major, as well as identify the information available to students at the time of the major decision. Our framework and identification strategy is related to Carneiro et al. (2003) and Cunha et al. (2005), as it utilizes a similar factor-based structure. Individuals are endowed with two types of human capital, which we label as math and verbal human capital. An individual's math and verbal human capitals evolve as a result of schooling, where the nature of the accumulation varies across major. However, at the time of the major decision, students do not observe their math and verbal human capitals directly, but instead observe their total math and verbal abilities. Total math ability is assumed to be the sum of the math human capital and an independent math specific scholastic ability which is not priced in the labor market. Total verbal ability is similarly defined. Scholastic outcomes, such as SAT scores and GPA, are assumed to be functions of the total math and verbal abilities. Upon exiting college, individuals enter the labor market, learn their true human capitals, and choose whether to work in a job related to their field of study. Wages then depend on an individual's level of human capital and the type of job selected. While math and verbal human capital affect wages in all jobs, the skill prices can vary according to major and whether the job is related to the field of study.

We estimate our model by maximum likelihood and show that we are able to replicate the key findings from our reduced-form analysis using data simulated from the model. The parameter estimates indicate that students have very little information about their underlying verbal human capital at the time the college major decision is made. However, students receive more precise information about their underlying math human capital through their observed SAT and GPA measures. Students with high expected math human capital tend to sort into science related majors, while students with high verbal scholastic ability tend to choose fields other than science or business. The price of math and verbal human capital varies considerably both across and within major. As an example, a one-standard deviation increase in math human capital increases wages by 22% for a science major working in a related job, but has essentially no effect on wages for a science major working in a non-related job.

While we find evidence of significant skill price heterogeneity in our baseline model, the results are predicated on the assumption that individuals do not know their true underlying human capitals when making their major choice. Consequently, we estimate a more general version of our model that nests both uncertainty and perfect knowledge about an individual's human capital. We are able to reject the model of perfect certainty in favor of the model that allows for human capital uncertainty.[5]

Using data simulated from our model, we calculate the "average" returns to college major and working in a related job. We find that the returns to obtaining a business or science degree are smaller than what is indicated by a simple OLS regression, but remain quite large, on the order of 0.15 and 0.12 log points respectively. The true benefit of working in a related job increases for the residual major group relative to OLS, but decreases for science majors. These patterns can be explained by the differential sorting of individuals into related jobs across fields.

Finally, we examine how important the specificity of human capital is across majors

[5]Other papers have looked at different dimensions of imperfect information for major choice, focusing on its role for major switching, time to degree, and dropping out. See for example Altonji (1993), Arcidiacono (2004), Beffy et al. (2009), and Montmarquette et al. (2002).

through a simple wage decomposition exercise. First, we determine how much of the observed wage variability within major is related to human capital. We find that 33% of the wage variation for science majors is related to our definition of initial labor market human capital. Smaller fractions are obtained for business and all other majors. We then decompose the wage variability arising from human capital into a general and specific component. The specific nature of human capital as it relates to job type explains at least 24% of the wage variability associated with human capital in the science field. For business and all other fields, specificity is responsible for only 5% and 9% of the wage variation related to human capital. These results, in conjunction with the evidence supporting human capital uncertainty at the time of the major choice, suggest that science related majors may be a riskier human capital investment for students relative to other fields of study. This risk could aid in explaining why students tend to shy away from math and science related majors.

The remainder of the paper is as follows. Section 2 describes the B&B data in detail and provides reduced-form evidence of the specificity of human capital through schooling. In section 3 we describe our model of major choice and discuss identification. Section 4 discusses estimation and presents the structural estimates along with some simple validation exercises. In section 5 we perform several counterfactual experiments to understand the relative importance of the various mechanisms driving major choices. Section 6 concludes.

2 Data and Reduced-Form Analysis

2.1 Data

We use data from the first cohort of the B&B to investigate the links between human capital, major choice, and wages. The initial B&B cohort consists of a representative sample of 11,192 graduating seniors drawn from the 1993 National Postsecondary Student Aid Study (NPSAS). This group of college graduates is subsequently surveyed in 1994, 1997, and 2003. The original NPSAS data along with the follow up B&B surveys allow us to construct detailed information on students' backgrounds, schooling outcomes, and labor market outcomes. While the sample is representative, the B&B provide sampling weights, which we use in both the descriptive and structural analysis. Additional details regarding the sample are provided below.

2.1.1 Student Background and Schooling Data

The 1993 NPSAS and 1994 wave of the B&B collected detailed background data on each student. Using this data we are able to construct measures of respondent race, gender, and age. We limit the sample to males to avoid the complications fertility expectations have on female human capital accumulation and labor supply. This reduces the number of students to 4,834. In addition, there is a significant number of graduating seniors in 1993 who are older than 30. Older graduates tend to have lower (higher) math (verbal) SAT scores, choose business majors more often, and earn more than their fellow graduates who are significantly younger. These students likely also differ in other unobserved dimensions. In order to keep a relatively homogenous sample we drop anyone above the age of 30 in 1993, reducing the sample to 4,264 individuals. In addition, any individual with missing parental information or missing SAT scores is excluded from the sample, resulting in a final sample of 2,476 individuals.[6]

Detailed data on college major choice, as well as major specific grade point average (GPA) are readily available. In order to keep the model tractable, we collapse major choice into three broad categories: business, science, and other. The aggregate business major includes fields such as economics, accounting, and general business. The science major includes engineering, the physical and natural sciences, and computer science. All remaining fields, such as social sciences, education, psychology, humanities, etc., fall into the other category. For the estimation of the structural model, we also construct variables that are assumed to be exogenous to the model. Using the school identifiers available in the B&B, we merge information from the Integrated Postsecondary Education Data System (IPEDS) and then link each student in our sample to the share of students in each major in his graduating college. This variable will act as an exogenous shifter in the choice of college major.

In addition to major choice, information regarding student participation and performance across the various majors is also available. Total credits and GPA are separately available for business, social science, science and engineering, education, math, and foreign language courses, among others. We use these disaggregated variables to construct individual specific

[6] Many individuals in our sample take the ACT rather than the SAT. We are unable to incorporate these individuals since only the composite ACT score is available and in our empirical analysis we treat separately the SAT math and verbal tests.

GPA measures for each of our broad major categories.[7] Note that we can only construct these measures if a student ever took a course in one of these subjects. GPA measures for business, science, and other majors are available for 45%, 95%, and 98% of our sample respectively.

Table 1 provides basic summary statistics describing schooling outcomes. Overall we see that science and business majors account for approximately 50% of the sample, with students split rather evenly across the two categories. A quarter of the sample eventually obtains a post-graduate degree.[8] When we examine the characteristics of the students across each major significant differences appear. Science majors have higher SAT math and verbal scores than either the business or residual group. In addition, their science GPA is also significantly higher. Note that both business and other majors take a significant amount of science credits, approximately 20 and 18 respectively. Across fields, GPA is always highest on average for those who chose to major in that field. This likely reflects both selection and the accumulation of additional human capital most relevant for the chosen field. Finally, across majors there also appear to be important differences in family background. Individuals who major in science are more likely to come from households where both parents are foreign born.

2.1.2 Labor Market Data

In survey years 1994 and 1997, respondents were asked about their primary employment during the month of April.[9] In 2003, respondents were asked about their current job. Across all surveys, individuals provided information regarding their hours, wages, and whether their

[7]We trim the subject-specific GPAs according to the following procedure. We find the percentile x at which all individuals above this percentile receive a 4.0. We then find the GPA associated with the 1-x percentile, and replace all lower GPAs with this value. Thus, the top and bottom of the GPA distribution are trimmed in a similar fashion. The trimming does not affect the reduced form analysis and we do it to reduce the importance of the outliers which would be problematic once we estimate the structural model.

[8]In survey years 1994, 1997, and 2003, respondents are asked about post-BA degree receipt. For tractability, we treat all graduate degrees identically and do not allow individuals to switch their major at this point. This is largely consistent with the fact that 70% of the individuals who eventually obtain a graduate degree choose a graduate field of study that falls in the same broad major category as their undergraduate field of study. Note that if an individual reports obtaining a graduate degree by 1994, we utilize the graduate degree major rather than the undergraduate major. This occurs for 31 individuals.

[9]Note that in 1994, individuals were also asked about their primary job. When possible, missing information for the April job is replaced with information from the primary job.

job is related to their field of study.

Individuals can report wages either hourly, daily, weekly, monthly, or annually. We convert all wages to full-time yearly equivalents for those individuals who report working at least 30 hours per week. For anyone working fewer than 30 hours per week, we treat their annual salary as missing, since we do not model labor supply explicitly. All salaries are measured in 2000$. The most important labor market variable for our purposes is whether an individual's job is related to their field of study. For the 1994 and 1997 surveys, a job is defined as being related to an individual's field of study if the respondent reported that the April job was either closely or somewhat related to their field of study. In 2003, the relationship question is altered slightly. A respondent is considered to be working in a job related to their field of study if their undergraduate education is very important in their current job or if their graduate education is very important in their current job and the respondent reports obtaining a graduate degree. Because of the change in the wording of the question, if an individual reports being in the same job as they were in 1997, we use the relationship variable from 1997. Approximately 13% of the valid relationship entries in 2003 are changed as a result. If information about whether the job is related to the field of study is missing, to the greatest extent possible we use information from the subsequent surveys to fill it in. For example, if the respondent reports starting their 2003 job prior to 1997, we replace the missing relationship variables in 1997 with their 2003 values. If a relationship variable is missing and cannot be imputed, we also consider as missing the relationship variables for all subsequent years. We do this since in our model the likelihood of working in a related job is a function of whether the individual was working in a related job in the previous period.

Using the B&B we also construct state specific deviations with respect to the national average of the share of workers in a related job conditional on year, major, post-secondary degree, and average wages in a related job. Although agents do move across states for economic reasons, see for example Kennan and Walker (2011), in our model we assume that geographical location is exogenous. This assumption is mitigated by the fact that 75% of the individuals in our sample live in the same state in which their parents lived in 1993 or in the same state as their college.

Table 2 provides summary statistics for the labor market outcomes for the sample as a whole and by field of study. Overall, we see that approximately 70% of college graduates are working in jobs that are related to their college major. Not surprisingly, annual salaries are significantly different according to whether an individual is working in a job that is related to their field of study. This difference could reflect sorting, meaning that higher ability individuals are more likely to work in a related job, or it could reflect the idea that human capital is priced differently across different types of jobs.

Looking at labor market outcomes across fields of study illustrates important differences in outcomes by major. First, business and science majors earn significantly higher salaries relative to the residual group. This pattern is not unique to the B&B and can be found in the National Longitudinal Surveys of Youth (1979 and 1997) and more recently in the 2009 American Community Survey.[10] Second, and more important for our purposes, is the huge impact that working in a related job has on the salary of science majors. Science majors who work in a job related to their field of study earn close to a 30% premium relative to science majors who work in an unrelated job. The gaps for business and the residual major are only 3% and 11% respectively. These patterns are consistent with varying returns to skill, but could also reflect differing degrees of sorting into related jobs across majors or issues with aggregation. As an example, the wage premium associated with working in a related job for science majors could simply reflect that engineering majors are both more likely to work in a related job and earn more as compared to less remunerative science fields such as biology.

2.2 Reduced-Form Evidence of the Specificity of Schooling Human Capital

Table 2 illustrates that not only do wages vary considerably across college major, but also within major according to whether an individual works in a related job. While these patterns are consistent with heterogenous returns to college through major choice and the existence of skill price differentials by job type, sorting or aggregation bias introduced by our course characterization of majors could also rationalize the data. In this section we perform some simple reduced form analysis to shed light on the underlying mechanisms that drive the observed patterns and to help motivate our modeling choices regarding human capital and college major choice.

To study the wage differences across majors, we begin with some simple OLS regressions presented in Table 3. Column 1 indicates that conditional on year effects and graduate degree

[10] Additional details available upon request.

receipt, business and science majors earn 0.19 and 0.23 log point higher wages relative to the residual major category. Adding controls for observable measures of ability like SAT and GPA in columns 2 and 3 mildly decreases these differentials.[11] Thus, the wage variation across major evident in the summary statistics does not appear to be driven entirely by differences in observed math and verbal ability. Column 4 shows that the estimated major returns are not sensitive to whether we use hourly wages rather than our measure of annualized income.

The primary advantage of the B&B in examining the returns to college major is that it contains direct information on whether an individual's job is related to their field of study. In Table 4 we look at the relationship between this variable and the earnings of a worker. Although the resulting patterns could be generated by the presence of sorting on unobservables, we interpret these regressions as evidence of skill price heterogeneity. In the first column we note that, even after controlling for observable measures of human capital using SAT and GPA, wages are significantly larger in jobs that are related to the field of study, and this relationship is significantly stronger for science majors.[12] In column 2 we show that this pattern is not generated by our aggregation of majors. Including dummy variables for each of the 28 majors observed in the data, we see that the impact of being in a related job for the residual major decreases, yet we still observe large returns to working in a related job for business majors and science majors.

The third column of Table 4 examines how the inclusion of worker fixed effects alters the estimated returns to working in a related job. The importance of the job's relationship with the field of study decreases significantly for business and science majors, indicating that sorting across majors on unobserved dimensions can help explain the relationship between wages and labor market outcomes. However, even in this case science majors receive a much larger increase in their wage in related jobs when compared to workers with different majors. The inclusion of worker fixed effects ensures that the returns to working in a related job are

[11] An interesting result that emerges in Table 3 is that conditional on math SAT scores, verbal SAT scores negatively impact wages. This result is robust to controls for major, relatedness, and occupation. Further, the negative impact of verbal test scores on wages can be replicated using other data sources, such as the National Longitudinal Surveys of Youth. While we do not pursue in detail the root of this negative relationship, we choose a specification of our human capital model that is flexible enough to replicate this empirical regularity.

[12] These results are consistent with recent findings from Silos and Smith (2012) who find that wage growth is positively related to how applicable the skills obtained in college are to the current job.

identified by workers who switch job types. The fact that these workers switch, however, suggests that being in a related job is less salient for them as compared to workers who do not switch. As a result, the estimated returns in column 3 likely understate the average return to working in a related job.

The fourth column of Table 4 investigates whether the returns to observable measures of skill vary with relatedness. In other words, is there any direct evidence of skill price heterogeneity. The results indicate that the returns associated with major specific GPA increase dramatically when working in a related job, though the returns to SAT math and verbal are largely unaffected. Because the differential returns to SAT by relatedness could move in opposite directions across major, this result is not surprising. However, it is clear that the returns to the specific skills learned in each major vary considerably with the worker's job type.

While the results from Table 4 suggest that the human capital developed through schooling is specific to particular types of jobs, no definitive conclusions can be drawn since SAT and GPA are noisy measures of student ability. Thus, we cannot rule out the possibility that the important patterns we find in the reduced-form are either inflated or attenuated by the presence of sorting on unobservables. For this reason, we present and then estimate a model of college major choice that explicitly allows for sorting into field of study based on unobservable attributes. An additional benefit of the structural model is that it allows us to determine whether the choice of college major is made in an environment characterized by skill uncertainty.

3 The Model

Agents begin their life in college, before deciding their major field of study. After choosing a major and finishing college they enter the labor market. While in the labor market, individuals choose whether to work in jobs that are related to their field of study. After working for T years, individuals retire for T_R years.

3.1 Human Capital and Returns to College

In our model, individuals are characterized by two types of human capital: mathematical human capital and verbal human capital. These two types of human capital are "general" in the sense that workers bring these skills with them to all types of jobs. At the same time the human capitals are "specific" since their value in the labor market depends on the particular job chosen.

When an individual starts his life, he is endowed with a vector of human capital. Initial math and verbal human capitals, denoted $H_{m,0}$ and $H_{v,0}$, are drawn from independent distributions $F_j(\cdot)$ for $j = \{m, v\}$. While in college, an individual's math and verbal human capital evolve as a function of the chosen major according to

$$H_{j,1} = H_{j,0} + \mu^H_{j,f^*} \;, \text{ for } j = \{m, v\} \tag{1}$$

where f^* indicates the chosen field of study. The accumulation of human capital through major choice is assumed to be independent of the initial levels of the math and verbal skill. It is theoretically possible to identify heterogenous returns given the available data, however, in practice identification would be rather weak.

3.2 Wages

Once in the labor market individuals receive a wage that depends on their major field of study, their post-schooling human capitals, and whether the chosen job is directly related to the field of study of the worker:

$$\ln w_{r,f^*,t} = p_{r,f^*,m} H_{m,t} + p_{r,f^*,v} H_{v,t}. \tag{2}$$

The t subscript indicates calendar time, while the r subscript indicates whether the job is related to the studies of the individual. The coefficient on each type of human capital depends on both the individual's major and the relatedness status of the job. This reflects the idea that workers from different schooling backgrounds sort into different occupations that possibly reward each type of human capital differently. Similarly, related jobs may reward skills differently than non-related jobs.

Although it is reasonable to expect that human capital evolves once in the labor market as a result of accrued experience and/or post-secondary education investment, we assume that human capital remains constant since we have relatively few wage observations in the time dimension. Thus, $H_{j,t} = H_{j,1} \; \forall \; j$. However, in order to allow for the fact that wages increase both with time and the acquisition of a graduate degree, we also include in our wage specification a time dummy and a graduate degree dummy that depends on the major chosen.

Under the assumption that human capital is fixed upon entering the labor market, we can utilize Equation (1) to re-write the wage as a function of the initial levels of human capital:

$$\ln w_{r,f^*,t} = p_{r,f^*,c} + p_{r,f^*,m} H_{m,0} + p_{r,f^*,v} H_{v,0}. \tag{3}$$

As a result we obtain a wage constant that is a function of both the chosen major and job type. Finally, we assume that the econometrician observes $\ln w^{obs}_{r,f^*,t} = \ln w_{r,f^*,t} + \epsilon_{r,f^*,t}$, where $\epsilon_{r,f^*,t} \sim F_\epsilon(\cdot)$ and the variance of $\epsilon_{r,f^*,t}$ is allowed to differ across majors and relatedness.[13]

3.3 Measurement Equations

A crucial component of our empirical analysis is the availability of observable measures of individual ability, such as SAT scores and college GPAs. These measurements allow the econometrician to have a direct, although imperfect, look at the individual's human capital. Math and verbal SATs are taken prior to starting college and therefore are useful for identifying math and verbal human capital prior to choosing a college major. Subject-specific GPAs provide additional information regarding the post-collegiate human capital that individuals will ultimately bring to the labor market.

Although we assume that measures of scholastic aptitude such as the SAT and GPA contain information about an individual's labor market skills, we want to allow for the possible distinction between being a good student and being a good worker. Therefore, we introduce the concept of scholastic ability, which we denote by $\nu_{j,0}$ for $j = \{m, v\}$. The math and verbal scholastic abilities are orthogonal to each other and to the student's human capital. Similar to the math and verbal human capital, we let the scholastic abilities evolve with major choice according to

$$\nu_{j,1} = \nu_{j,0} + \mu^{\nu}_{j,f^*}, \text{ for } j = \{m, v\}. \tag{4}$$

Scholastic performance, as measured by the SAT and GPA, is assumed to be a function of the sum of human capital and scholastic ability: $A_{j,0} = H_{j,0} + \nu_{j,0}$ for $j = \{m, v\}$. We refer to $A_{j,0}$ as the pre-college total math or verbal ability. Utilizing our definition of total ability,

[13] It is possible that a fraction of the random variable $\epsilon_{r,f^*,t}$ actually captures an idiosyncratic component of wages rather than strictly measurement error. This component would be an additional source of major-specific risk that we do not consider in our model. For example, the empirical framework developed by Nielsen and Vissing-Jorgensen (2010) is aimed to capture this component of risk.

we assume that the SAT math and verbal measurements are determined according to:

$$
\begin{aligned}
SAT_m &= \eta_{m,c} + \eta_{m,m} A_{m,0} + u_m \\
SAT_v &= \eta_{v,c} + \eta_{v,v} A_{v,0} + \eta_{v,m} A_{m,0} + u_v.
\end{aligned} \tag{5}
$$

The residual components, u_m and u_v, are measurement error and are orthogonal to the total abilities. The math SAT score is a dedicated measure of the student's total math ability, while the SAT verbal score is a function of both total verbal and total math ability. We include total math ability in the verbal score since SAT scores are highly correlated in the data.[14] The subject-specific GPAs are similarly defined:

$$GPA_{f,f^*} = \eta_{f,f^*} + \eta_{f,m} A_{m,1} + \eta_{f,v} A_{v,1} + u_f^{gpa}$$

where $A_{j,1}$ represents the post-schooling level of total ability. Notice that we allow the constant of the GPA equation to depend on the major actually chosen. This feature can capture the fact that students may take easier courses in fields outside their major field of study. Similar to the wage equation, we can express the GPA measurement equations as a function of the initial level of total ability, utilizing the accumulation equations for human capital and scholastic ability.

$$GPA_{f,f^*} = \tilde{\eta}_{f,f^*} + \eta_{f,m} A_{m,0} + \eta_{f,v} A_{v,0} + u_f^{gpa} \tag{6}$$

The constant in the above equation now includes the subject-specific shifters for human capital and scholastic ability illustrated in Equations (1) and (4).

Our assumption that human capital and scholastic ability evolve as a function of the major field of study may not be ideal in the presence of double majors. Fortunately in our data set only 5% of students have multiple majors. The GPA measurements are assumed to be missing at random, and the random components u_f are assumed to be mean zero, independent of the human capitals, but with potentially different distributions across fields.[15]

[14] The exclusion of verbal human capital from the SAT math measurement is without loss of generality. Its inclusion would simply require a re-labeling of the "math" human capital. Note that the math and verbal labels are themselves arbitrary.

[15] It should be noted that GPA measurements are missing for those students that have not taken classes in that particular subject and therefore the assumption of random attrition, although convenient, is quite strong.

3.4 Dynamic Problem

3.4.1 Labor Market

Upon entering the labor market individuals decide whether to work in a job related to their field of study.[16] When choosing whether to work in a job related to the field of study, workers consider the monetary and non-monetary components of the utility associated with each type of job. Although we assume that related jobs are always available, a large non-monetary cost could capture the fact that these jobs are more difficult or costly to obtain. We assume that the non-monetary component of the utility is separable with respect to the monetary component and is given by:

$$C_{r,f^*,t} = c_{r,f^*,t}(H_{m,1}, H_{v,1}, Z_{r,t}, r_{t-1}) + u^C_{r,t}. \tag{7}$$

$u^C_{r,t}$ is assumed to be an idiosyncratic preference shifter that is uncorrelated with the human capitals. The variable $Z_{r,t}$ is an exogenous shifter for the non-monetary preferences for related jobs. In the empirical implementation we utilize state-specific deviations from the national averages of the fraction of workers in related jobs, controlling for average wages and demographic composition. r_{t-1} is an indicator variable for whether the individual worked in a related job in the previous period.[17] We include this to capture the fact that switching job types can be costly.

We assume that the value of retirement is equal to the present discounted value of the utility derived from pension payments, which are assumed to be equal to the last wage earned in the market. While working, the value function at the beginning of a period t is given by:

$$V_t(f^*, H_{m,1}, H_{v,1}, Z_t, r_{t-1}) = E_{u^C_{r,t}}\{\max_r\{C_{r,f^*,t} + U(w_{r,f^*,t}(\cdot))\}\} + \beta V_{t+1}(\cdot)\,. \tag{8}$$

[16] Although the choice of acquiring a graduate degree is clearly related to the human capital of an individual, we decided to treat graduate studies as an exogenous variable known to the agent since the beginning of life, and it is suppressed in the presentation of the model for expositional clarity. This choice simplifies greatly the computation of the likelihood function and reduces the number of the parameters to be estimated. We believe that the results are not strongly affected by this simplification given that in our sample workers with graduate degrees do not earn appreciably larger wages than workers without them (around 7%) and the sorting into those studies does not look very different across majors. Previous preliminary estimations of a version of this model with endogenous graduate studies support this hypothesis.

[17] We assume that $r_0 = 0$ for all graduating seniors.

Once the time t preference shocks are revealed, individuals choose relatedness to maximize current utility. Thus, the model predicts that the probability a person works in a related job at time t is:

$$\begin{aligned} Pr(r_t = 1) &= Pr(c_{1,f^*,t}(\cdot) + u^C_{1,t} + U(w_{1,f^*,t}(\cdot)) > c_{0,f^*,t}(\cdot) + u^C_{0,t} + U(w_{0,f^*,t}(\cdot))) \\ &= Pr(u^C_{0,t} - u^C_{1,t} < \Upsilon(f^*, H_{m,1}, H_{v,1}, Z_t, r_{t-1})). \end{aligned} \tag{9}$$

In our empirical implementation, we estimate the above reduced-form choice probability rather than attempt to separate the parameters governing the non-monetary and monetary components of utility. Notice that we can replace the post schooling human capitals with the pre-schooling values using the accumulation process outlined in Equation (1): $Pr(r_t = 1) = Pr(u^C_{0,t} - u^C_{1,t} < \widetilde{\Upsilon}(f^*, H_{m,0}, H_{v,0}, Z_t, r_{t-1}))$.

3.4.2 Schooling

Prior to entering the labor market, individuals must decide which field of study to pursue. We assume that this choice is made under uncertainty, i.e., individuals do not observe their human capitals $(H_{m,0}, H_{v,0})$ but only their total abilities $(A_{m,0}, A_{v,0})$. As a result, individuals will utilize the available information to infer what their future human capital will be. We assume that the true human capitals are revealed upon entering the labor market.

When choosing a college major, students consider the the non-monetary components of the utility associated with each major and the impact their major choice will have on future labor market outcomes. We assume that the non-monetary component of the utility over major choice is separable with respect to the future labor market returns and is given by:

$$C^{UG}_f = c^{UG}_f(gpa^P_f, Z_f) + \varepsilon^{UG}_f. \tag{10}$$

The non-monetary component of the major choice is a function of the GPA that a student would achieve if he decides to choose that major, defined as gpa^P_f, a vector of variables Z_f that are assumed to be exogenous conditional on observables, and an idiosyncratic preference shifter, ε^{UG}_f, that is uncorrelated with the human capitals.[18] In our empirical implementation

[18] To derive the potential GPA in field f, we first recover the random component of the GPA, u_f, using the realized GPA and the actual major choice. Thus, we are assuming that the draw on the random component is both fixed and known at the time of the major choice. For those individuals with missing GPA measures, we draw u_f from the appropriate distribution to generate the potential GPA.

Z_f is the share of students in each field at the chosen college and whether the student's parents are foreign born. While the selected college may not be chosen randomly, we are implicitly assuming that conditional on the SATs, the share of students in each major affects student choices through channels other than human capital.

The value function associated with the major choice while in college is given by:

$$V_0\left(A_{m,0}, A_{v,0}, gpa^P\right) = E_{\varepsilon^{UG}}\left\{\max_f \left\{C_f^{UG} + E_{H_0}[V_1(f,\cdot)]\right\}\right\}. \tag{11}$$

The first expectation reflects the uncertainty over the idiosyncratic preference shifter prior to making the major choice. The second expectation captures the idea that students are uncertain about their underlying human capitals. Once the idiosyncratic component of utility is revealed, students choose the major that yields the highest expected utility. Thus, the probability that a student chooses major j, $Pr(f^* = j)$, is equal to:

$$Pr\left(c_j^{UG}(\cdot) + E_{H_0}[V_1(j,\cdot)] + \varepsilon_j^{UG} \geq c_{j'}^{UG}(\cdot) + E_{H_0}[V_1(j',\cdot)] + \varepsilon_{j'}^{UG}\right) \ \forall j'$$

The value functions associated with the future labor market outcomes are defined over the human capitals and the exogenous shifter for the non-monetary preferences for working in a related job. We can thus re-write the above probability of choosing major j as

$$Pr\left(\varepsilon_j^{UG} - \varepsilon_{j'}^{UG} \geq E_{H_0}\left[\Psi_{j'}(H_0, gpa_{j'}^P, Z_{j'})\right] - E_{H_0}\left[\Psi_j(H_0, gpa_j^P, Z_j)\right]\right) \ \forall j' \tag{12}$$

where Z_j includes all the exogenous variables of the model. In our empirical implementation, we estimate the above reduced-form choice probability as opposed to explicitly solving the dynamic problem.

3.5 Identification

In this section, we discuss how several key parameters of the model are identified. We focus primarily on the identification of the wage parameters of Equation (3), the measurement parameters of Equations (5) and (6), and the standard deviations of the unobserved human capitals. Our strategy is based on an infinity argument as in Carneiro et al. (2003) and we refer to that paper for all the technical details. This assumption implies that we can move some element of the vector of exogenous variables Z in such a way that the resulting probability of selecting a certain combination of major choice and relatedness is equal to one. Clearly the infinity argument is a strong argument when compared to the patterns that we observe in the data. Our point of view is that this identification strategy indicates whether theoretically, the data can be rich enough to identify the model even without any restrictions implied by distributional assumptions. Of course, imposing a set of distributional assumptions will help in the actual implementation of the estimation.

In our empirical implementation we assume that agents can choose among 3 majors: business (B), science (S) and other (O). We normalize the factor loadings of the math and verbal total abilities in the SAT math and verbal measurement equations $(\eta_{m,m}, \eta_{v,v})$ to one. Additionally, we normalize all random variables to have mean zero. These are done without loss of generality.

We start by considering the identification of the parameters of the GPA measurement equations, the variances of the total abilities $\left(\sigma^2_{A_m}, \sigma^2_{A_v}\right)$, and the parameter $\eta_{v,m}$ that governs the correlation between the math and verbal SAT measures. In order to prove the identification of these parameters consider the following covariances

$$\begin{aligned}
A &= cov\left(SAT_m, SAT_v\right) = \eta_{v,m}\sigma^2_{A_m} \\
B_f &= cov(GPA_f, SAT_m) = \eta_{f,m}\sigma^2_{A_m} \\
C_f &= cov(GPA_f, SAT_v) = \eta_{v,m}\eta_{f,m}\sigma^2_{A_m} + \eta_{f,v}\sigma^2_{A_v} \\
D_{f,f'} &= cov(GPA_{f'}, GPA_f) = \eta_{f,m}\eta_{f',m}\sigma^2_{A_m} + \eta_{f,v}\eta_{f',v}\sigma^2_{A_v}
\end{aligned}$$

calculated for fields f, f' and f'' and all their combinations. Working with these 10 equations it is possible to show that:

$$\eta_{v,m} = \frac{D_{f,f''}C_{f'} - C_{f''}D_{f,f'}}{B_f \frac{B_{f''}C_{f'} - C_{f''}B_{f'}}{A} + D_{f,f''}B_{f'} - B_{f''}D_{f,f'}}$$

Once $\eta_{v,m}$ is identified we can also recover $\left(\sigma^2_{A_m}, \sigma^2_{A_v}\right)$ and $(\eta_{f,m}, \eta_{f,v})$ for all f. The measurement constants can then be identified by looking at the average SAT scores and the average GPAs across major for individuals that have a probability one of selecting each of the three majors.

Treating the measurement parameters and variances of the total abilities as known, we can now study the identification of the labor market parameters $(p_{r,f^*,m}, p_{r,f^*,v})$ and the human capital variances $\left(\sigma^2_{H_m}, \sigma^2_{H_v}\right)$. Consider first the set of agents that choose with probability one major f or major f' and work in a job unrelated to their field of study.[19] Calculating the covariances of the wages with the SAT measures and the auto-covariances of the wages we obtain:

$$
\begin{aligned}
a_f &= cov(\ln w_{r,f,t}, SAT_m) = p_{r,f,m}\sigma^2_{H_m} \\
b_f &= cov(\ln w_{r,f,t}, SAT_v) = \eta_{v,m}p_{r,f,m}\sigma^2_{H_m} + p_{r,f,v}\sigma^2_{H_v} \\
c_f &= cov(\ln w_{r,f,t}, \ln w_{r,f,t-1}) = \left(p_{r,f,m}\right)^2\sigma^2_{H_m} + \left(p_{r,f,v}\right)^2\sigma^2_{H_v}.
\end{aligned}
$$

Working with this set of 6 equations we can solve for the value of $p_{r,f,m}$:

$$
p_{r,f,m} = \frac{a_f c_f \left(b_{f'} - \alpha a_{f'}\right)^2 - a_f c_{f'} \left(b_f - \alpha a_f\right)^2}{a_f^2 \left(b_{f'} - \alpha a_{f'}\right)^2 - a_{f'}^2 \left(b_f - \alpha a_f\right)^2} \tag{13}
$$

and subsequently all other parameters. Now that we have the variance of the human capitals we can also calculate the variance of the scholastic abilities by subtracting the former from the variance of total ability. The identification structure relied on $\eta_{v,m}$ being different from zero. However, we can also identify all the parameters even if $\eta_{v,m} = 0$.[20] Once the identification of these parameters is achieved we can apply Kotlarski's theorem (1967) as in Carneiro et al. (2003) to show the non-parametric identification of the distribution of the unobserved variables.

Notice that we have shown identification of the "reduced form" equations, in the sense that we have not identified the structural parameters that govern the evolution of the human capitals. As an example, we have shown identification of Equation (3), where the constant in this equation incorporates the major specific returns to math and verbal human capital. For the purposes of this paper, the parameters we have identified are sufficient to describe the model. With some additional normalizations we could identify the underlying structural parameters, however, we do not follow this approach. What we have not shown here but it is also crucial for our estimation is the identification of the parameters describing the functions $\widetilde{\Upsilon}(f^*, H_{m,0}, H_{v,0}, Z_t, r_{t-1})$ and $\Psi_j(H_0, gpa_j^P, Z_j)$. In this case we approximate these functions

[19] Using the same infinity argument we can also identify the wage constants by looking at the average wages. Similarly, to identify the variances of all measurement errors we can utilize the variances of the associated observable measures.

[20] Results available upon request.

with linear specifications and apply the standard normalizations (for example $\Psi_j = 0$ for $j =$ Other). By looking at the correlations between the choices and the measurements we can identify these functions.

4 Estimation and Model Fit

In this section we discuss estimation of the model parameters. We estimate our model by maximum likelihood. Let the data for an individual i be:

$$Y_i = \{f_i^*, SAT_i, GPA_i, G_{i,t}, r_{i,t}, w_{i,t}, Z_i\} \qquad \text{for } t \in \{1994, 1997, 2003\} \tag{14}$$

where $f_i^* = \{$Science(S), Business(B), Other(O)$\}$, $r_{it} = \{1$ if Related, 0 if Non-Related$\}$, SAT_i includes both math and verbal scores, and GPA_i includes grades across the three major fields. Z is the vector of exclusion restrictions while G is a dummy for graduate studies (which are assumed to be exogenous to the model). Ω is the vector of parameters that describe the model.

For expositional convenience, suppose for a moment that the econometrician is able to observe the human capitals $H_i = \{H_{i,m,0}, H_{i,v,0}\}$ and the total abilities $A_i = \{A_{i,m,0}, A_{i,v,0}\}$. Under this assumption, the individual contribution to the likelihood function $L(Y_i|Z_i, H_i, A_i; \Omega)$ can be written as follows:

$$\begin{aligned} L(Y|Z,H,A;\Omega) \;=\; & \Pr(f^*|Z,A,gpa^P;\Omega) \\ & \times \prod_{j\in\{m,v\}} f_{u_j^{SAT}}(SAT_j|A;\Omega)) \times \prod_{j\in\{S,B,O\}} f_{u_j^{GPA}}(GPA_j|A,f^*;\Omega)) \\ & \times \Pr(r_{1994}|Z,H,f^*,G_t;\Omega) \times \prod_{t\in\{1997,2003\}} \Pr(r_t|Z,H,f^*,G_t,r_{t-1};\Omega) \\ & \times \prod_{t\in\{1994,1997,2003\}} f_{u_t^{wage}}(w_t|Z,H,f^*,G_t,r_t;\Omega) \end{aligned} \tag{15}$$

where for ease of presentation we suppress the individual subscripts. Note that in the probability of working in a related job the lag related measure refers to the previous survey year. In order to calculate the terms in the above likelihood function we need to impose some parametric assumptions. As noted earlier, we assume that the utility for each major choice and each relatedness choice contains an idiosyncratic extreme-value shock, which yields a simple logit-type probability for both choices. We also assume that the measurement errors in the

GPAs, SATs, and wages are normally distributed. We let the dispersion of the idiosyncratic component of log wages vary by education and type of job.

Of course the econometrician cannot observe H_i or A_i, so we cannot evaluate the above likelihood function directly. The unobserved random variables must be integrated out of the likelihood function:

$$L(Y|Z;\Omega) = \int L(Y|Z,H,A;\Omega)dF_{H,A}(H,A;\Omega) \tag{16}$$

We assume that the unobserved human capitals and scholastic abilities are normally distributed, and take 20,000 draws from the vector of the unobserved human capitals and scholastic abilities to evaluate the above integral.

In order to calculate the probability of choosing a certain major, we need to calculate the expected human capitals conditional on the total abilities and the model parameters. Recall that when making the major choice, students only observe the sum of their scholastic ability and human capital. Under the assumption of normality, we can derive closed form solutions for these expectations rather simply. For each iteration of the likelihood function we update these expectations and plug them into the utility function over major choice. If the GPA is missing we integrate the probability over the distribution of its measurement error.

The full set of 83 parameter estimates and standard errors from the model are reported in Appendix Tables A.1, A.2, and A.3.[21] Note that when we estimate the model we restrict the loading on the total math ability in the verbal SAT equation ($\eta_{v,m}$) to zero and utilize a residualized version of the verbal SAT score as a dedicated measure of total verbal ability. The residual verbal SAT score is obtained as the residual from a regression of the verbal SAT score on the math SAT score. We found that the model fit the data better utilizing this restriction, since it essentially allows the original measurement errors of the SAT equations to be correlated.[22] Finally, we restrict the coefficients on the math human capital in the wage

[21] We calculate standard errors by inverting a numerical approximation of the Hessian.

[22] The SAT math and verbal tests are taken at essentially the same time, so it is likely that the components of the test scores unrelated to ability are correlated. Utilizing the residualized version of the SAT verbal score as a dedicated measure of verbal ability is consistent with our original model specification if we assume that the projection coefficient of total verbal ability on total math ability is equal to the projection coefficient of the verbal SAT measurement error on the math SAT measurement error. Note that it is not possible to identify the correlation in the error term and therefore our assumption cannot be tested.

equation and the coefficients on potential GPA in the major specific utility equations to be positive.[23]

There are a few results worth pointing out directly. The standard deviation of the verbal human capital is quite small, while the standard deviation of the scholastic verbal ability is positive and economically significant. This implies that there is little information about the verbal human capital in the SAT verbal and GPA measures. As a result, the variability in the expected verbal human capital is quite small relative to the variability in the true verbal human capital. More precisely, $\frac{\sigma_{E[H_{v,0}]}}{\sigma_{H_{v,0}}} = 0.10$. Thus, there is minimal sorting into major based upon verbal human capital despite the rather large coefficients in the major choice equation.

In contrast, the standard deviations of the math human capital and math scholastic ability are both positive and economically significant. This suggests that the SAT math and GPA measures contain useful information about math human capital. In fact, the ratio of the standard deviations of expected and actual math human capital is much larger then the corresponding verbal ratio, $\frac{\sigma_{E[H_{m,0}]}}{\sigma_{H_{m,0}}} = 0.32$. Additionally, the coefficients indicate that math human capital plays an interesting role in the science field. Math human capital yields no return for science majors who work in non-related jobs, but yields a large return in related jobs. The gap in wage returns to math human capital across job type for the other two major categories is considerably smaller. Similarly, the baseline return to working in a related job is much higher for science majors than for the others. These two sets of estimates are quite consistent with our reduced form analysis.

For additional evidence regarding the fit of our estimated model, we perform a set of simple validation exercises. Using the estimated model parameters we simulate major choices and labor market outcomes for a large number individuals (20 times the size of the original sample). We then compare the simulated data to the actual data, the results of which are shown in Tables 5 and 6. In Table 5 we can see that the model does a good job fitting the basic patterns in the data. The model captures the fact that students who choose to major in a science related field are strongly positively selected on SAT math scores and negatively selected on residual SAT verbal scores. Business majors have low math and residual verbal

[23]When calculating standard errors, we assume that the parameters that hit the non-negativity constraints are equal to zero.

SAT scores in both the data and model. Table 6 indicates that the model also captures the key patterns from our reduced-form wage regressions quite well. In particular, we match the unconditional mean wages across major as well as the differential in the returns to relatedness across major in the wage specification controlling for SAT and GPA.

One moment we fail to match precisely is the negative returns to SAT verbal in the labor market. We are able to capture the sign, but not the magnitude of the conditional relationship between SAT verbal and wages. The model has difficulty replicating the wage impact of our three measures of ability, SAT math, verbal, and GPA, with only two types of human capital. A third type of human capital that affects GPA only would help fit the data better, however this adds to the computational burden and is left for future research.

5 Results

While the parameter estimates and model fit are informative, the primary advantage of the structural approach is that we can utilize our model to estimate the returns to college major and quantify the specificity of human capital when individuals differ in their unobserved abilities. However, our structural framework is predicated on the assumption that individuals are uncertain about their underlying human capitals at the time of the college major decision. Thus, we first want to examine whether our key informational assumption is consistent with the data prior to investigating returns and human capital specificity.

5.1 Testing the Information Set

Our baseline model assumes that individuals are uncertain about their true underlying human capitals when they make their major choice. Specifically, one of the arguments of our linearly approximated choice function $E(\Psi)$ is $E(H_0|A)$, which implies that students can only extract some information about their human capital from their total ability. Alternatively, students may be able to observe H_0 directly and therefore base their major decision on this rather than on expectations of H_0. Similar to Carneiro et al. (2003), Cunha et al. (2005), Navarro (2011), or Guvenen and Smith (2010), testing the information set of the agent is essentially equivalent to testing whether the individuals choices are a function of the human capitals or their expected values.

Rather than assuming that individuals either know or do not know their human capitals, we specify a more general setting in which agents base their decision on the value of:

$$(1-\alpha)\, E[H_{j,0}|A;\Omega] + \alpha H_{j,0} \ \text{ for } j \in \{m, v\},$$

where $0 \leq \alpha \leq 1$. When $\alpha = 0$ we obtain the main specification of our model and when $\alpha = 1$ we obtain an alternative specification with perfect knowledge about human capital. When we estimate this model with α as a free parameter, we maximize the likelihood function at $\alpha = 0$, obtaining the same estimates reported in Section 4. This result indicates that the model with uncertainty provides a better fit of the underlying data. To determine whether the model with uncertainty is statistically different from the alternative we construct a confidence interval for α. In order to calculate the confidence interval for α, we have to account for the fact that the parameter space is bounded and that the estimate of α hits one of the boundaries. Assuming that α is the only parameter that lies on the boundary, we can use the results of Chant (1974) and Self and Liang (1987) that show that the estimated parameter is distributed as a truncated normal. The 95% confidence interval for α is $[0, 0.005]$ and therefore we strongly reject that $\alpha = 1$.

While we are able to easily reject the full information setting, it is interesting to consider what would happen if we instead estimate a model that assumes perfect knowledge. Although we do not present the parameter estimates and model fit from the perfect knowledge model, we find that this model is not able to capture the patterns in the data as well as the model under uncertainty. More precisely, the sorting into major based on math and verbal ability is not as strong as in the data, nor as strong as in the benchmark model. In fact, there is essentially no sorting into major associated with SAT verbal scores, a persistent feature of schooling data.

5.2 Counterfactuals

Having established that individuals make their major choices with limited information, we now estimate the true returns to major and relatedness, and quantify the role of varying skill prices. In all of the following regressions we use data simulated from the model as described at the the end of Section 4.

5.2.1 The Returns to Major Choice and Relatedness

Here we provide our estimates of the average return to each major if we forced students into each major (Average Treatment Effect - ATE), the average return to each major for those who selected that major (Treatment of the Treated - TT), and finally the average return for those that selected a different major (Treatment of the Untreated - TUT). In Table 7, we report the results jointly with the OLS estimates from Table 6. Across the population, the average return to a business degree declines to 0.15 log points from an OLS estimate of 0.19. The estimated return to obtaining a science degree drops from 0.23 log points in the OLS specification to 0.12 in the structural model. In contrast, controlling for observable measures of ability using SAT and major specific GPA reduces the OLS returns by only one or two log points, highlighting the importance of accounting properly for the presence of measurement error in the available ability measures. Yet, despite the declines, the average return to obtaining a business or science degree continues to be quite large.

The reduction in the wage returns to major relative to the OLS model is a result of sorting into field of study based on the expected underlying human capitals. Given that students have very imprecise signals of their verbal human capital, most of the sorting is related to math human capital. In particular, the average math human capital for those who choose science in our simulated data is 0.042, which is 0.25 standard deviations above the unconditional mean. The corresponding values for business and other majors are -0.019 (0.11 SD) and -0.005 (0.03 SD). The high returns to math human capital in the science field combined with the sorting on math skill helps explain the large observed wage gap between science and the residual major group.

In the remaining columns of Table 7, we report the average returns for those who selected a business, science, or residual major, respectively. The average returns to a business or science degree do not vary much across the population. Science majors actually have the highest return to obtaining a business degree relative to the other major, reflecting the large return to math human capital for business majors. The main conclusion to draw from Table 7 is that there are real differences in the monetary returns to majoring in business or science. However, individuals continue to major in the residual category, reflecting the fact that the non-monetary preferences over field of study and job type play a large role in individual schooling choices.

Within our framework, we are unable to separate the effect of human capital on major choices through preferences versus future returns. However, both Arcidiacono (2004) and Beffy et al. (2009) find that much of the sorting into technical fields is related to preferences as opposed to expected income.

The reduced form evidence presented in Section 2 indicates that not only do individuals majoring in science related fields earn a significant wage premium, but also that this premium varies greatly according to the type of job the individual ultimately chooses. The evidence from Table 7 indicates that the average return to science is smaller than the reduced form evidence indicates. Table 8 examines whether the wage returns associated with working in a related job are robust to sorting based on unobserved skill.

The first column of Table 8 presents the OLS estimates of the returns to working in a related job across all fields estimated on the simulated data.[24] Note that the numbers in Table 8 are not the coefficient estimates themselves, but rather the total returns which require adding the business and science coefficients to the coefficient for the residual major group. The results are quite similar to those presented in Section 2, however, the OLS estimates are likely biased by both sorting into major and relatedness. One method for minimizing this bias is to incorporate individual fixed effects, a strategy we pursued in the reduced form analysis of Section 2. The second column of Table 8 lists the relatedness returns estimated on the simulated data when individual fixed effects are incorporated. The results are quite similar to the fixed-effect estimates presented in Table 4, an indication that our model is able to replicate the returns to working in a related job for those workers who are on the margin between working in a related and non-related job.

As noted in Section 2, the problem with the fixed effect approach is that the estimated returns for the marginal worker likely understate the returns for the average worker. This is borne out by the results in column 3 of Table 8, which presents the average return to working in a related job when there is no selection into major. In other words, these numbers reflect the population returns to working in a related job for each major. For all three majors, the average return to working in a related job is higher than the corresponding fixed-effect estimate. When we compare the average returns to the OLS estimates, we find that for the other major group the return to relatedness is slightly higher than the OLS results would suggest, while it is

[24]In each regression we also control for major, year, and graduate degree effects.

lower for business and science majors. In particular, the return to a related job for students choosing science drops from 0.28 in the OLS regression to 0.24 in the structural human capital model. The difference in the returns to relatedness for the other major relative to the OLS estimates arises from the fact that high verbal human capital individuals are more likely to work in related jobs. The loading of verbal human capital in the wage equation for other majors is large and negative for non-related jobs and negative but smaller for related jobs. Essentially, individuals with low verbal skills tend to work in unrelated jobs and actually earn more then they would have in a related job, shrinking the gap in wages between related and non-related jobs in the OLS regression. The reduction in the return to relatedness for science majors relative to OLS is instead driven by positive sorting into related jobs based upon math human capital.

Columns 4 and 5 of Table 8 then consider the same returns, but condition on the individuals actual choice of relatedness. Not surprisingly we see that the group that chooses to work in related jobs tends to benefit more financially. This is not true, however, for business majors, where the return to relatedness is slightly smaller among those choosing to work in related jobs. Again, preferences over job types likely induce some workers to choose related jobs even though they could earn more in non-related jobs.

The final three columns of Table 8 present the average returns to relatedness conditional on major choice. Overall, the patterns in the returns to relatedness are quite similar to those presented in columns 3 through 5. There are small increases in the returns for science majors, a result of the fact that high math human capital students tend to sort into the science field and high math human capital individuals benefit the most from working in a related job.

The results presented in Table 7 indicate that the return to college major is significantly smaller then what is observed in the OLS regressions. However, Table 8 indicates that the returns to working in a related job are quite similar to those produced by the OLS regressions. The sizeable returns to working in a related job indicates that the specificity of human capital within major is a distinct feature of the data. In the next section we investigate further the specificity of schooling human capital.

5.2.2 The Importance of the Specificity of Human Capital

As a final exercise, we decompose the *within-major* variation in wages related to human capital into general and specific components.[25] Recall that the wage equation is given by

$$\ln w_{r,f^*,t} = p_{r,f^*,c} + p_{r,f^*,m} H_{m,0} + p_{r,f^*,v} H_{v,0}.$$

We view the wage variation that arises strictly through $H_{m,0}$ and $H_{v,0}$ as variation in the general skills that individuals bring to the labor market. In contrast, the variation in wages due to p is related to how math and verbal general skills apply to a specific type of job.[26]

The variation in wages related to general human capital can be pinned down by fixing the p's for each major using a weighted average of the related and non-related coefficients. We use the predicted proportion of individuals within each major working in related jobs as our weight. Combining these coefficients with the within-major variation in human capital we can easily generate the overall variation in wages related to general human capital. Table 9 provides the decomposition of wages into initial human capital and other components, followed by a decomposition of the impact of initial human capital into its general and specific components. We find that initial human capital plays the largest role in the science field, explaining 33% of the overall variation in wages. Additionally, we find that variation in math and verbal human capital alone is responsible for 95%, 76%, and 91% of the overall variation in wages related to human capital for business, science, and other majors. Thus, not only does human capital play the largest role in the science field, but the specificity of its application is also large relative to business or other majors.[27] The results in Table 9 also complement the evidence

[25] The initial human capitals and their coefficients are responsible for approximately a third of the overall wage variation. The remaining variation arises through the graduate degree and year dummies, as well as the measurement error. Obtaining a graduate degree and/or accumulating additional work experience certainly boosts human capital. Thus, the wage variation arising from human capital is most certainly larger than what is presented here.

[26] Notice that we only explore the within major variation in wages when decomposing wage variability into its general and specific components. However, the between major variation could potentially be decomposed into general and specific components given that the price of human capital differs by major. We are unable to do this since we cannot decompose the cross-major variation in the wage constants into specific and general components. This is a consequence of estimating the reduced-form parameters of the model rather than the structural parameters governing the accumulation of human capital. Therefore, the results of this section can be seen as a lower bound of the importance of specific human capital.

[27] It should be noted that the residual major group has the greatest overall variability in wages. Therefore,

on the returns to relatedness presented in Table 8. Science majors have the highest returns to working in a related job and also the greatest dispersion in wages induced by the presence of specificity. In contrast, the specificity of human capital plays the smallest role in both returns and wage dispersion for business majors.

The existence of both human capital specificity and uncertainty about one's human capital has the potential to make the college major decision a risky one. Individuals who are uncertain about their underlying human capital may want to choose majors whose wages vary less with human capital. The specificity of human capital is an additional source of risk in the sense that even within major the returns to human capital can vary considerably. While we are unable to estimate directly the role that risk plays in the choice of major, we find evidence that in addition to taste, there is room for risk to help explain why so few students choose to major in science related fields. The fact that relatively few students also tend to choose a business degree is unlikely to be explained by risk, since as noted above the specificity of human capital plays the smallest role in the wage dispersion for business majors.

6 Conclusion

It is well documented that the returns to obtaining a college degree vary significantly across fields of study, with business and science majors earning a significant wage premium relative to all other fields. In this paper we illustrate that there also exists significant variation in wages within major according to the type of career an individual pursues. The question is whether the observed differences in wages are driven by selection, or if there truly differential returns both within and across field.

We estimate a structural human capital model that allows for sorting into major and job type based on observed and unobserved characteristics to determine whether the observed wage gaps are driven primarily by selection. Our findings indicate that selection plays a significant role in generating the observed wage gaps across major, particularly in the sciences. However, the large wage gaps across job type are instead explained by true differences in returns. Even

there is still much to learn about what generates such variability. The residual group is certainly the most heterogenous in terms of the underlying group of majors and we cannot rule out that a large fraction of what we impute as the effect of measurement error is in reality a dimension of heterogeneity that we do not account for.

when we account for selection into major, the returns to business and science majors are still economically significant. Thus we are still left with the question of why do individuals pursue less remunerative majors?

One reason individuals may not pursue a science degree is a lack of knowledge about the true returns. Future success in the labor market depends on the skills a worker accumulates and the type of job pursued. If individuals do not know precisely their skills when making their major choice, then they face two sources of uncertainty. First, there is a risk that the human capital students accumulate will be de-valued if they do not obtain a job related to their field of study. Second, upon obtaining a related job, individuals face an uncertain wage since they do not know their skill level exactly. This story is consistent with our finding that both dimensions of uncertainty may make science and math related majors unattractive relative to business or other majors. Human capital plays the largest role in wages for science majors, and the specificity of human capital is also most severe for graduates with a science degree. The impact of skill uncertainty on major choice has thus far not been considered when designing policies to increase the number of students majoring in science and technology fields. However, reducing risk by incrementing information might be more valuable than utilizing monetary incentives, such as those provided by the SMART grant.

While we find evidence of both human capital uncertainty and specificity related to the college major decision, we are unable to estimate precisely the effect risk has on individual decisions. Allowing for a flexible human capital framework and individual uncertainty leads to a computationally intensive model, and as a result, we do not separately estimate monetary and non-monetary preferences over major and working in a related job. However, this makes it difficult to convincingly estimate the curvature of the utility function, which is crucial for understanding the importance of risk. Yet, we believe that risk is an important piece of the puzzle for understanding college major choice. Future research should seek to estimate the effect of risk on major choice more directly, while also incorporating additional sources of risk not considered here, such as risk related to college dropout, the business cycle, and wage variability over the life cycle.

References

Altonji, Joseph G., "The Demand for and Return to Education When Education Outcomes are Uncertain," *Journal of Labor Economics*, 1993, *11* (1), pp. 48–83.

Arcidiacono, Peter, "Ability sorting and the returns to college major," *Journal of Econometrics*, 2004, *121* (1-2), 343–375.

Becker, Gary S., "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 1962, *70* (5), pp. 9–49.

Beffy, Magali, Denis Fougre, and Arnaud Maurel, "Choosing the Field of Study in Post-Secondary Education: Do Expected Earnings Matter?," *Forthcoming in Review of Economics and Statistics*, 2009, (4127).

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman, "2001 Lawrence R. Klein Lecture Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice," *International Economic Review*, 05 2003, *44* (2), 361–422.

Chant, D., "On Asymptotic Tests of Composite Hypotheses in Nonstandard Conditions," *Biometrika*, 1974, *61* (2), 291–298.

Cunha, Flavio, James J. Heckman, and Salvador Navarro, "Separating Heterogeneity from Uncertainty in Life Cycle Earnings," *Oxford Economic Papers*, 2005, *57* (2), 191–261.

Gathmann, Christina and Uta Schnberg, "How General Is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 2010, *28* (1), 1–49.

Guvenen, Fatih and Tony Smith, "Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach," 2010. Working Paper.

Keane, Michael P and Kenneth I Wolpin, "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, *105* (3), 473–522.

Kennan, John and James R. Walker, "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 2011, *79* (1), 211–251.

McCall, Brian P., "Occupational Matching: A Test of Sorts," *Journal of Political Economy*, 1990, *98* (1), pp. 45–69.

Montmarquette, Claude, Kathy Cannings, and Sophie Mahseredjian, "How do young people choose college majors?," *Economics of Education Review*, 2002, *21* (6), 543–556.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "The Allocation of Talent: Implications for Growth," *The Quarterly Journal of Economics*, 1991, *106* (2), pp. 503–530.

Navarro, Salvador, "Using Observed Choices to Infer Agent's Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance," 2011. Working Paper 2011-8 CIBC Working Paper Series.

Neal, Derek, "Industry-Specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, 1995, *13* (4), pp. 653–677.

Nielsen, Helena Skyt and Annette Vissing-Jorgensen, "The Impact of Labor Income Risk on Educational Choices: Estimates and Implied Risk Aversion," 2010. Working Paper.

Oi, Walter Y., "Labor as a Quasi-Fixed Factor," *Journal of Political Economy*, 1962, *70* (6), pp. 538–555.

Parent, Daniel, "Industry?Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics," *Journal of Labor Economics*, 2000, *18* (2), pp. 306–323.

Pavan, Ronni, "Career Choice and Wage Growth," *Journal of Labor Economics*, 2011, *29* (3), pp. 549–587.

Poletaev, Maxim and Chris Robinson, "Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984-2000," *Journal of Labor Economics*, 2008, *26* (3), 387–420.

Self, S.G. and K. Liang, "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions," *Journal of the American Statistical Association*, 1987, *82* (398), 605–610.

Silos, Pedro and Eric Smith, "Human Capital Portfolios," 2012. Working Paper.

Sullivan, Paul, "A Dynamic Analysis Of Educational Attainment, Occupational Choices, And Job Search," *International Economic Review*, 2010, *51* (1), 289–317.

Yamaguchi, Shintaro, "Tasks and Heterogeneous Human Capital," *Forthcoming in Journal of Labor Economics*, September 2011, (2011-06).

Table 1: Schooling Statistics

Overall		**by Major**			
			Business	Science	Other
Age in 1993	23.31	SAT Math	525	596	525
	(1.48)		(113)	(106)	(111)
SAT Math	542	SAT Verbal	450	496	474
	(116)		(98)	(105)	(102)
SAT Verbal	473	Business GPA	3.14	3.07	2.71
	(103)		(0.44)	(0.75)	(0.78)
		Business Credits	37.04	2.91	5.70
% Business Major	26.12		(18.98)	(7.55)	(12.16)
% Math/Science Major	24.66	Science GPA	2.75	3.23	2.62
			(0.65)	(0.47)	(0.7)
% Other Major	49.22	Science Credits	20.01	78.43	17.93
			(12.61)	(29.34)	(17.24)
% Graduate Degree	27.34				
		Other GPA	2.89	3.14	3.23
Both Parents Foreign Born	0.08		(0.51)	(0.53)	(0.45)
		Other Credits	43.00	29.34	62.69
			(25.51)	(19.88)	(33.47)
		Both Parents Foreign Born	0.08	0.13	0.04

Table 2: Labor Market Statistics

	Overall	by Major: Business	Science	Other
Job is Related to Major Field of Study	70.46	78.35	79.81	61.01
Log(Annualized Salary)	10.52	10.59	10.65	10.41
	(0.62)	(0.60)	(0.57)	(0.65)
Log(Annualized Salary) if Related	10.57	10.59	10.70	10.45
	(0.59)	(0.58)	(0.52)	(0.62)
Log(Annualized Salary) if NOT Related	10.40	10.56	10.42	10.34
	(0.69)	(0.66)	(0.69)	(0.69)

Table 3: Returns to College Major

Dep. Var.	Log(Inc)	Log(Inc)	Log(Inc)	Log(Wage)
Business	0.191*	0.181*	0.185*	0.162*
	(0.024)	(0.024)	(0.023)	(0.021)
Science	0.234*	0.210*	0.215*	0.198*
	(0.021)	(0.022)	(0.022)	(0.021)
SAT Math/100		0.045*	0.042*	0.044*
		(0.011)	(0.011)	(0.010)
SAT Verbal/100		-0.040*	-0.045*	-0.039*
		(0.013)	(0.014)	(0.013)
Major Specific GPA			0.070*	0.065*
			(0.020)	(0.019)
Year/Graduate Degree Effects	Y	Y	Y	Y
N	4,927	4,927	4,927	4,927

*,** Indicate a coefficient is statistically significant at a 5 and 10% significance level. Heteroskedastic-robust standard errors reported in parentheses.

Table 4: Returns to College Major Accounting for Utilization

Dep. Var.	Log(Inc)	Log(Inc)	Log(Inc)	Log(Inc)
Business	0.143*			0.142*
	(0.051)			(0.050)
Science	0.019			0.023
	(0.047)			(0.047)
Job is Related to Studies	0.063**	0.019	0.048	-0.346*
	(0.035)	(0.030)	(0.030)	(0.165)
Business*Job is Related to Studies	0.039	0.096**	0.003	0.042
	(0.058)	(0.055)	(0.048)	(0.056)
Science*Job is Related to Studies	0.229*	0.227*	0.118*	0.223*
	(0.051)	(0.049)	(0.054)	(0.052)
SAT Math/100	0.041*	0.031*		0.030
	(0.011)	(0.011)		(0.026)
SAT Verbal/100	-0.042*	-0.033*		-0.047
	(0.013)	(0.012)		(0.032)
Major Specific GPA	0.064*	0.085*		0.000
	(0.020)	(0.019)		(0.093)
SAT Math/100*Job is Related to Studies				0.015
				(0.028)
SAT Verbal/100*Job is Related to Studies				0.007
				(0.034)
Major GPA*Job is Related to Studies				0.093*
				(0.046)
Year/Graduate Degree Effects	Y	Y	Y	Y
Worker Fixed Effects	N	N	Y	N
Detailed Major Effects	N	Y	N	N
N	4,927	4,927	4,927	4,927

*,** Indicate a coefficient is statistically significant at a 5 and 10% significance level. Heteroskedastic-robust standard errors reported in parentheses.

Table 5: Model Validation: Summary Statistics

	Business Majors		Science Majors		Other Majors	
	Data	Model	Data	Model	Data	Model
% Selected	26.1	26.1	24.7	24.9	49.2	49.1
SAT Math	529	539	596	582	525	527
Verbal Residual	-12.0	-10.9	-7.9	-14.8	10.4	13.7
Business GPA	3.15	3.16	3.11	3.14	2.78	2.83
Science GPA	2.76	2.77	3.23	3.21	2.64	2.67
Other GPA	2.91	2.92	3.16	3.14	3.24	3.25

Table 6: Model Validation: Wage Regressions

	Data	Model	Data	Model	Data	Model
Business	0.191	0.187	0.185	0.190	0.143	0.149
Science	0.234	0.228	0.214	0.222	0.018	0.044
Job is Related					0.064	0.057
Related x Bus.					0.038	0.041
Related x Sc.					0.228	0.214
SAT Math/100			0.019	0.013	0.020	0.012
SAT Residual Verbal/100			-0.043	-0.005	-0.040	-0.004
Major GPA			0.069	0.061	0.063	0.054
Year/Grad	Y	Y	Y	Y	Y	Y

Table 7: Returns to College Major

		Sample by Chosen Major			
	OLS	All	Business	Science	Other
Business	0.191	0.150	0.134	0.163	0.152
Science	0.234	0.123	0.121	0.122	0.124
Year/Grad controls	Y	Y	Y	Y	Y

Table 8: Returns to Relatedness by Major

			No Selection into Major			Sorting into Major		
	OLS	FE	All	Rel=0	Rel=1	All	Rel=0	Rel=1
Other	0.062	0.065	0.089	-0.009	0.146	0.086	-0.012	0.146
Business	0.102	0.050	0.058	0.060	0.057	0.049	0.044	0.050
Science	0.279	0.201	0.244	0.088	0.295	0.289	0.127	0.333
Major/Year/Grad	Y	Y	Y	Y	Y	Y	Y	Y

Table 9: Log-Wage Decomposition

	Business	Science	Other
Wage Variability	0.352	0.321	0.417
Fraction Related to Initial Human Capital	27.8%	32.7%	27.0%
General Skill	95.4%	76.3%	91.4%
Specificity	4.6%	23.7%	8.6%

A Parameter Estimates - Uncertainty

Table A.1: Skill Distributions

Standard Deviation Math Human Capital, σ_{H_m}	0.171	(0.024)
Standard Deviation Verbal Human Capital, σ_{H_v}	0.020	(0.005)
Standard Deviation Math Scholastic Ability, σ_{ν_m}	0.505	(0.025)
Standard Deviation Verbal Scholastic Ability, σ_{ν_v}	0.194	(0.015)

Table A.2: Major and Relatedness Choice Parameters

Major Choice			Relatedness Choice		
Constant, Business	2.639	(0.307)	Constant, Business	1.755	(0.138)
E($H_{m,0}$), Business	44.496	(7.086)	$H_{m,0}$, Business	3.129	(0.448)
E($H_{v,0}$), Business	-3014.164	(175.771)	$H_{v,0}$, Business	40.411	(4.381)
Potential GPA, Business	0	-	Graduate, Business	1.739	(0.260)
Parents Foreign, Business	2.395	(0.444)	Lag Related, Business	1.008	0.075
Constant, Science	0.316	(0.253)	Constant, Science	1.224	(0.111)
E($H_{m,0}$), Science	49.067	(6.209)	$H_{m,0}$, Science	4.624	(0.943)
E($H_{v,0}$), Science	-3618.706	(51.552)	$H_{v,0}$, Science	6.042	(5.346)
Potential GPA, Science	2.924	(0.148)	Graduate, Science	0.874	(0.240)
Parents Foreign, Science	3.074	(0.218)	Lag Related, Science	1.754	(0.122)
Potential GPA, Other	0	-	Constant, Other	0.494	(0.071)
			$H_{m,0}$, Other	4.465	(0.721)
Own Share	5.893	(0.200)	$H_{v,0}$, Other	21.004	(2.586)
			Graduate, Other	1.639	(0.149)
			Lag Related, Other	2.047	(0.074)
			1997 Dummy	-0.814	(0.073)
			2003 Dummy	-1.458	(0.078)
			Share Related	6.521	(0.788)

Table A.3: Wage and Measurement Parameters

Wage Coefficients			Measurement Coefficients		
$H_{m,0}$, Non-Related, Other	0.594	(0.330)	SAT Math Constant	5.444	(0.019)
$H_{m,0}$, Non-Related, Business	2.178	(0.323)	SAT Verbal Constant	0.001	(0.014)
$H_{m,0}$, Non-Related, Science	0	-			
$H_{m,0}$, Related, Other	1.268	(0.193)	Constant, Business	3.307	(0.029)
$H_{m,0}$, Related, Business	1.532	(0.127)	$A_{m,0}$, Business	0.611	(0.036)
$H_{m,0}$, Related, Science	1.323	(0.184)	$A_{v,0}$, Business	0.847	(0.100)
$H_{v,0}$, Non-Related, Other	-18.557	(0.909)	$\eta_{O,B}$	-0.492	(0.015)
$H_{v,0}$, Non-Related, Business	-8.576	(2.439)	$\eta_{S,B}$	-0.284	(0.026)
$H_{v,0}$, Non-Related, Science	-19.239	(1.719)			
$H_{v,0}$, Related, Other	-10.337	(1.064)	Constant, Science	2.676	(0.044)
$H_{v,0}$, Related, Business	-5.442	(1.236)	$A_{m,0}$, Science	1.115	(0.026)
$H_{v,0}$, Related, Science	-9.251	(0.946)	$A_{v,0}$, Science	0.038	(0.104)
Constant, Other	9.964	(0.028)	$\eta_{O,S}$	0.193	(0.013)
Constant, Business	10.149	(0.027)	$\eta_{B,S}$	0.247	(0.012)
Constant, Science	9.966	(0.050)			
Related, Other	0.090	(0.016)			
Related, Business	0.058	(0.013)	Constant, Other	3.058	(0.024)
Related, Science	0.241	(0.019)	$A_{m,0}$, Other	0.389	(0.028)
Graduate, Other	0.083	(0.030)	$A_{v,0}$, Other	1.884	(0.058)
Graduate, Business	0.162	(0.024)	$\eta_{B,O}$	0.121	(0.011)
Graduate, Science	0.048	(0.028)	$\eta_{S,O}$	0.190	(0.015)
1997 Dummy	0.346	(0.014)	σ_{u_m}	0.985	(0.015)
2003 Dummy	0.459	(0.011)	σ_{u_v}	0.769	(0.009)
			σ_{u_B}	0.423	(0.009)
			σ_{u_S}	0.344	(0.020)
			σ_{u_O}	0.263	(0.014)