10-30-2019 2:00 PM

Essays on College Majors and Skills

Yuki Onozuka, *The University of Western Ontario*

Supervisor: Robinson, Chris, *The University of Western Ontario*
Co-Supervisor: Lochner, Lance J, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Economics

Abstract

My thesis consists of three chapters that study relationships between college majors and multi-dimensional skills.

Chapter 2 examines the sources of wage penalties for working outside one's major field of study. Previous papers show that workers in a job which is unrelated to their major field of study tend to earn significantly lower wages than those in a related job. I use the 1993 National Survey of College Graduates and the O*NET to divide the sources of wage penalty into the levels of basic skills required in a job and the mismatch in major-specific knowledge. I find that the average wage penalty is 9% after conditioning on individual characteristics, such as degree type and field of study. Around 45% of the wage penalty stems from differences in the required levels of basic skills between related and unrelated jobs. I also find that the results are heterogeneous across degree types and fields of study. A mismatch in major-specific knowledge has a large effect on wages of workers with an advanced or specialized degree and on those who majored in Computer and Math Sciences or Engineering.

Chapter 3 estimates skill growth during college by major using the National Longitudinal Survey of Youth 1997 (NLSY97) and the O*NET. To capture both the type and quantity of accumulated skills, I assume that each major increases a general cognitive skill and a major-specific skill. I further allow for individual heterogeneity in skill growth. I take a task-based approach and use occupation choice to estimate skill growth in general cognitive skill. To deal with noisy skill measurements and endogeneity, a dynamic factor model is constructed. The results show a substantial growth of general cognitive skill in all majors, but with large differences across majors. I find different effects of pre-college skill levels on skill growth by major, but the differences are not large. The contribution of major-specific skill growth to wage growth is small compared to that of general cognitive skill growth.

Chapter 4 examines what skills are most closely associated with male college workers becoming managers from the perspective of human capital theory, with a focus on cognitive and social skills. I first construct a social task intensity measure using the American Community Survey 2010-2017 and the O*NET and document that management jobs tend to have high levels of social and cognitive task intensity. I then use the NLSY97 and analyze the transition patterns into management jobs. Most workers start their careers in non-management jobs, but workers who become managers relatively quickly tend to have jobs involving intense social tasks in an early stage of their careers, which may lead to a greater increase in their social skill. Business & Economics majors are more likely to become managers, but the results suggest that this mostly stems from skills other than social skill.

Keywords: College majors, Multi-Dimensional skills, Task-Based approach

Summary for Lay Audience

College education increases individual's productivity in the labour market. However, what students learn substantially depends on their chosen major. My thesis consists of three chapters that empirically study labour-market correlates and consequences of college majors using US detasets.

Chapter 2 examines why workers in jobs that are unrelated to their college major tend to earn lower wages than those in related jobs. Will workers suffer from low wages in any type of jobs unrelated to their major? This question is important in what kind of jobs workers should have. I show that the wage penalties partially stem from the fact that unrelated jobs workers tend to hold require lower levels of general type of cognitive skills than related jobs workers tend to hold. Mismatch in major-specific type of skills is not as important as it looks. I also find heterogeneous results across degree types and fields of study.

Chapter 3 estimates what skills college students increase by major. I assume that each major increases a general cognitive skill and a major-specific skill. The literature mostly focuses on wage differences across majors, but this multi-dimensional skills framework will be helpful to understand differences across majors deeply. I show that increase in general cognitive skill is a key feature of college education. I find large differences in the skill growth across majors.

Chapter 4 examines what skills are most closely associated with male college workers becoming managers. I focus on social skill, which is the capacity to work with others to achieve goals, and cognitive skill. I first empirically document that management jobs tend to require higher levels of cognitive and social skills than non-management jobs. I then show that most workers start their careers in non-management jobs but that workers who become managers relatively quickly tend to have jobs that require a high level of social skill in an early stage of their careers, which may have a greater increase in their social skill. Business & Economics majors are more likely to become managers. The results suggest that this mostly stems from skills other than social skill.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Chris Robinson and Lance Lochner, for their support and guidance throughout the duration of my doctoral studies. I am particularly indebted to Chris Robinson for his patient guidance since my second-year paper project. I would also like to thank Salvador Navarro for his support.

I must acknowledge the financial support of Japan Student Service Organization (JASSO) throughout my first year to my fifth year at the University of Western Ontario. The completion of the dissertation would not have been possible without their generous support.

I also owe a great deal of thanks to my fellow students, particularly to my cohort, Samantha Black, Matthew George Carew, Yifan Gong, Zhuang Liu, Fidel Perez Macal, Yuxi Yao, and Sha Wang. They supported me mentally throughout my Ph.D. life at the University of Western Ontario.

Finally, I am deeply grateful to my family members. My parents, Nobutaka and Yoko Onozuka, have always supported my studies. I would also like to extend my thanks to my siblings and my grandparents for their support over many years.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Introduction

College graduation rates have been dramatically increasing the last several decades in many developed countries. Altonji et al. (2016) use US data and document that the share of individuals who earn a Bachelor's degree within each cohort has increased by 86% over the period 1985 to 2013. Traditionally, labour economists pay a lot of attention to education level, such as high school or college. However, as more people go to college, college majors are attracting attention. In a human capital framework, people can accumulate human capital or skills through education. College students choose their college major, and their course-taking behaviour substantially depends on their chosen major. Hence, human capital or skill accumulation of college students may significantly differ by college major.

My thesis consists of three chapters that study relations between college major and skills. I consider multi-dimensional skills. In Chapter 2, I examine the sources of wage penalties for working outside one's major field of study. Previous papers show that a substantial proportion of college workers hold a job that is not related to their major field of study and that these workers tend to earn significantly lower wages than those in a job that is related to their major field of study. This suggests that a substantial amount of human capital may be underutilized. Identifying the source factors that generate the wage penalty is important for understanding how students should choose a college major, what type of human capital students accumulate in college, and how the inefficient use of human capital can be reduced. I use the 1993 National Survey of College Graduates and the O*NET to divide the sources of wage penalty into the levels of basic skills required in a job and the mismatch in major-specific knowledge. The results show that the average wage penalty is 9% after conditioning on individual characteristics, such as degree type and field of study. Around 45% of the wage penalty stems from differences in the required level of basic skills between related and unrelated jobs. The results are heterogeneous across degree types and fields of study. A mismatch in major-specific knowledge has a large effect on wages of those with an advanced or specialized degree and those who majored in Computer and Math Sciences or Engineering.

The results in Chapter 2 provides some evidence on major specificity of skills, but worker's skills are taken as given and skill growth differences across college majors are not studied. In Chapter 3, I estimate skill growth during college by college major. There is a growing literature on differential wages across college majors, but few studies focus on skill growth by college major. If each major increases a common single type of skill and there is no heterogeneity in the wage returns to the skill across occupations, then wage differentials across majors will directly reflect skill growth differences across majors. However, as suggested in Chapter 2, a one-dimensional skill framework may be inappropriate for college majors. Courses college students take vary significantly by their college major, and the course differences will result in students accumulating different types and amounts of skills. A role of college major or differences across college majors will be understood better in a multi-dimensional skill framework. To capture both the type and quantity of accumulated skills, I assume that each major increases a general cognitive skill and a major-specific skill. I further allow for individual heterogeneity in skill growth. I take a task-based approach and use occupation choice to estimate skill growth in general cognitive skill. To deal with noisy skill measurements and endogeneity, a dynamic factor model is constructed. I estimate the model using the National Longitudinal Survey of Youth 1997 (NLSY97) and the O*NET. The results show substantial growth of general cognitive skill in all college majors, but with large differences across majors. I find different effects of pre-college skill levels on skill growth by major, but the differences are not large. The contribution of major-specific skill growth to wage growth is small compared to that of general cognitive skill growth.

Chapter 3 estimates skill growth during college, but it is silent on a worker's career after entering the labour market. In Chapter 4, I examine what skills are most closely associated with male college workers becoming managers from the perspective of human capital theory, with a focus on cognitive and social skills. I analyze the transition patterns into managerment and provide new support for the importance of social skill to becoming a manager. The literature states that social tasks are an important factor of management jobs, but no study actually constructs a data-based social task measure and empirically supports the statement. Hence, the first part of my analysis constructs a social task intensity measure using the American Community Survey (ACS) 2010-2017 and the O*NET and documents that management jobs tend to have high levels of social and cognitive task intensity. Although the ACS has an advantage in its large sample size, it is a cross-sectional dataset. Hence, the second part of my analysis uses the NLSY97 and analyzes the transition patterns into management jobs. Most workers start their careers in non-management jobs, but workers who become managers at age 30 tend to have jobs involving intense social tasks in an early stage of their careers. The workers may increase their social skill at a relatively high rate in these jobs, and, thus, become managers relatively quickly. Cognitive task intensity after conditioning on social task intensity does not have the power to predict whether a worker becomes a manager at age 30. These results suggest that increasing social skill is essential for becoming a manager and that being good at only cognitive tasks is unlikely to be enough for management. Business & Economics majors are more likely to become managers, but the result suggests that this mostly stems from skills other than social skill.

Bibliography

Altonji, J. G., P. Arcidiacono, and A. Maurel, 2016, The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects, *in* Handbook of the Economics of Education: Elsevier B.V., 7, 305–396.

Chapter 2

Basic Skills or Major-Specific Knowledge? Sources of Wage Penalties for Working Outside the Major Field of Study

2.1 Introduction

In a human capital framework, education brings high wages because people can accumulate human capital through education. This, however, implies that if accumulated human capital is not utilized well in a job, a worker may suffer from the penalty of having low wages. In order to earn high wages, workers would therefore need to take a job that requires the skills they possess.[1]

Unlike in high school or lower levels of education, students choose their major field of study in college and accumulate human capital according to that field of study. The accumulated human capital is expected to be well-utilized in certain occupations, and this leads to a strong relatedness between major fields of study and occupations after graduation. For example, human capital accumulated in engineering is designed to be useful for working as an engineer, and a majority of engineering majors become engineers after graduation.

However, this trend does not extend to everyone, and people might not obtain a job that is related to their major field of study. This could be due to various reasons, including labour market frictions and their preference against the field. Some college students might realize that they do not like their field enough to have a career after graduation. The proportion of the mismatch is dependent on how the term "unrelated jobs" is defined, but most of the literature report that 20%-30% of college workers are in a job that is mismatched with their major field of study. Furthermore, a further 10%-30% are in jobs that are weakly mismatched (somewhat related). In Abel and Deitz

[1] There is a large literature on vertical mismatch, such as overeducation and being overskilled. See Leuven and Oosterbeek (2011) for a survey.

(2015), who use the 2010 American Community Survey (ACS) and the CIC-SOC crosswalk provided by the National Center for Education Statistics, only 27% of college workers hold jobs which are directly related to their major field of study.

The large proportion of mismatch between major fields of study and jobs implies that a substantial amount of human capital may be underutilized. Previous papers argue that wages among workers in a job unrelated to their major field of study tend to be lower than those among workers in a job closely related to their major field of study (see, e.g., Robst (2007,b); Nordin et al. (2010), and Kinsler and Pavan (2015)). This wage penalty can be considered as a result of the mismatch in human capital between workers and jobs.

However, this does not necessarily mean that workers will suffer from a large wage penalty when in any type of job that is unrelated to their major field of study. Montt (2017) and Sellami et al. (2017) study the wage penalty that was associated with a field mismatch and a mismatch in education qualification. The results show that college workers will suffer from a large wage penalty of working outside their major field of study if they are overqualified.[2]

In this paper, I examine the sources of wage penalty of working outside one's major field of study by decomposing it into the levels of basic skills that would be required in one's job and the mismatch in major-specific knowledge. I use a worker's self-assessment relatedness measure as a relatedness measure. Basic skills are skills that are, to some extent, useful in any job. Examples of basic skills include math and English skills. I consider that every job can be performed using a combination of basic skills and that major-specific knowledge can help a worker to employ their basic skills effectively in cases where they are working in a job related to their major field of study.

Given that college jobs tend to require higher levels of basic skills than non-college jobs, papers by Montt (2017) and Sellami et al. (2017) are related to my study. However, since they use education qualification, they cannot consider heterogeneity within college jobs. Indeed, related college jobs and unrelated college jobs may differ in other than the requirement of major-specific knowledge. My paper considers two-dimensional basic skills, which allows me to capture differences across jobs in terms of the types of basic skills required in work.

Jobs that are related and unrelated to major field of study can differ from each other in the requirements pertaining to major-specific knowledge and levels of basic skills. Take, for an instance, the following four types of jobs: engineer, dishwasher, English teacher, and actuary. Out of the four aforementioned types of jobs, only the job of engineer can be classified as a related job for engineering majors because it requires engineering knowledge. The other three types of jobs will be considered to be unrelated to engineering majors, and the job of dishwasher, in particular, will require a much lower level of basic skills than the other two types of jobs. Consequently, one can

[2]My work is also, to some extent, related to Robst (2008), who shows that wage penalty for working in a non-college job is large if the job is not related to one's major field of study.

easily imagine that wage differences between engineers and dishwashers will be larger than those between engineers and English teachers or actuaries. Furthermore, the jobs of English teacher and actuary will differ from each other in terms of the types of basic skills required in work. While the job of English teacher would require a very high level of English skills, the job of actuary will require a very high level of math skills. Since math skills are highly required when studying engineers, wage differences between engineers and actuaries might be smaller than those between engineers and English teachers.

My paper also examines the heterogeneity across fields of study and degree types as well. Moreover, unlike Montt (2017) and Sellami et al. (2017), my paper explicitly controls for the levels of basic skills that are required in work. Through this approach, the remaining wage penalty is more likely to come from a mismatch in major-specific knowledge. The identification of the wage effect of a mismatch in major-specific knowledge is important when looking to provide an efficient allocation of workers or when providing career advice to college students. If major-specific knowledge is not a crucial factor when determining wage, then college students should be advised to avoid asking whether a job requires knowledge that is specific in their major field of study and instead focus on the intensity of the basic skills required in the work of a particular job. Furthermore, the student's college major choice would also be related to this topic. If major-specific knowledge is important to wage, students should be extremely careful about their preferences on the fields of study when choosing a major field of study, because switching fields after graduation might lead to low wages. This can be related to research which shows that preferences or tastes on fields of study play a significant role in students' college major choice (see, e.g., Arcidiacono (2004) and Wiswall and Zafar (2015)).[3]

I use the 1993 National Survey of College Graduates (NSCG) and the O*NET. The O*NET is used to construct measures of basic task intensity, which represent the levels of basic skills required in a job. In my analysis data, 15% of male workers and 17% of female workers report that their jobs are not related to their major field of study. The average wage penalty for working in an unrelated job that cannot be explained by differences in observed individual characteristics, such as degree type and field of study, between workers in closely related jobs and those in unrelated jobs, is 9% in both male and female cases. If the constructed task intensity measures are additionally controlled for, then the wage penalty decreases by around 45%, from 9% to 5%, for both males and females. With regard to the heterogeneity across degree types, workers with an advanced or a specialized degree tend to suffer from a large wage penalty, even when controlling for the

[3]Lemieux (2014) uses Canadian datasets and shows that college workers who hold a related job tend to earn higher wage than those who hold an unrelated job conditional on occupation. He considers this wage difference as an effect of occupation-field of study match on wage. However, his occupation classification is coarse compared to my measures of basic skills required in work. Moreover, his interest is not the wage penalty for working outside one's major field of study that is documented in the literature.

individual characteristics and the basic skills required in a job. There is also a wide variation across the degree fields that are studied. There is, on average, no wage loss stemming from mismatch in major-specific knowledge for workers who majored in Life and Related Sciences, Physical and Related Sciences, or Social and Related Sciences.

The rest of this paper is as follows. Section 2.2 describes my data. Section 2.3 explains my estimation models. Section 2.4 is for analysis. I conclude with section 2.5.

2.2 Data

I use two US datasets. I combine individual data from the 1993 NSCG with occupation data from the O*NET.

2.2.1 1993 NSCG

The individual data for this paper come from the 1993 NSCG. The target population of this survey is individuals who were under the age of 76, living in the US, and had a Bachelor's degree or higher as of the survey date. The sample for the 1993 NSCG is selected from the 1990 Census. The original sample size is 148,905, containing 87,649 males and 61,256 females. Some of the previous studies on wage penalty among college graduates, such as Robst (2007,b, 2008), use the 1993 NSCG. One of the advantages of this survey is that the survey asks the respondents about the relatedness between their principal job and their highest degree field. The respondents are allowed to choose between the options of closely related, somewhat related, or not related. I use this worker's self-assessment relatedness measure as my relatedness measure.

My analysis sample is composed of workers who were between 25 and 59 years old at the time of the survey.[4] After I restrict the sample to those who responded to the questions I use, the sample size is 102,114, with 64,782 males and 37,332 females.

2.2.2 O*NET

I combine the 1993 NSCG and the O*NET. The O*NET is used to construct measures of basic task intensity, which represents levels of basic skills required in certain jobs. In the O*NET, each occupation is characterized by hundreds of standardized measures. The measures describe the day-to-day aspects of the job, along with the qualifications and interests of the typical workers in the occupation. I use 31 elements that are related to cognitive skills from the O*NET. For each type of

[4]Robst (2007) does not put any age restriction, but I am concerned about selection of workers into the labour market based on unobservables because a substantial part of the people who were younger than 25 years old or who were older than 59 years old were not working.

standardized measure, the level is recorded with a range of 0–7, based on ratings by analysts. The score indicates the degree to which the descriptor is required or is needed for the performance of the designated occupation.

Following papers that take a task-based approach, I employ a Principal Component Analysis (PCA) to reduce the dimensions of the occupational classification scores from 31 to two. For the PCA, I use job information of workers who were aged between 25-59 in the 1993 NSCG and are included in my analysis sample. Table A.1 in the appendix presents the results of the PCA after a varimax rotation.[5] Based on the component loadings, I call the first component communication-related task intensity and the second component math-related task intensity. However, it is to be noted that while I name the two types of basic tasks in this manner, the tasks are constructed to be orthogonally distributed with each other. I standardize each type of task intensity to have a mean of 0 and a standard deviation of 1 over my analysis sample.[6]

2.2.3 Summary statistics

Table 2.1 reports the basic statistics of selected variables by relatedness, and the basic statistics of the other variables are reported in the appendix (see Table A.2 in the appendix for males and Table A.3 in the appendix for females). In my analysis sample, 15% of male workers and 17% of female workers report that their job is not related to their highest degree field, and 27% of male workers and 21% of female workers report that their job is somewhat related to their highest degree field. The mean annual wage varies by relatedness status. The mean annual wage among workers who are in a closely related job is found to be the highest, amounting to \$58,445 for males and \$41,849 for females. Workers in an unrelated job suffer from low wages, and the difference between a closely related job and an unrelated job is \$12,988 for males and \$7,098 for females.

People with an advanced degree are more likely to work in a related job. Most of the workers in an unrelated job have a Bachelor's as the highest degree, while people with a Bachelor's as a highest degree only account for approximately half of the workers in a closely related job. In contrast, the proportion of workers who have a more advanced degree is larger in closely related jobs than in unrelated jobs.

With regard to the major fields of study, I show the six categories of fields of study despite

[5] Varimax rotation is an orthogonal rotation, and this is the most commonly used rotation. When using this rotation, each component will tend to have either large or small loadings on any particular element, which makes it easier to identify each element with a single component.

[6] In order to utilize the occupation descriptors in the O*NET, I need to connect the job codes in the NSCG with those in the O*NET. The NSCG uses an original 3-digit job codes. Since the NSCG particularly focuses on those within the science and engineering workforce, the classification of science and engineering areas tends to be more detailed compared to other areas, such as health area. The number of jobs contained in the NSCG is 115. The job classification is, in general, rougher than that in the O*NET. When multiple job codes in the O*NET are found to correspond to a job code in the NSCG, I take the mean of the occupation descriptors among the multiple occupations.

using 31 minor categories in the regression analysis below. The correspondence table for the categories of fields of study is reported in Table A.4 in the appendix. For example, the three largest minor categories among Non-Science and Engineering (Non-S&E) degrees are management & administration, teaching, and health & related. Non-S&E degrees account for more than half of the cases for both males and females, regardless of the status. Engineering is the second largest group when looking at the field of study for males, while Social and Related Sciences is the second largest group for females. Some majors are unequally distributed across related statuses. For example, workers from the Engineering major are more likely to take a closely related job than those from Social and Related Sciences majors.

Large differences can be seen in levels of basic skills required in jobs. For both males and females, closely related jobs tend to be more demanding in both types of basic skills than somewhat related and unrelated jobs. Closely related jobs are higher in both types of basic task intensity by one standard deviation in comparison to unrelated jobs and by 0.5 standard deviation in comparison to somewhat related jobs.

I also look at the kernel density of each type of task intensity measure by relatedness using male workers. Figure 2.1 demonstrates the density of the communication-related task intensity, while Figure 2.2 illustrates the density of the math-related task intensity. In both types of task intensity, the distribution of closely related jobs (black solid line) is located to the right of those of somewhat related jobs and unrelated jobs, and the distribution of unrelated jobs (black dashed line) is located to the left of those of the other two types of jobs. Especially, in the case of communication-related task intensity, the distribution of unrelated jobs has a mass around the levels that are -3 to -2 standard deviations below the mean of college workers, and this probably corresponds to non-college jobs. The distributions of females show a similar pattern; similar in the sense that the distribution of closely related jobs is located to the right of the other two distributions, and the distribution of unrelated jobs is located to the left of the other two distributions in both types of task intensity.

2.2.4 Workers in unrelated jobs

Before showing regression results, I describe characteristics of workers in unrelated jobs through the use of graphs. At this point, I do not separate the analysis by gender. The results do not change much by separating by gender. As mentioned above, workers with a more advanced degree tend to hold a closely related job. Figure 2.3 shows the distribution of relatedness statuses by degree type. Among those whose highest degree is a Bachelor's, about half are in a closely related job and about 20% are in an unrelated job. Among workers with a Professional degree, the proportion of those working in a closely related job is higher than 90% and that of individuals working in an

Table 2.1: Basic statistics of selected variables by gender and by job relatedness

	Males			Females		
	Closely related	Somewhat related	Not related	Closely related	Somewhat related	Not related
Proportion	0.58	0.27	0.15	0.62	0.21	0.17
	(0.49)	(0.45)	(0.36)	(0.48)	(0.41)	(0.37)
Annual wage	58445.09	52228.38	45457.36	41849.04	39572.84	34751.19
	(31723.50)	(25464.39)	(26479.23)	(22420.80)	(20500.91)	(20398.65)
Degree type						
Bachelor's	0.48	0.70	0.78	0.51	0.71	0.81
	(0.50)	(0.46)	(0.41)	(0.50)	(0.45)	(0.39)
Master's	0.28	0.25	0.18	0.37	0.25	0.16
	(0.45)	(0.43)	(0.38)	(0.48)	(0.44)	(0.37)
Doctorate	0.10	0.04	0.02	0.05	0.03	0.01
	(0.30)	(0.19)	(0.15)	(0.22)	(0.16)	(0.11)
Professional	0.14	0.01	0.02	0.07	0.01	0.01
	(0.35)	(0.12)	(0.14)	(0.26)	(0.11)	(0.10)
Field of study						
Computer and Math Sciences	0.07	0.06	0.04	0.05	0.06	0.03
	(0.26)	(0.25)	(0.19)	(0.22)	(0.24)	(0.16)
Life and Related Sciences	0.05	0.04	0.06	0.05	0.06	0.05
	(0.23)	(0.21)	(0.24)	(0.23)	(0.23)	(0.22)
Physical and Related Sciences	0.05	0.05	0.04	0.02	0.02	0.02
	(0.21)	(0.22)	(0.21)	(0.13)	(0.16)	(0.13)
Social and Related Sciences	0.06	0.09	0.17	0.09	0.15	0.18
	(0.23)	(0.29)	(0.37)	(0.29)	(0.35)	(0.38)
Engineering	0.23	0.24	0.10	0.03	0.05	0.02
	(0.42)	(0.43)	(0.29)	(0.17)	(0.22)	(0.13)
Non-S&E degrees	0.54	0.51	0.59	0.75	0.66	0.71
	(0.50)	(0.50)	(0.49)	(0.43)	(0.47)	(0.45)
Task intensity						
Communication-related	0.38	-0.09	-0.75	0.10	-0.35	-0.82
	(0.83)	(0.99)	(1.24)	(0.77)	(0.95)	(1.02)
Math-related	0.34	0.11	-0.35	-0.17	-0.36	-0.71
	(0.98)	(1.09)	(0.98)	(0.73)	(0.97)	(0.93)
N	37,498	17,633	9,651	23,251	7,791	6,290

Standard deviations are in parentheses.

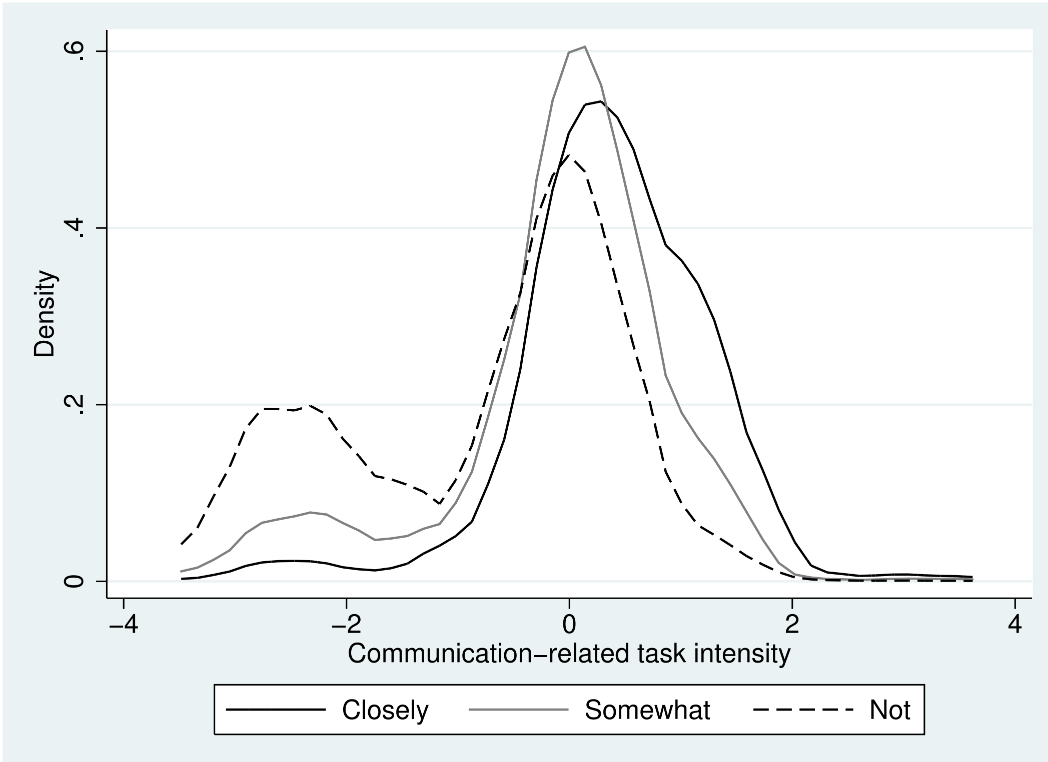

Figure 2.1: Kernel density of communication-related task intensity by relatedness: male

unrelated job is less than 5%. It is possible that individuals accumulate specific knowledge more than general skills in the process of earning a graduate or professional degree and are consequently more likely to work in a job that is related to the specific knowledge.

I also look at the ratios of relatedness statuses by major field of study. Figure 2.4 shows some differences across the fields of study. Computer and Math Sciences and Engineering show that there is a high percentage of closely related jobs, close to 60%, and a low percentage of unrelated jobs, around 10%. Compared to them, Life and Related Sciences, Physical and Related Sciences, and Non-S&E show a similarly high percentage of closely related jobs and a somewhat higher percentage of unrelated jobs. Their percentage of unrelated jobs is over 15%. Social and Related Sciences is clearly different from the others. The percentage of closely related jobs is only slightly above 40%. The percentage of unrelated jobs is close to 30%, which is almost the same as that of somewhat related jobs.

I then examine the differences in the distributions of degree types by field of study. Figure 2.5 shows some differences across the fields of study. Life and Related Sciences and Physical and Related Sciences show a relatively large proportion of doctorate. Non-S&E degrees show a larger proportion of professional degree than the other fields of study.

Given that a degree type will affect whether students work in a related job in future, the degree

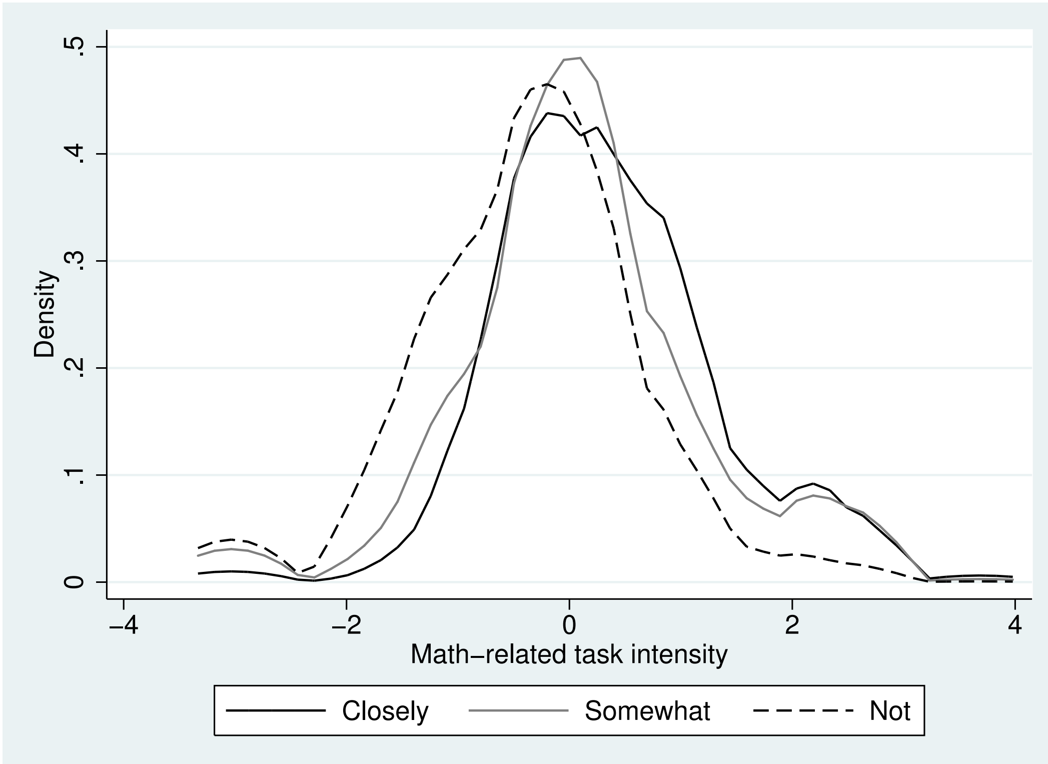

Figure 2.2: Kernel density of math-related task intensity by relatedness: male

type compositions across the fields might result in the differences in relatedness across the fields of study. Figure 2.6 shows the distribution of the relatedness status across the fields of study conditional on the highest degree being a Bachelor's degree. The distributions are quite different across Bachelor fields of study, and the tendency seen in Figure 2.4 is more obvious: Computer and Math Sciences and Engineering show a high percentage of closely related job and a low percentage of unrelated job; Physical and Related Sciences and Non-S&E show a similarly high percentage of closely related and a somewhat higher percentage of not related; Social and Related Sciences show a clearly different pattern: the percentage of unrelated job is the largest.

Figure 2.7 is for workers whose highest degree is a Master's degree. Although there are some differences, the distributions are similar across the fields of study as compared to the Bachelor's case. Hence, Figures 2.6 and 2.7 suggest that the differences in relatedness statuses across the fields of study mostly come from workers whose highest degree is a Bachelor's degree and that human capital acquired during an undergraduate study, along with the importance of working in a related job may depend on the field of study.

The description results suggest that controlling for degree types and fields of study is important for an analysis of wage penalty for working in an unrelated job. Furthermore, wage penalty might be heterogeneous across degree types and fields of study. Hence, the heterogeneity is examined in

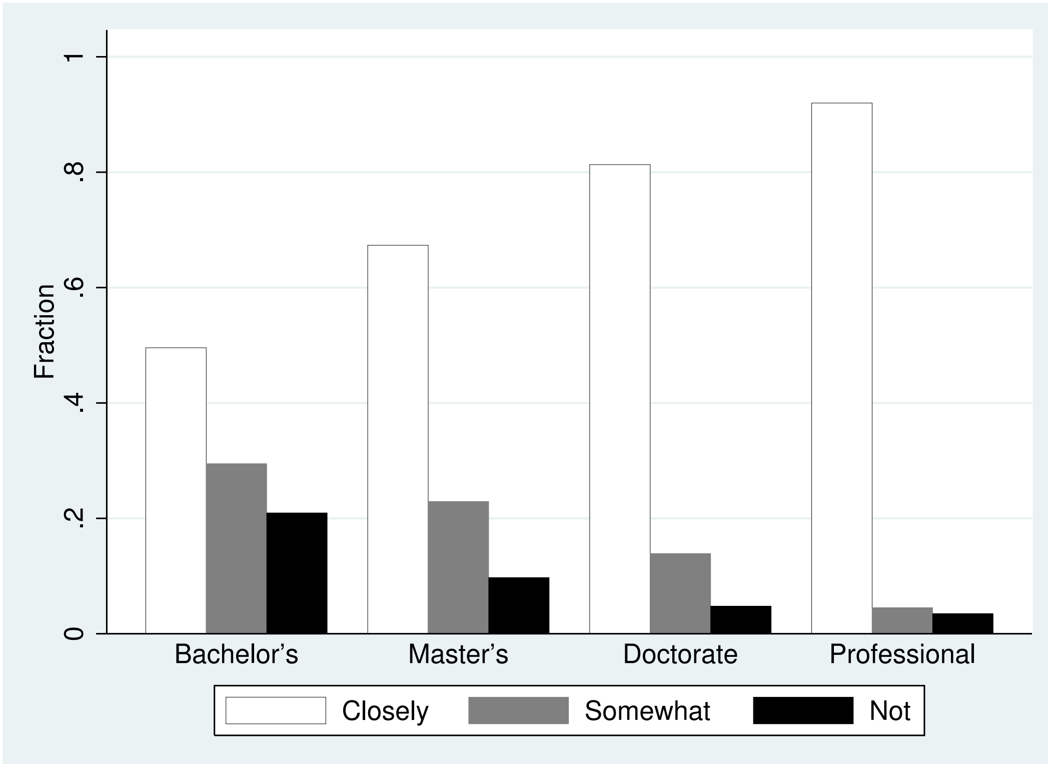

Figure 2.3: Distribution of relatedness status by degree type

subsections 2.4.2 and 2.4.3.

2.3 Estimation models

I estimate two types of regression models. Regressions are run separately by gender. The first regression for individual i is as follows:

$$\log w_i = \beta_{01} + \beta_{11} D_{1i} + \beta_{21} D_{2i} + x_i' \gamma_1 + \varepsilon_{1i}, \tag{2.1}$$

where w denotes annual wage, D_1 takes 1 if the relatedness status is not related and 0 otherwise, D_2 takes 1 if the relatedness is somewhat related and 0 otherwise, and ε_1 denotes an idiosyncratic error term. Following Robst (2007), vector x denotes individual characteristics, containing degree dummies, years of full-time and part-time experience, race dummies, a disabled dummy, a marital status dummy, training dummies, and the most recent degree field dummies.[7] The estimates of β_{11} and of β_{21} represent wage penalties incurred for working in an unrelated job and for working in a somewhat related job, respectively, after controlling for the individual characteristics. These would also correspond to the wage penalty documented in the previous studies, such as Robst (2007,b);

[7]The number of the categories is 31.

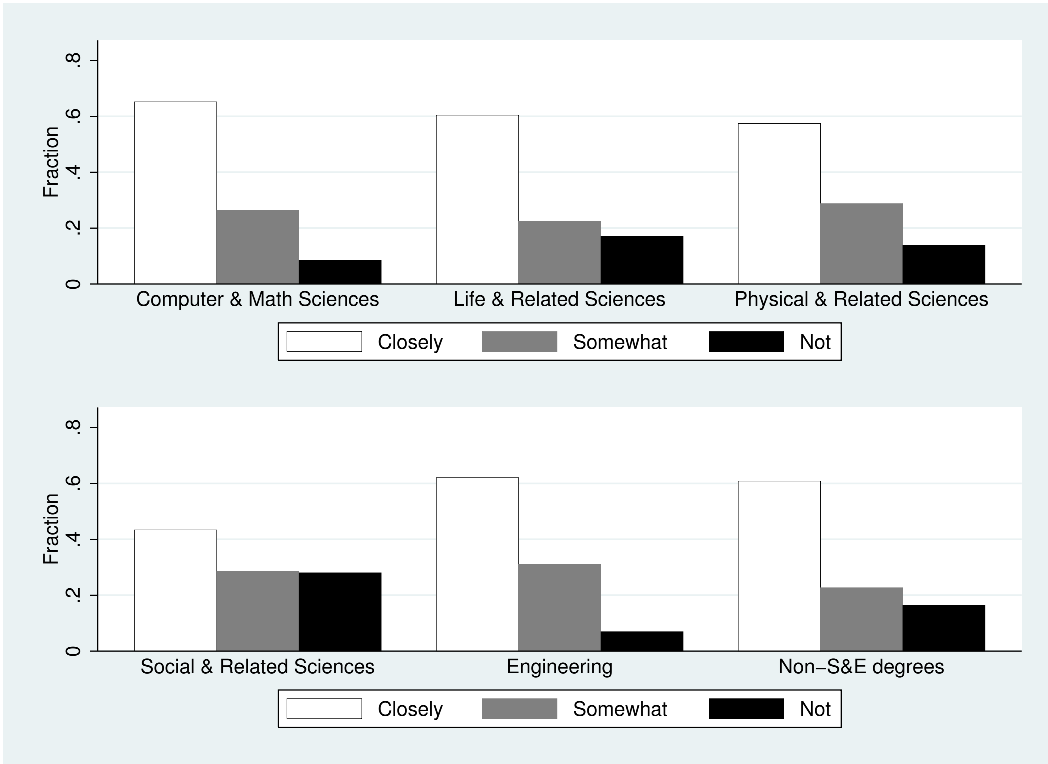

Figure 2.4: Distribution of relatedness status by field of study

Nordin et al. (2010), and Kinsler and Pavan (2015).

The second regression additionally includes the basic task intensity in one's occupation:

$$\log w_i = \beta_{02} + \beta_{12} D_{1i} + \beta_{22} D_{2i} + \beta_{32} \tau_i^c + \beta_{42} \tau_i^m + x_i' \gamma_2 + \varepsilon_{2i}, \tag{2.2}$$

where τ^c and τ^m represent the intensity of communication-related tasks and math-related tasks, respectively, and ε_2 is an idiosyncratic error term. The task intensity is constructed in the data section. The estimates of β_{12} and of β_{22} now indicate the wage penalties after controlling for the individual characteristics and the levels of the basic skills required in the worker's job. I consider that the wage penalties stem from a mismatch in major-specific knowledge.

My interests are the coefficients of D_1 and of D_2. I am interested in the extent to which the effects of differences in basic task intensity between closely related and unrelated jobs will be contained in the wage penalty that has been traditionally documented in previous papers, such as Robst (2007,b); Nordin et al. (2010), and Kinsler and Pavan (2015). I am also interested in the size of the effect of the mismatch in major-specific knowledge on wage.

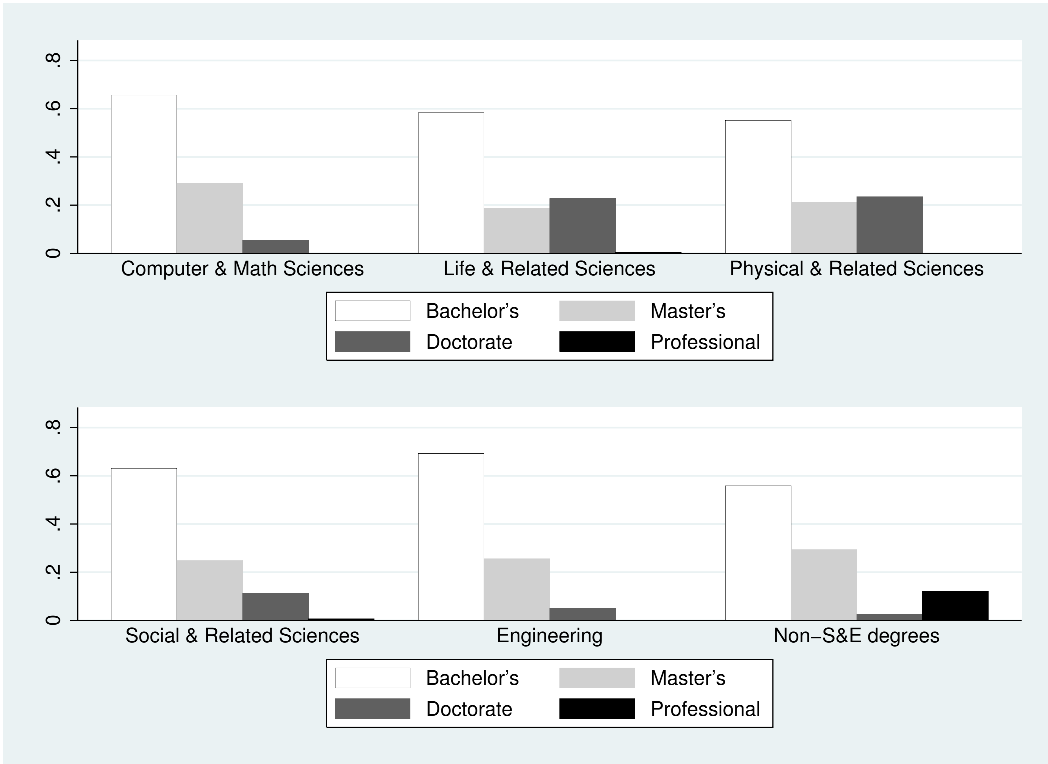

Figure 2.5: Distribution of degree type by field of study

2.4 Analysis

2.4.1 Main regression results

As explained in section 2.3, I am interested in how much the negative effects of working outside one's major field of study change by controlling for the basic task intensity measures in one's job, and I am also interested in the amount of the remaining negative effects. Columns 1 and 3 in Table 2.2 regress the log annual wage on dummies of the category of relatedness and the individual characteristics. Workers in an unrelated job tend to earn a significantly lower wage than those in a related job. Both males and females who are in an unrelated job earn 9% lower than those in a closely related job, on average. The negative effect of working in a somewhat related job is 2%. These negative effects are much smaller than the raw negative effects, which can be seen in Table 2.1. The difference in the mean raw wages between closely related and unrelated jobs is 27% for males and 21% for females. Even the difference between somewhat related and unrelated jobs is 9% for males and 6% for females. This significant reduction of the wage gaps in Table 2.2 is because degree types and fields of study are controlled.

The intensity of communication-related and math-related tasks are added in columns 2 and 4, and it can be seen that the wage penalties substantially reduce. The negative effect of working in

Table 2.2: Log wage regressions

	Male		Female	
Not related	-0.0888***	-0.0513***	-0.0883***	-0.0471***
	(0.0053)	(0.0055)	(0.0062)	(0.0064)
Somewhat related	-0.0218***	-0.0084*	-0.0193***	-0.0012
	(0.0041)	(0.0041)	(0.0055)	(0.0056)
Communication-related task intensity		0.0430***		0.0305***
		(0.0025)		(0.0037)
Math-related task intensity		0.0060*		0.0595***
		(0.0026)		(0.0041)
Constant	10.5496***	10.5487***	10.4497***	10.4344***
	(0.0116)	(0.0116)	(0.0152)	(0.0151)
Individual characteristics	Y	Y	Y	Y
N	64,782	64,782	37,332	37,332

Independent variable is log annual wage.

Standard errors are in parentheses.

$*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Individual characteristics are highest degree dummies, years of full-time and part-time experience, race dummies, disabled dummy, marital status dummy, training dummies, and most recent degree field dummies.

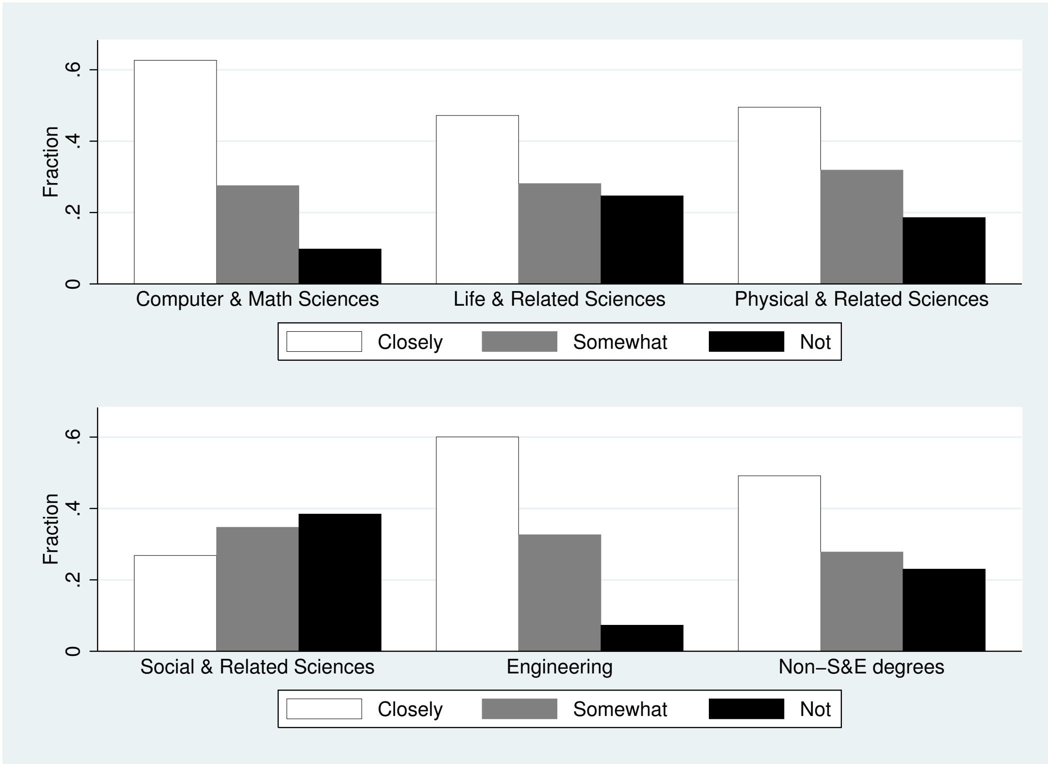

Figure 2.6: Distribution of relatedness status by field of study: Bachelor's

an unrelated job reduces by more than 40%, from 9% to 5% for both males and females. When it comes to somewhat related jobs, the effect is now only -1% for males and is insignificant for females. These estimation results imply that some of the negative effects of working outside the major field of study stem from the fact that unrelated jobs tend to require lower levels of basic skills than related jobs. Hence, if related and unrelated jobs that workers tend to take are similar in terms of required levels of basic skills, the wage differences between related and unrelated jobs will not be large.

Due to the data limitation, this paper does not take into account endogeneity of occupation choice. It is possible that workers in unrelated jobs are negatively selected. Individuals working in unrelated jobs might have been unable to find a closely related job because they had a lower level of major-specific knowledge. In this case, not all of the wage differences between closely related and unrelated jobs after controlling for basic task intensity stem from the mismatch in major-specific knowledge, and a part of the wage differences stem from differences in levels of major-specific knowledge between workers in closely related jobs and those in unrelated jobs.

Endogentiy of occupation choice should be addressed in future work, but the results from some previous studies might suggest that workers' skill levels or knowledge is not too important when estimating the average wage penalty for working outside a major field of study. Kinsler and Pavan

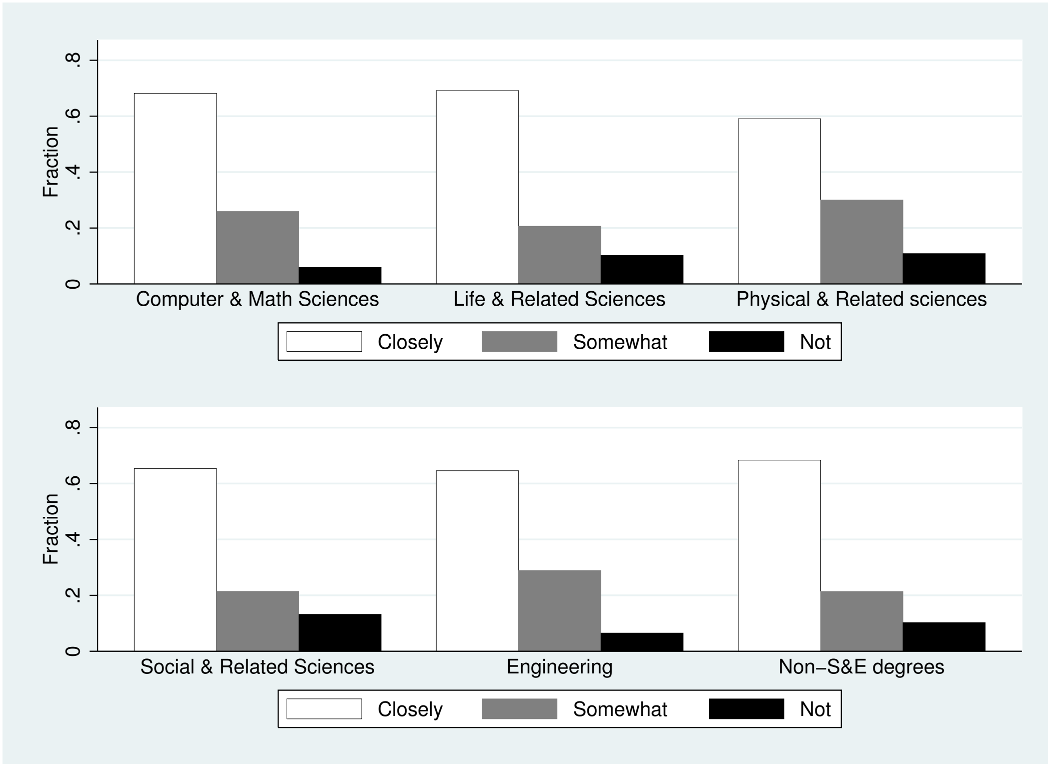

Figure 2.7: Distribution of relatedness status by field of study: Master's

(2015) use SAT scores, GPA, and an exogenous shifter of preference on working in a related job, and the wage penalty does not change much by accounting for workers' skills or knowledge.

2.4.1.1 College/Non-college jobs and intensity of basic tasks

As explained above, papers such as Montt (2017) and Sellami et al. (2017) study wage penalty associated with a mismatch by field and a mismatch by education qualification. Hence, this subsection examines how much the results change through the use of education qualification as a variable instead of the basic task intensity. Since the 1993 NSCG does not ask the respondents about education qualification for their job, I use the O*NET to define college and non-college jobs. In the O*NET, incumbents or occupational experts are asked for the required level of education to qualify for their jobs. I define college jobs as jobs in which more than half report that the required level of education is equal to or higher than Bachelor's degree. In my analysis sample, 80% of people are classified to work in a college-level job.

Table 2.3 regresses the log annual wage on a constant term, the individual characteristics, and a college job dummy. The coefficients of the somewhat related job dummy and of the unrelated job dummy are very similar with those in columns 2 and 4 in Table 2.2, which include the individual characteristics and the task intensity as regressors. This similarity comes from the fact that the distributions of the task intensity do not vary much across the relatedness statuses once jobs are restricted to college jobs (see Figures 2.8 and 2.9 for males). That is, unrelated college jobs and related college jobs that workers tend to hold are similar in the sense of the required levels of the basic skills. The similarity between the results does not necessarily mean that wage penalty for working in unrelated jobs will be homogeneous across college jobs. The non-zero coefficients of the task intensity measures in Table 2.2 suggest that wage penalty for working in unrelated college jobs would be large if the unrelated jobs workers tend to hold required lower levels of basic skills.

Table 2.3: Log wage regressions including college job dummy

	Male	Females
Not related	-0.0579***	-0.0485***
	(0.0054)	(0.0064)
Somewhat related	-0.0097*	-0.0000
	(0.0041)	(0.0056)
College job dummy	0.1063***	0.1216***
	(0.0048)	(0.0058)
Constant	10.4527***	10.3350***
	(0.0123)	(0.0161)
Individual characteristics	Y	Y
N	64,782	37,332

Independent variable is log annual wage.
Standard errors are in parentheses.
* $p < 0.05$,** $p < 0.01$, *** $p < 0.001$.
Individual characteristics are highest degree dummies, years of full-time and part-time experience, race dummies, disabled dummy, marital status dummy, training dummies, and most recent degree field dummies.

2.4.2 Heterogeneity across degree types

The next two subsections examine the heterogeneity of wage penalty. In this section, I examine the heterogeneity across degree types. Importance of basic skills and of major-specific knowledge can be different across degree types. For example, graduate and professional programs may encourage students to accumulate more specialized knowledge than an undergraduate program. In this case, the wage penalty for working outside one's major field of study among workers from these programs will be large and may not decrease by a drastic amount when controlling for the levels of basic skills required in one's job.

In order to examine the heterogeneity across degree types, I include the interactions between the relatedness dummies and highest degree type dummies as regressors. Table 2.4 summarizes

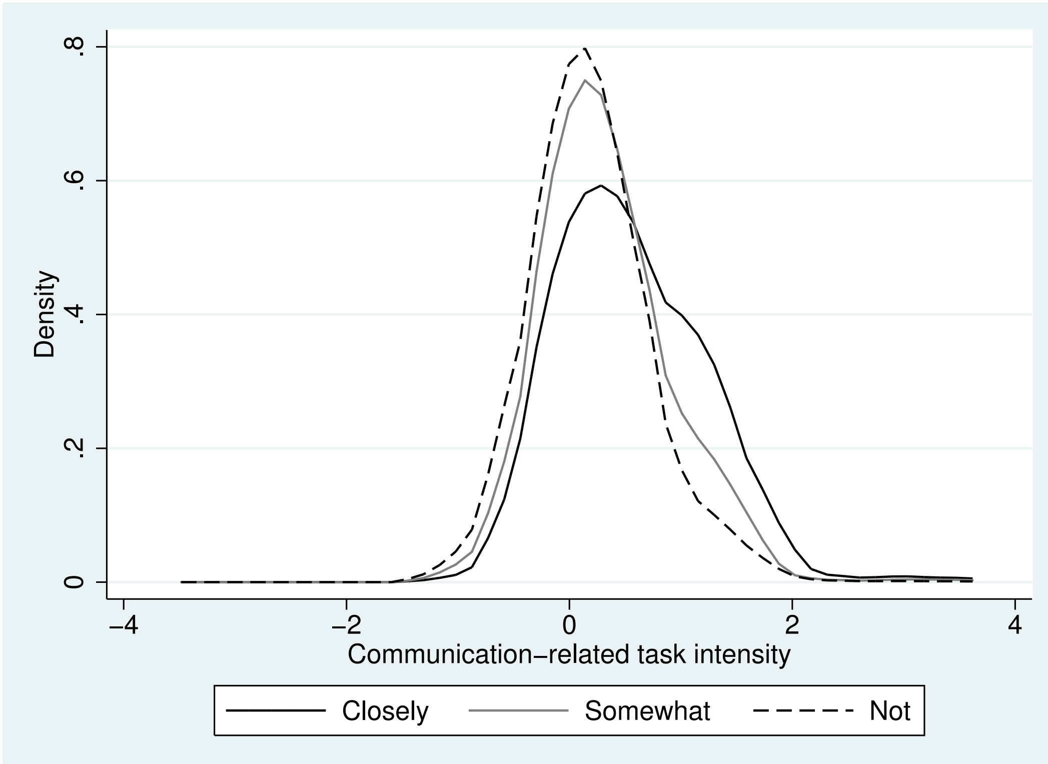

Figure 2.8: Kernel density of communication-related task intensity of college jobs by relatedness: male

the effects of unrelatedness on wages. For each type of degree, columns 1 and 3 show the wage differences between workers in a closely related job and those in an unrelated job after controlling for the individual characteristics. Professional degree shows a huge wage loss when working in an unrelated job, and this amount to 39% and 46% for males and females, respectively. The wage penalty in other types is around 8%.

In columns 2 and 4, the measures of basic task intensity are additionally controlled. The large wage loss for individuals with a professional degree who work in an unrelated job does not disappear and still amounts to about 33%. This implies that their large wage penalty mostly stems from underutilizing major-specific knowledge. This is plausible given that knowledge that is taught in a professional program is highly specialized and directly linked with certain occupations. On the other hand, in the case of the Bachelor's degree, the remaining wage loss is only 4% for males and females. The remaining wage penalty is smaller than that among workers with a Master's degree.

2.4.3 Heterogeneity across major fields of study

I also examine the heterogeneity across major fields of study. The quality of knowledge that students learn in college may depend on their field of study. In order to examine heterogeneity across major fields of study, I include the interaction terms between the relatedness dummies and

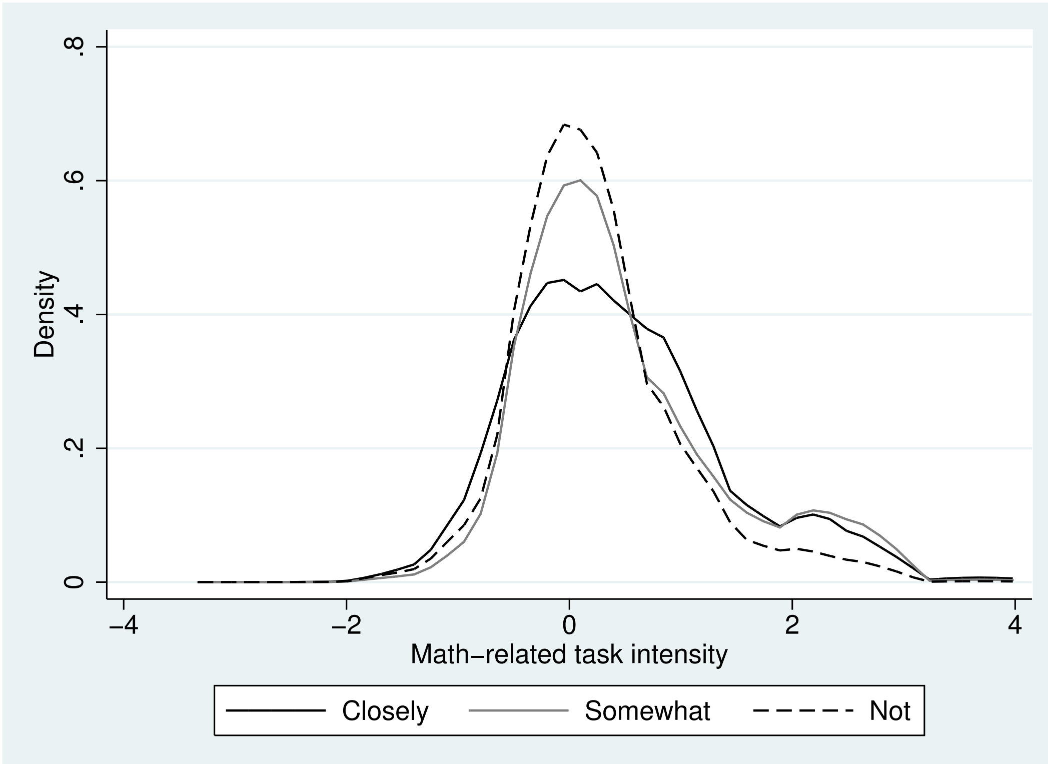

Figure 2.9: Kernel density of math-related task intensity of college jobs by relatedness: male

dummies of highest degree field of study. I use the six categories of fields of study in order to make the interpretation easier.

Table 2.5 reports the wage effects of working in an unrelated job as compared to working in a closely related job for each study field. Columns 1 and 3 show that wage losses significantly vary with the field. For males, the smallest is around 0% for Life and Related Sciences and Social and Related Sciences, while Engineering demonstrates 18% and Computer and Math Sciences shows 15% of wage loss, respectively. For females, Social and Related Sciences and Life and Related Sciences show small wage losses that are about 5%, while Computer and Math Sciences shows 31% and Engineering shows 22%.

The task intensity measures are added in columns 2 and 4. The wage loss of working outside a major field of study becomes smaller in every field. For both males and females, insignificant wage penalties are found in the fields of Life and Related Sciences, Physical and Related Sciences, and Social and Related Sciences, while the other three, Computer and Math Sciences, Engineering, and Non-S&E degrees, still have a wage penalty of more than 7%. Female workers from Computer and Math Sciences are shown to have 23% of wage penalty even when controlling for the basic task intensity.

As shown in the data description above, the difference in the distribution of related statuses

Table 2.4: Log wage penalties for working in an unrelated job across degree types

	Male		Female	
Bachelor's	-0.0786***	-0.0416***	-0.0841***	-0.0439***
	(0.0062)	(0.0063)	(0.0070)	(0.0098)
Master's	-0.0846***	-0.0555***	-0.0827***	-0.0578**
	(0.0112)	(0.0113)	(0.0134)	(0.0203)
Doctorate	-0.0708*	-0.0235	-0.1269**	-0.0888
	(0.0304)	(0.0303)	(0.0477)	(0.0693)
Professional	-0.3856***	-0.3350***	-0.4536***	-0.3292***
	(0.0308)	(0.0308)	(0.0518)	(0.0957)
Individual characteristics	Y	Y	Y	Y
Basic task intensity	N	Y	N	Y
N	64,782	64,782	37,332	37,332

Standard errors are in parentheses.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Individual characteristics are highest degree dummies, years of full-time and part-time experience, race dummies, disabled dummy, marital status dummy, training dummies, and most recent degree field dummies.

across the fields of study mostly stems from individuals whose highest degree is a Bachelor's degree. Based on these findings, I then do the same analysis by restricting the sample to those whose highest degree is a Bachelor's degree. Table 2.6 reports some differences from Table 2.5, but the results are similar overall. This similarity suggests that the results in Table 2.5 are driven by workers with a Bachelor's degree, a group which accounts for close to 60% of the main analysis sample.[8] In sum, the results across the major fields of study suggest that the importance of mismatch in major-specific knowledge to wage penalties can significantly depend on the field of study.

2.5 Conclusion

In college, students choose their major field of study, and the accumulated human capital may be somewhat specific to that field. Hence, some of this human capital that is accumulated in college may only be viable in jobs that are related to their major field of study. The literature shows that there exists a wage penalty for working in a job that is unrelated to one's major field of study. Understanding the source factors that generate the wage penalty is important for an efficient allocation of workers.

In this study, I examined the sources of wage penalty for working outside one's major field

[8]It might be interesting to see if wage penalties are similar across the fields of study by restricting the sample to those with a more advanced degree. However, since the fraction of people in an unrelated job to those who have an advanced degree is small, I cannot obtain precise estimates by fields of study based on my dataset.

Table 2.5: Log wage penalties of working in an unrelated job across degree fields

	Male		Female	
Computer and Math Sciences	-0.1495***	-0.1144***	-0.3062***	-0.2308***
	(0.0245)	(0.0244)	(0.0346)	(0.0342)
Life and Related Sciences	-0.0097	0.0425*	-0.0603*	0.0299
	(0.0203)	(0.0204)	(0.0261)	(0.0259)
Physical and Related Sciences	-0.0717**	-0.0134	-0.0964*	0.0047
	(0.0237)	(0.0237)	(0.0446)	(0.0441)
Social and Related Sciences	0.0128	0.0436**	-0.0401**	0.0018
	(0.0148)	(0.0149)	(0.0155)	(0.0155)
Engineering	-0.1819***	-0.1351***	-0.2215***	-0.1289**
	(0.0152)	(0.0153)	(0.0422)	(0.0417)
Non-S&E degrees	-0.1366***	-0.1066***	-0.1064***	-0.0664***
	(0.0068)	(0.0070)	(0.0071)	(0.0072)
Individual characteristics	Y	Y	Y	Y
Basic task intensity	N	Y	N	Y
N	64,782	64,782	37,332	37,332

Standard errors are in parentheses.

*$p < 0.05$,**$p < 0.01$, ***$p < 0.001$.

Individual characteristics are highest degree dummies, years of full-time and part-time experience, race dummies, disabled dummy, marital status dummy, training dummies, and most recent degree field dummies.

of study. I consider two channels: levels of basic skills required for work and a mismatch in major-specific knowledge. Basic skills are skills that are useful in any job, while major-specific knowledge is helpful only in related jobs. Unrelated jobs can be different from related jobs not only in the requirement of having major-specific knowledge but also in the required levels of basic skills. Indeed, distinguishing between basic skills and major-specific knowledge is important from the perspective of an efficient allocation of workers and career advice for students. Furthermore, this information could also impact upon the choices by a student on their college major.

There are previous papers that study the wage penalties associated with a mismatch by field and a mismatch by education qualification. The studies show that the wage penalty can be little in college jobs. Since non-college jobs will require lower levels of basic skills than college jobs, their results suggest the importance of levels of basic skills required in a job. However, using two-dimensional basic skills, my paper considers job heterogeneity within education qualification. Also, by directly controlling for the required levels of basic skills in one's job in my paper, the remaining wage penalty is more likely to represent the effect of a mismatch in major-specific knowledge, which cannot be avoided in any unrelated job. Moreover, my paper examines the heterogeneity of wage penalties across degree types and fields of study.

I combined two US datasets. Individual data come from the 1993 NSCG, and occupation

Table 2.6: Log wage penalties of working in an unrelated job across bachelor degree fields

	Male		Female	
Computer and Math Sciences	-0.1608***	-0.1165***	-0.3413***	-0.2514***
	(0.0282)	(0.0281)	(0.0376)	(0.0371)
Life and Related Sciences	0.0518*	0.1006***	-0.0640*	0.0222
	(0.0237)	(0.0236)	(0.0297)	(0.0293)
Physical and Related Sciences	-0.0649*	-0.0141	-0.0783	0.0224
	(0.0280)	(0.0279)	(0.0539)	(0.0530)
Social and Related Sciences	0.0273	0.0665***	0.0048	0.0529**
	(0.0191)	(0.0191)	(0.0192)	(0.0191)
Engineering	-0.1910***	-0.1504***	-0.2713***	-0.1511**
	(0.0174)	(0.0174)	(0.0491)	(0.0484)
Non-S&E degrees	-0.1369***	-0.0946***	-0.1085***	-0.0525***
	(0.0080)	(0.0081)	(0.0082)	(0.0083)
Individual characteristics	Y	Y	Y	Y
Basic task intensity	N	Y	N	Y
N	40,290	40,290	23,176	23,176

Standard errors are in parentheses.

$*p < 0.05$,$**p < 0.01$, $***p < 0.001$.

Individual characteristics are highest degree dummies, years of full-time and part-time experience, race dummies, disabled dummy, marital status dummy, training dummies, and most recent degree field dummies.

data come from the O*NET. The NSCG dataset includes a worker's self-assessment measure of relatedness, and I use the O*NET to construct measures of basic task intensity, which represent levels of basic skills required in a job. I showed that unrelated jobs tend to require lower levels of basic skills than related jobs, but I also found that unrelated college jobs and closely related college jobs that workers tend to hold are similar in relation to the requirement of basic skills.

I ran two types of log wage regressions, separately by gender. After controlling for the individual characteristics, average wage loss for working outside a major field study is 9% for both males and females. When the basic task intensity is additionally controlled for, the wage loss decreased by around 45%, from 9% to 5%, for both males and females. These results suggest that both major-specific knowledge and basic skills can play a role in wage penalties when working outside the major field of study.

Across degree types, workers with an advanced or specialized degree tend to show large wage penalties even after controlling for the basic task intensity. There are also wide variances across the fields of study. Major-specific knowledge seems to be very important to Computer and Math Sciences and Engineering. The heterogeneity across degree types and fields of study suggests that career advice for students should vary based on their degree type and field of study. According to my results, workers with a Bachelor's degree in Life and Related Sciences, Physical and Related

Sciences, or Social and Related Sciences will not suffer from wage loss stemming from a mismatch in their major-specific knowledge. This also suggests that they have not acquired major-specific knowledge much.

My current analysis does not take into account the endogeneity of occupation choice due to the data limitation and cannot exclude the possibility that workers in unrelated jobs are negatively selected. If workers in unrelated jobs have a low level of major-specific knowledge, a part of the wage differences between those in closely related and unrelated jobs will stem from the negative selection of workers. It is also possible that some of the estimated heterogeneity across degree types and fields of study reflect differences in the degree of the selection of workers by degree type and field of stuudy. Endogeneity of occupation choice should be addressed in future work.

In this paper, I showed that observed wage differences between related and unrelated jobs do not directly reflect underutilized major-specific knowledge and that the wage difference partially stems from the fact that unrelated jobs, on average, require lower levels of basic skills in work than related jobs. Dividing major specificity as it appears in the literature into basic skills and major-specific knowledge is a new point of view, and I believe that it will be useful in analyzing topics related to college major specificity, such as college major choice and occupation choice.

2.6 Bibliography

Abel, J. R., and R. Deitz, 2015, Agglomeration and Job Matching Among College Graduates: Regional Science and Urban Economics, **51**, 14–24.

Arcidiacono, P., 2004, Ability Sorting and the Returns to College Major: Journal of Econometrics, **121**, 343–375.

Kinsler, J., and R. Pavan, 2015, The Specificity of General Human Capital: Evidence from College Major Choice: Journal of Labor Economics, **33**, 933–972.

Lemieux, T., 2014, Occupations, Fields of Study and Returns to Education: Canadian Journal of Economics, **47**, 1047–1077.

Leuven, E., and H. Oosterbeek, 2011, Overeducation and Mismatch in the Labor Market, *in* Handbook of the Economics of Education: Elsevier B.V., **4**, 1–56.

Montt, G., 2017, Field-of-Study Mismatch and Overqualification: Labour Market Correlates and Their Wage Penalty: IZA Journal of Labor Economics, **6**.

Nordin, M., I. Persson, and D. O. Rooth, 2010, Education-Occupation Mismatch: Is There an Income Penalty?: Economics of Education Review, **29**, 1047–1059.

Robst, J., 2007a, Education and Job Match: The Relatedness of College Major and Work: Economics of Education Review, **26**, 397–407.

———, 2007b, Education, College Major, and Job Match: Gender Differences in Reasons for Mismatch: Education Economics, **15**, 159–175.

———, 2008, Overeducation and College Major : Expanding the Definition of Mismatch Between Schooling and Jobs: Manchester School, **76**, 349–368.

Sellami, S., D. Verhaest, W. Nonneman, and W. Van Trier, 2017, The Impact of Educational Mismatches on Wages: The Influence of Measurement Error and Unobserved Heterogeneity: B.E. Journal of Economic Analysis and Policy, **17**.

Wiswall, M., and B. Zafar, 2015, Determinants of College Major Choice: Identification Using an Information Experiment: Review of Economic Studies, **82**, 791–824.

Chapter 3

Heterogeneous Skill Growth Across College Majors

3.1 Introduction

Wage inequality across college majors has recently been attracting attention. Carnevale et al. (2015) report that the average difference in the lifetime earnings in the US between the highest-paying major, engineering, and the lowest-paying major, education, is \$3.4 million while the average difference between college and high school graduates is \$1 million. Inequality research has mostly focused on disparities between high school and college graduates, but given this earnings gap across majors, college majors are an equally important determinant of future career prospects.

Since college majors are not randomly chosen by students, sorting may contribute to these large income differences among majors. Students who choose a "high-paying" major might earn a lot even if they chose a "low-paying" major.[1] Many previous papers examine whether wages are different across college majors even when controlling for sorting or self-selection into major (see, e.g., Arcidiacono (2004); Hamermesh and Donald (2008); Arcidiacono et al. (2016b), and Kirkeboen et al. (2016)). Their consensus is that, even when controlling for selection or sorting, wages are different across college majors and that math-intensive majors, especially Science, Technology, Engineering, and Mathematics (STEM) majors, show high returns.

In this paper, I examine where wage differentials across majors come from using a multi-dimensional skill framework. I estimate how much college graduates increase their general type of skills and their major-specific type of skills during college depending on their major. Studying

[1] For example, the average SAT scores vary by college major. According to a report in 2015 by College Board (2015), the average scores of individuals who intended to major in engineering are 526 (Critical Reading), 575 (Mathematics), and 509 (Writing), while those of individuals who intended to major in education are 481 (Critical Reading), 480 (Mathematics), and 473 (Writing).

skill growth in these multiple dimensions will be helpful to understand differences across college majors more deeply.

There is a big argument about which majors public money should be spend on. According to the literature, STEM majors dominate the other majors in terms of wages. However, this does not necessarily mean that STEM majors dominate the other majors in every type of skills and that the government should focus only on STEM majors. For example, STEM majors might mostly increase highly-rewarded STEM major-specific skills, while humanities majors might mostly increase less-rewarded more general type of skills. In this case, these two majors provide workers with different types of skills, and thus, expanding only STEM majors may cause problems in the economy in future, because we might face a shortage of workers with high general type of skills. This kind of role of college major cannot be examined by focusing only on wage differentials or by using a single-dimensional skill framework.

Furthermore, major-specific type of skills imply a risk of choosing the relevant major. Suppose that major-specific type of skills are the main contributor to the high wages among STEM majors. In this case, their wages can become much lower than the current wages if the economy changes and most STEM occupations change or disappear. Deming and Noray (2018) argue that the technological change over the last decade made obsolete the skills possessed by applied science majors in an old cohort and caused a declining wage premium for them with age. In addition, some students might see this risk of choosing STEM majors and might avoid majoring in STEM. Kinsler and Pavan (2015) argue that skill specificity can play an important role in students' major choice. Despite the fact that STEM majors tend to earn high wages, many students avoid majoring in STEM. The governments in many countries are trying to create policies to increase the number of students in STEM majors. A multi-dimensional skills framework will help to further our understanding of students' college major choice better.

Wage returns are closely related to skill growth. If every major increases a common single type of skill and there is no heterogeneity in the wage returns to the skill across occupations, then wage differences will directly reflect skill growth differences across major. A one-dimensional skill framework may be appropriate in studying differences between high school and college graduates because a crucial difference between these two groups is years of education. Since college graduates spent more time on studying, they will possess a larger amount of skill, which will contribute to wage differential between college and high school graduates.

However, a one-dimensional skill framework may not be enough to analyze differences across college majors. Skill accumulation processes could be heterogeneous across college majors in both quantity and type. College majors are different in how demanding they are in terms of course burden and the number of credits required to graduate. For example, in the 1993 Baccalaureate and Beyond Longitudinal Study (B&B:93), the average number of total credits undergraduate students earn is 150 in engineering majors, 133 in business majors, and 131 in humanities majors. Hence, in some majors, students might increase their skills more than in other majors, in a general sense. Furthermore, courses college students take vary significantly by their college major. For example, engineering majors take more science courses than humanities courses, while the reverse occurs for education majors. These course differences will result in students accumulating different types and amounts of skills. A multi-dimensional framework capturing this skill growth heterogeneity in both quantity and type will provide us a richer framework in analyzing differences across majors.

In my model, each college major represents a different skill production function; skills students start with will evolve differently depending on their major. In order to capture skill growth differences in both quantity and type, I assume that each college major increases two types of skills: a general cognitive skill and a major-specific skill. General cognitive skill can increase in all college majors, but the amount of growth can vary by major. General cognitive skill growth captures the similarity of the skills produced in different college majors. In addition to general cognitive skill, each college major can increase a major-specific skill.[2] Major-specific skill growth captures the uniqueness of the skills accumulated in the major. I allow for individual heterogeneity in the growth of both types of skills. The growth in general cognitive skill can be affected by the pre-college level of general cognitive skill.

If there are test scores of each type of skill at both pre- and post-college periods, skill growth by major could be easily measured. However, although students take cognitive tests before entering college, such as high school graduation exams or college entrance exams, that are observed in various data sets, people usually do not take such tests after college. My data include only pre-college cognitive test scores, and the approach of comparing test scores at different periods, common in the literature on elementary schooling, cannot be applied in this case. Hence, I use other sources of information.

For general cognitive skill, I exploit post-education occupation choice, which reflects post-education skill level. In exploiting occupation choice, I do not use occupation as it is. High school graduates and college graduates, of course, tend to take different jobs. Furthermore, as Ransom (2014) and Altonji et al. (2012b) show, college majors and post-college occupations are closely related. If occupation or industry categories are used to control for jobs, the categories can be very coarse. Hence, I take a task-based approach and characterize each occupation by a low-dimensional task portfolio. Following Acemoglu and Autor (2011), I define task as a unit of work activity that produces output while skill is defined as a worker's endowment of capabilities for performing various tasks. Individuals possess skills and apply the skills in tasks to produce outputs. The task-based approach enables me to relate occupations to each other.

One type of the task intensity in this paper, which represent how intensely the relevant skills are

[2]As described later, I aggregate majors into three types, hence, there are three types of major-specific skills in total.

required in work, is general cognitive task intensity. General cognitive task intensity corresponds to general cognitive skill. Given pre-college general cognitive skill level, individuals who have increased their general cognitive skill more will take a job that has a higher level of general cognitive task intensity. Since most respondents in my dataset, including those who did not go to college, took the cognitive tests, the skill growth in certain majors can be identified by comparing the task intensity choice by individuals from the majors with that by high school graduates conditional on pre-college general cognitive skill levels. Furthermore, I can show that the general cognitive skill which is measured by the cognitive test scores corresponds to the general cognitive task intensity by using high school graduates, who are assumed to enter the labour market with the skill level measured by the pre-college cognitive test scores.

For major-specific skills, although I use college Grade Point Average (GPA) as measures of post-college major-specific skills, I do not have measures of pre-college major-specific skills.[3] Hence, I make some assumptions to specify pre-college major-specific skills. I consider that my cognitive test scores and college GPA will not perfectly measure skills. Therefore, I use a dynamic factor model to deal with the measurement errors.

I combine two US datasets: the National Longitudinal Survey of Youth 1997 (NLSY97) for individual level data and the Occupational Information Network (O*NET) for occupation data. The NLSY97 provides data on Armed Service Vocational Aptitude Battery (ASVAB) test scores, college majors, college GPA, wages, and occupations. Most respondents, including both those who eventually went to college and those who did not, took the ASVAB test while in junior high or high school. These test scores are assumed to reflect pre-college general cognitive skill levels. I aggregate college majors into three types of majors: STEM, Business & Economics, and Humanities & Social Sciences majors. The O*NET contains a number of standardized measures describing the day-to-day aspects of the job and qualifications and interests of the typical workers in occupations. I use the O*NET to construct task portfolios.

The results show that all college majors significantly increase a student's general cognitive skill and that there are also large differences in the growth across college majors. This is reflected in the wage premium for college graduates and the large wage differences across majors. For those having a population average level of pre-college skill, the skill growth in STEM majors is 16 points higher than that in Humanities & Social Sciences majors using a log wage point metric. Although the skill growth varies somewhat with pre-college general cognitive skill levels, STEM majors still have a larger production of the skill and associated wage returns in the labour market. Wage returns to a major-specific skill are allowed for depending on whether the occupations are related to the corresponding major. The estimates suggest that only Business & Economics and STEM majors have a positive wage benefit from major-specific skill growth. The wage effect is

[3] GPA is allowed to be affected by general cognitive skill as well.

the largest among STEM majors, but the effect is still only about one quarter of that from general cognitive skill growth. Overall, growth in general cognitive skill is most important during college and also plays an important role in wage differentials across majors.

Although there is a growing number of papers estimating wage returns to college majors, relatively few papers in the previous literature examine skill growth across majors. Lemieux (2014) is the closest to my study. He considers three channels to increase wages: general skill growth, occupation, and match between college major and occupation. The major-occupation match matters because of major-specific skill growth. He uses Canadian datasets and decomposes wage gap between college workers and high school workers by college major into the three channels. However, there is no sorting into majors or into occupation, and there is no individual heterogeneity in skill growth. My model allows for both sorting and individual heterogeneity. In addition, my task-based approach avoids the relatively coarse occupation classification used in Lemieux (2014), in which people classified in the same occupation category might be doing very different work.

The model developed by Kinsler and Pavan (2015) is similar to my model in that they assume students bring low-dimensional skills into college and the skills change depending on their college major. Unlike my model, their model assumes two types of general cognitive skills, verbal and math, and there is no major-specific skill. In their model, different wage returns to the general cognitive skills represent major specificity. Since their main interest is whether wage returns vary by job, they do not estimate skill changes. Furthermore, since they classify jobs based on only one dimension, related or not related, for each college major, and since related and unrelated jobs will be different depending on major, they cannot compare wage differences across majors controlling for occupations. Arcidiacono et al. (2017) use data on subjective expectations that were collected at Duke University and argue that there exist sizable complementarities between some college majors and occupations.

My paper is also related to studies showing positive effects of schooling on cognitive test scores, such as Gormley and Gayer (2005); Aaronson and Mazumder (2011), and Fitzpatrick et al. (2011). Hansen et al. (2004) and Cascio and Lewis (2006) analyze the effects on ASVAB scores. Although their focus is on an earlier level of schooling than post-secondary education, they show that schooling has positive impacts on cognitive test scores. Aucejo and James (2016) examine math and verbal skills changes during primary and secondary education using UK data.

The rest of this paper is organized as follows: section 2 describes my data. My model is explained in section 3. Section 4 explains the identification of skill growth parameters. Section 5 discusses the estimation. Results are reported in section 6. Section 7 concludes.

3.2 Data

I use two US datasets, the NLSY97 and the O*NET. The NLSY97 provides individual level data on test scores, college majors, college GPA, wages, and occupations. The O*NET is used to characterize occupations by task portfolios. I combine individual data from the NLSY97 with occupational data from the O*NET.

3.2.1 NLSY97

The NLSY97 is a panel survey conducted by the US Bureau of Labor Statistics. It started in 1997, was conducted annually up to 2011, and has been conducted biennially since then. In the first round, 8,984 males and females, who were between 12 and 17 years old at that time, were interviewed.

Test scores

Most of the respondents of the NLSY97 took an ASVAB test during the first round, between summer 1997 and spring 1998, when they were between 12 and 18 years old. I use ASVAB scores as noisy measurements of pre-college general cognitive skill instead of perfectly accurate measurements. Since there may be an age effect, I adjust the test scores for age using the method of Altonji et al. (2012).[4]

The ASVAB test is composed of many sections, eight of which are used in this paper: Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, Mathematics Knowledge, Numerical Operation, Mechanical Comprehension, Auto & Shop Information, and Electronics Information. Word Knowledge and Paragraph Comprehension are considered as verbal tests and Arithmetic Reasoning, Mathematics Knowledge, and Numerical Operation are considered as math tests. As Armed Forces Qualification Test (AFQT) scores are calculated based on these five test scores, they are often used to construct cognitive skill. Mechanical Comprehension, Auto & Shop Information, and Electronics Information are used to construct mechanical skill (see Prada and Urzúa (2017) and Speer (2017a)).

Education

Education levels are categorized into high school, some college, and college. High school includes GED. Some college includes those with an Associate degree and those who went to college for two

[4] In this method, ASVAB score at the qth percentile in the distribution of the taker's age is assigned to the ASVAB score at the qth percentile in the age 16 distribution. This method implicitly assumes that age effects do not change the rank in ASVAB scores although age effects can be nonconstant.

years or more but who did not earn Bachelor's degree. College is further categorized into three college majors. The definition of college majors in this paper is explained below.

Education group is defined by the individual's highest education with some exceptions. First, college graduates are defined as those who obtained a Bachelor's degree before 2010. This is because of a large change in the college major classification recorded in the NLSY97. Second, since my focus is on undergraduates, I drop those with a more advanced degree, although I keep those whose highest degree is a Master's if they have full-time work experience after their Bachelor's degree but before their Master's degree.[5] I restrict my sample to those who have at least a high school diploma.

College majors

In the NLSY97, students are asked to report their college major each term. Among the reported majors, I take the final one as their college major. As mentioned above, the classification of college majors changed substantially in 2010. Hence, I use Bachelor's degree major earned before 2010. I aggregate college majors into three majors based on the extent to which they are math intensive: Humanities & Social Sciences, Business & Economics, and STEM majors.[6] A similar classification is used in Kinsler and Pavan (2015).

College GPA

College GPA is reported each term, and I calculate annual GPA as the average GPA across all semesters in the year. I use the last two years of reported GPA as noisy measures of post-college major-specific skill.[7] I choose the last two years because major-specific skill will be acquired mainly in these years. American college students typically take general academic subjects in their first two years and take specialized subjects after that.

Occupation and wages

I consider only full-time jobs, which are defined as equal to or more than 35 hours worked per week. Part-time jobs are not used because the wage structure may be different from that of full-time jobs. I use the job information of the first year. I construct occupation variables on an annual

[5]Those who are dropped account for only around 2% of people whose education is classified as high school or higher.

[6]STEM majors include Agriculture & natural resource sciences, Biological sciences, Architecture/environmental design, Computer/information science, Engineering, Mathematics, Physical sciences, and Nutrition/Dietics/Food science. Health is dropped. Business & Economics majors include Business management, Economics, and Hotel/Hospitality management. Humanities & Social Sciences majors include all the other majors.

[7]I assume GPA is affected by evolved general cognitive skill as well. However, as seen in the estimation section later, GPA is not used to identify the distribution of pre-college general cognitive skill.

basis. If a worker had several jobs within a year, the one with the most weeks worked is taken as the occupation in that year. Using the occupation codes, the NLSY97 and the O*NET are connected.[8] In the following analysis, I use hourly compensation rates as wages. The wages are adjusted to dollars in year 2000. I restrict to the rates between $1 and $100 by assuming the others are misreported wages.

3.2.2 O*NET

I use the O*NET to construct task portfolios, which represent how intensely or importantly each type of skill is used in work. The O*NET is sponsored by the US Department of Labor/Employment and Training Administration and started as a successor to the Dictionary of Occupational Titles (DOT). Both the O*NET and the DOT characterize occupations by standardized measures and have been used in many papers taking a task-based approach (see, e.g., Poletaev and Robinson (2008) and Guvenen et al. (2016)). The O*NET contains a number of standardized measures describing the day-to-day aspects of the job and qualifications and interests of the typical workers in the occupations.[9]

The measures chosen from the O*NET are reported in Table 4.1. The selection of general cognitive measures is mainly based on a technical report by the ASVAB Career Exploration Program (ASVAB Career Exploration Program (2011)) and that of mechanical measures is based on Speer (2017a). For each type of element, I employ a Principal Component Analysis (PCA) and take the first component as task intensity of that type. I assume that the constructed task intensity is an accurate measurement.[10] Each type of task intensity is standardized to have a mean of 0 and a standard deviation of 1 over all full-time job observations in the NLSY97.

Although many papers on the wage penalty of working in unrelated jobs use a worker's self-assessed relatedness measure, not many datasets include one and my dataset does not have one. Hence, I define relatedness using O*NET measures. Using measures from the O*NET, I construct major-specific task importance, which represents how important major-specific skill is to perform

[8]The occupation codes in the NLSY97 are based on Census 2002 occupation codes, while those in the O*NET I use are based on Standard Occupation Classification (SOC) 2010. I use crosswalks to connect these two types codes. The crosswalks between SOC 2010 and 2009, between SOC 2009 and 2006, and between SOC 2006 and 2000 are provided by O*NET resource center. The crosswalk between SOC 2000 and Census 2002 occupation codes is distributed by the National Crosswalk Service Center.

[9]For each element, Importance and Level are recorded. I use Level for general cognitive task and mechanical task, which is recorded with a range of 0-7 based on ratings by analysts or job incumbents. I use Importance, which is recorded with a range of 1 (Not important) to 5 (Extremely important), for major-specific task. I use Importance for major-specific task because I want to capture the type of jobs, but, since Importance and Level are strongly correlated with each other, using Importance instead of Level will not change my results much.

[10]If the constructed task intensity is a noisy measure of the "true" task intensity and if the measurement errors are classical, my estimates of skill growth parameters are still consistent, because the effects of general cognitive skill on occupation choice are consistently estimated.

Table 3.1: O*NET elements used to construct task intensity/importance

General cognitive	
Oral comprehension	Oral expression
Written comprehension	Written expression
English language	Reading comprehension
Speaking	Writing
Mathematical reasoning	Number facility
Mathematics	Mathematics skill
Deductive reasoning	Inductive reasoning
Analyzing data or information	
Mechanical	
Handling and moving objects	
Controlling machines and processes	
Repairing and maintaining mechanical equipment	
Repairing and maintaining electrical equipment	
Inspecting equipment, structures, or material	
Operating vehicles, mechanized devices, or equipment	
Equipment maintenance skill	
Mechanical knowledge	
Humanities & Social Sciences	
Communications and media	English language
Sociology and anthropology	Geography
Therapy and counseling	Foreign language
Public safety and security	Fine arts
History and archeology	Psychology
Philosophy and theology	Education and training
Business & Economics	
Administration and management	Sales and marketing
Economics and accounting	Customer and personal service
Personnel and human resources	
STEM	
Computers and electronics	Design
Engineering and technology	Mathematics
Physics	Chemistry
Biology	

the job. Then, I categorize jobs into related jobs and unrelated jobs based on the constructed task importance. I define jobs as related or unrelated instead of using the constructed continuous measures of major-specific task importance for two reasons. One is comparability with the previous papers. Most of the previous papers use two or three relatedness categories. Another reason is a comparison of major-specific skill growth between majors. Since major-specific task importance is measured differently by major, the measures cannot really be compared between majors directly and so the results are difficult to interpret.[11]

The selection of measures for major-specific tasks follows Freeman and Hirsch (2008). In Freeman and Hirsch (2008), each college major is connected with one measure of knowledge in the O*NET. Since college majors are aggregated into three types in my study, I construct major-specific task importance as follows. I select O*NET knowledge measures related to detailed majors contained in each aggregated college major category. Then, I employ a PCA to the selected measures and take the first component as the major-specific task importance.

The histograms of the constructed task importance are shown in Figure 3.1. Each major-specific task importance is standardized to have a mean of 0 and a standard deviation of 1 over all full-time job observations in the NLSY97. They show that college graduates tend to take higher task importance corresponding to their college major. Since college graduates are assumed not to increase major-specific skill other than that of their own major, I expect that most of them take jobs unrelated to the other majors. Based on Figure 3.1, I consider jobs as related jobs if the corresponding task importance is equal to or higher than 1, while I consider jobs as unrelated jobs if the corresponding task importance is less than 1. For example, the major specific task importance of the job of Accountants is -0.65 (Humanities & Social Sciences), 1.46 (Business & Economics), and -0.74 (STEM). Hence, the job of Accountants is related to Business & Economics majors, but unrelated to Humanities & Social Sciences majors and to STEM majors.[12]

I assume that high school and some college graduates take a job unrelated to any major. In addition, I assume that college graduates can choose job relatedness only regarding their own major and have a job unrelated to majors other than that. The first assumption is made because high school and some college graduates do not have any "major". The second assumption is made because papers on the match quality between college majors and work only examine a student's major. They do not consider whether the work is related to outside of the student's major. I could potentially construct a model, in which individuals can choose job relatedness regarding each major, but I do not do that for simplicity. Although there are some individuals who have jobs that are defined as "related" in the data, I assume that their employers do not care their major-specific

[11] The O*NET measures and a worker's self-assessed relatedness are positively correlated in the 1993 National Survey of College Graduates.

[12] Jobs can be related to more than one major in the data. This might explain why some people are in jobs that are related to another major as observed in Figure 3.1.

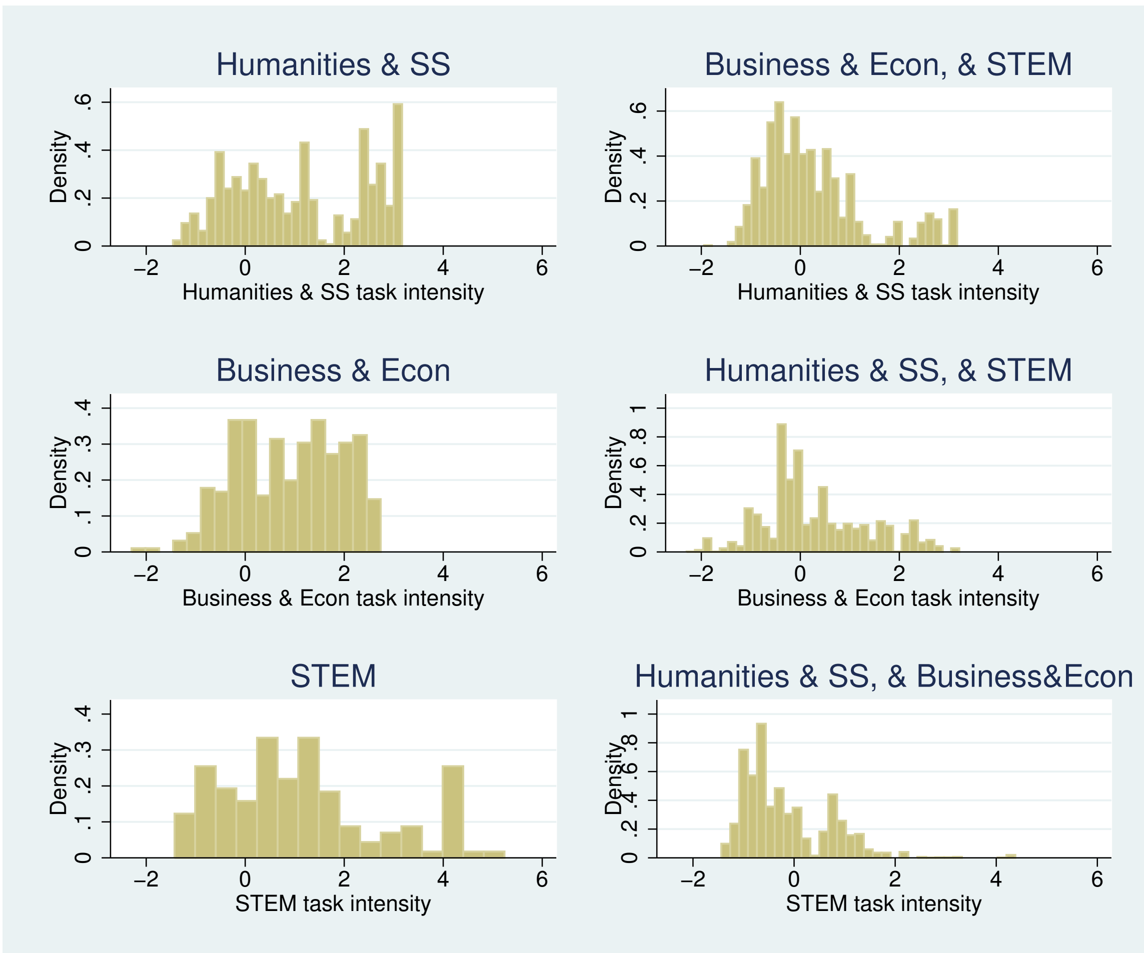

Figure 3.1: Histograms of major-specific task importance by college majors

Note: Each type of task importance is standardized to have a mean of 0 and a standard deviation of 1 over all full-time job observations in the NLSY97.

skill levels because the skills are too low.

3.2.2.1 Mechanical skill and task

Mechanical skill is introduced to allow high school graduates to choose an occupation based not only on cognitive skill. A recent paper by Prada and Urzúa (2017) shows that a higher level of mechanical skill reduces the probability of attending college given cognitive skill and also shows that wage returns to mechanical skill are large for high school graduates.[13] Figure 3.2 shows histograms of mechanical task intensity by education level. Remember that task intensity is standardized to have a mean of 0 and a standard deviation of 1 over all full-time job observations in the NLSY97. High school shows two humps. One is between -1 and 0, and another is between 1 and 2. In the case of some college, there is a second hump between 1 and 2, but it is small and not as obvious as high school. College shows only one hump between -1.5 and 0. This implies that the mechanical dimension is not important to college graduates. Given these observations, I define jobs as mechanical if the mechanical task intensity is higher than 0.5 for high school and some college graduates. All other jobs are defined as cognitive type jobs. Mechanical skill and task intensity do not matter in cognitive type jobs.[14]

An advantage of dividing jobs into mechanical and cognitive types is to make the model easy to interpret since mechanical skill does not matter anymore given cognitive type jobs. This reduces the computation burden as well. Furthermore, it makes explicit that the mechanical dimension is not relevant for college graduates.

3.2.3 Summary statistics

Table 3.2 shows summary statistics. Both cognitive and mechanical test scores are standardized to have a mean of 0 and a standard deviation of 1 over the population. High school shows the lowest and STEM majors show the highest average score, and people with a higher level of education tend to have higher test scores. Similarly, general cognitive task intensity is higher for a higher level of education, with high school the lowest and STEM majors the highest. Interestingly, this is not the case for mechanical task. In mechanical task intensity, high school shows the highest, while Humanities & Social Sciences and Business & Economics majors show the lowest. Even STEM

[13] In earlier related research, Yamaguchi (2012) documents that less educated people tend to take a job involving intense physical tasks. Physical task measures and mechanical task measures in the O*NET are highly correlated. Since my dataset does not have a good measure of physical skill, I introduce mechanical skill instead of physical skill. Given the close relatedness between physical and mechanical task measures, I expect that a selection between physical and mechanical skills will not matter much.

[14] If the log wage is regressed on cognitive and mechanical test scores and task intensity measures, and other controls using high school graduates in cognitive type jobs, the coefficient of mechanical test scores is negative and that of mechanical task intensity is insignificant.

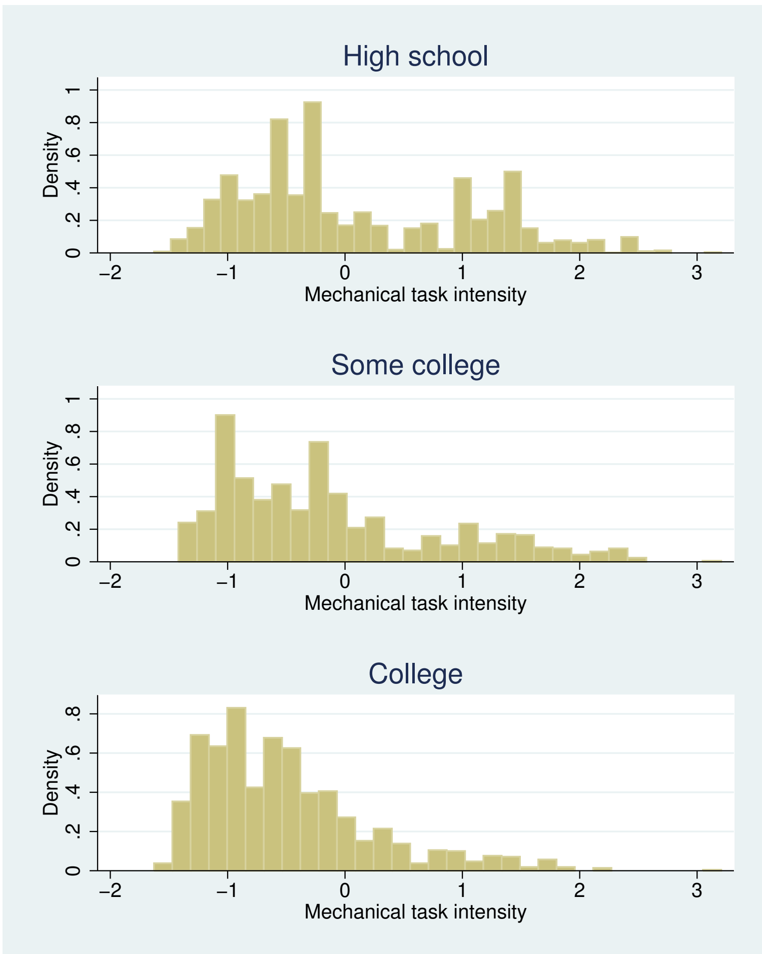

Figure 3.2: Histograms of mechanical task intensity in first year by education level

Note: Each type of task intensity is standardized to have a mean of 0 and a standard deviation of 1 over all full-time job observations in the NLSY97.

Table 3.2: Summary statistics

	High school	Some college	Humanities&SS	Business&Econ	STEM
ASVAB test scores					
Cognitive	-0.3255	0.0524	0.5306	0.5794	0.7804
	(0.8082)	(0.7093)	(0.6523)	(0.5918)	(0.5851)
Mechanical	-0.1852	0.0437	0.2157	0.3253	0.6869
	(0.8578)	(0.8246)	(0.6755)	(0.6833)	(0.7929)
Task intensity					
Cognitive	-0.4085	0.1906	0.6700	0.9107	1.1192
	(0.7933)	(0.8832)	(0.8097)	(0.8696)	(1.0027)
Mechanical	0.0869	-0.1439	-0.6213	-0.6419	-0.0369
	(0.9723)	(0.9503)	(0.5903)	(0.6881)	(0.8627)
Related job			0.5103	0.4809	0.4652
			(0.5002)	(0.5004)	(0.4997)
Wages	8.8881	11.9696	12.4170	14.3962	15.4904
	(4.8242)	(7.4956)	(6.8789)	(7.0064)	(7.9891)
N	3155	988	729	341	273

Notes: ASVAB test score is the average test score: Word Knowledge and Paragraph Comprehension, Arithmetic Reasoning, Mathematics Knowledge, and Numerical Operation are for general cognitive, and Mechanical Comprehension, Auto & Shop Information, and Electronics Information are for mechanical.

majors, who show much higher mechanical test scores, show lower mechanical task intensity than high school graduates. This suggests that mechanical skill is not important to them after college. About 50% of people took an occupation related to their major. Wages are higher for higher levels of education, and as expected, STEM majors show higher wages than the other majors.

3.3 Model

I first present an outline of the model. Then, I explain details of the empirical implementation.

3.3.1 Two-period model outline

It is well known that cognitive skills matter to whether people go to college or not. Furthermore, students are significantly sorted into college major based on their cognitive skill levels. I assume that students accumulate cognitive skills in college. College students take courses, most of which require cognitive skills. Course taking varies significantly by college major, and, thus, majors may differ in the accumulation of cognitive skills.

I divide cognitive skills into general cognitive and major-specific skills. General cognitive skill can increase in any major, and the amount can vary with major. Conversely, major-specific skill can increase only in its relevant major. Hence, general cognitive skill captures the similarity of the skills accumulated in different majors, and major-specific skill reflects the uniqueness of the skill accumulated in each major. For example, business majors will study business cases to learn business models of companies. They will increase their general cognitive skill through understanding and interpreting the cases. However, knowledge on the business cases per se, such as when and what company introduced the business model or the business history, is business major specific. In my model, further education represents a set of skill production functions, each taking pre-college general cognitive and major-specific skills as inputs and evolved general cognitive and major-specific skills as outputs.

The model has two periods. It starts at high school graduation and there are multiple decision stages. The time frame is as follows:

1. Figure 3.3 shows the decision flow in period 1. In the beginning of period 1, high school graduates are endowed with five-dimensional skills: general cognitive, mechanical, Humanities & Social Sciences major specific, Business & Economics major specific, and STEM major specific skills. Given these skills, they choose education level: work, some college, or college. If they choose work, they choose a job type, either cognitive or mechanical. In cognitive type jobs, mechanical skill does not matter.[15] After choosing a type, they choose task intensity corresponding to their chosen type of jobs; if they choose a cognitive type, they choose general cognitive task intensity.[16] High school graduates are assumed to have a job unrelated to any major. Their wages depend on their chosen job type, task intensity, and their skill levels. If they choose some college, they do not have a choice anymore in this period. If students decide to go to college, they learn how much major-specific skills they will have at college graduation for each major. I call them potential post-college major-specific skills. Potential post-college major-specific skill represents how good a students is at the relevant type of major. Based on their skills and potential post-college major-specific skills, they choose a college major from Humanities & Social Sciences, Business & Economics, or STEM majors. Skills of students choosing post-secondary education will evolve depending on their choice of education level and major.

2. Figure 3.4 presents the decision flow in period 2. In period 2, high school graduates continue working. Those who took post-secondary education finish their education and enter the labour market. Their skill levels have changed from pre-college skill levels depending on their education choice in the previous period. Some college could increase general cogni-

[15]Since I do not estimate the model part regarding mechanical type jobs, I do not assume which type of skill matters in mechanical type jobs.

[16]Workers with higher general cognitive skill will choose higher general cognitive task intensity because the wage return may be higher or they may be able to do the task more easily than workers with lower skill.

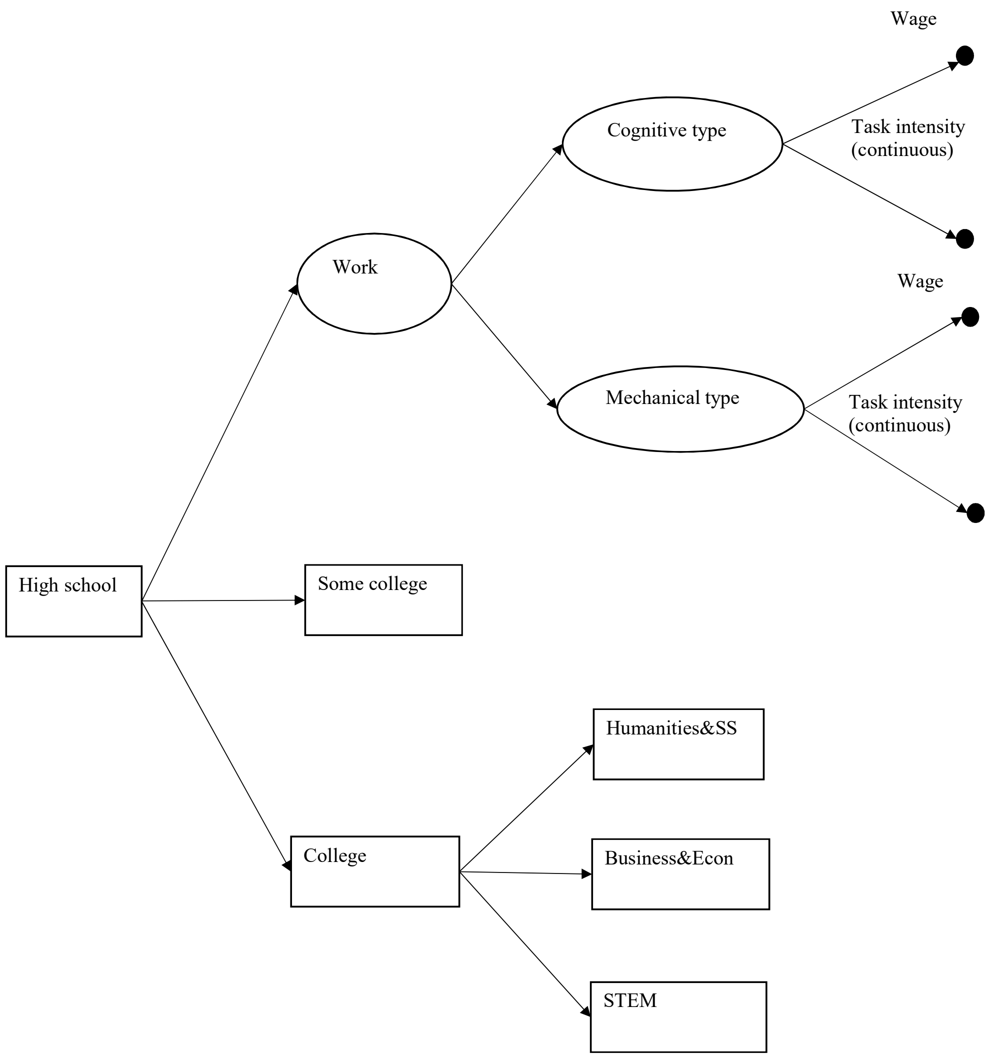

Figure 3.3: Decision flow; period 1

tive and mechanical skills. As with high school graduates, those who went to some college first choose either a cognitive type or mechanical type job and then choose task intensity. They are assumed to have a job unrelated to any major. Their wages depend on their chosen job type, task intensity, and their post-education skill levels. On the other hand, each college major increases general cognitive skill and its relevant major-specific skill. All college graduates are assumed to choose a cognitive type job. Hence, they choose general cognitive task intensity. Furthermore, they now have an option to choose job relatedness to their own major, either related or unrelated. Returns to major-specific skill can depend on the relatedness. Their wages depend on their chosen general cognitive task intensity, job relatedness to their major, and post-college skill levels.

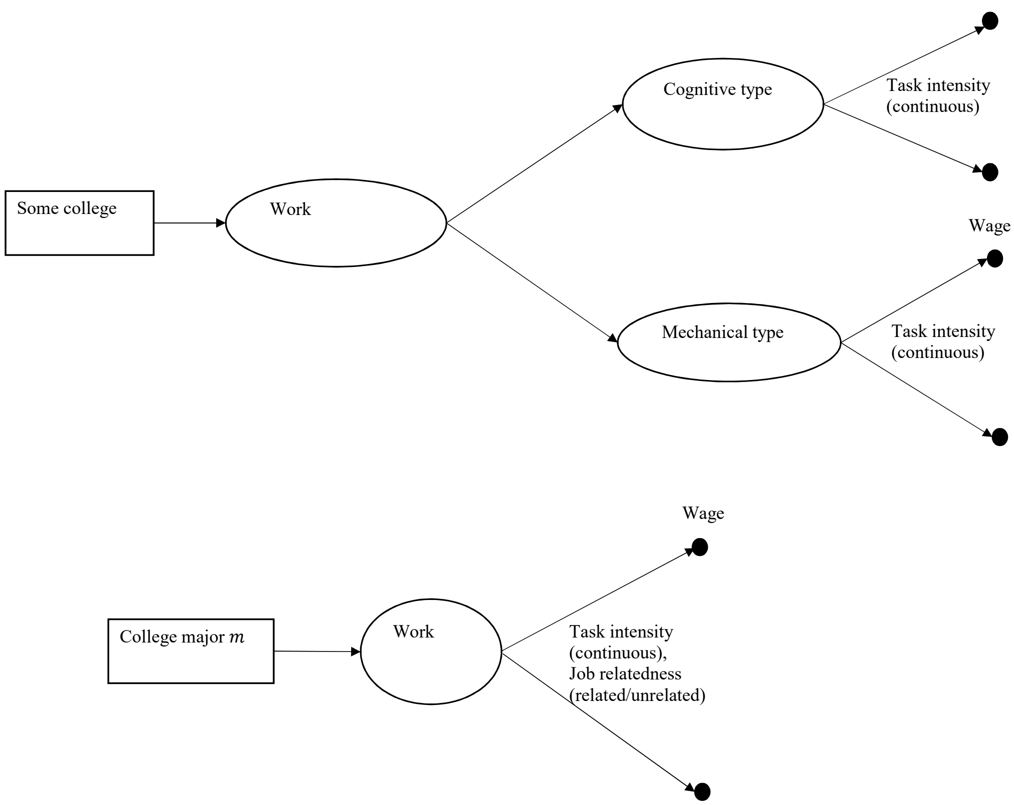

Figure 3.4: Decision flow; period 2

3.3.2 Empirical model

This subsection explains the details of my empirical model. In several decision stages, I assume a linear latent utility form instead of fully specifying a dynamic discrete choice model. A linear latent utility form is used in papers using a factor model in a dynamic treatment effect model (see, e.g., Heckman et al. (2016a,b), and Fruehwirth et al. (2016)). This specification is used mainly for simplicity, and one cannot say anything about the full underlying decision model. However, it has other advantages over fully specified dynamic discrete choice models. Fully specified dynamic discrete choice models require strong assumptions on agent preferences, constraints, and information sets. On the other hand, this simplified specification captures some essential features of dynamic discrete choice models without imposing the strong assumptions.

In the following, I first explain skills and then describe each stage in my model.

3.3.2.1 General cognitive and mechanical skills

High school graduates are assumed to have five types of skills, general cognitive, mechanical, and three types of major-specific skills. I here describe general cognitive and mechanical skills. Let $s_{1i} = (s^c_{1i}, s^{mech}_{1i})$ denote individual i's pre-college general cognitive and mechanical skill levels. The skills are modeled as follows:

$$s^c_{1i} = x'_{si}\alpha^c + \theta^c_i \tag{3.1}$$
$$s^{mech}_{1i} = x'_{si}\alpha^{mech} + \theta^{mech}_i. \tag{3.2}$$

This implies that the skills are the result of characteristics observed by the econometrician, x_s, and unobserved components, $\theta = (\theta^c, \theta^{mech})$. Unobserved components θ are orthogonal with observed components x_s. Vector x_s contains a constant, a female dummy, race dummies, father's education, mother's education, dummies for regions of residence in 1997, a dummy for living in an urban area in 1997, a dummy for broken home in 1997, household income in 1997, and the number of siblings. Parent's education is categorized into high school, some college, college, or graduate degree. Household income is divided into quartile groups. This skill specification is similar to Aucejo and James (2016).[17]

The unobserved components are joint normally distributed and may be correlated with each other as in Prada and Urzúa (2017):

[17] They do not assume any structure on corerlation between θ's. I allow θ^c and θ^{mech} to be correlated. The approach is different, because I do not have many skill measurements.

$$\begin{pmatrix} \theta^c \\ \theta^{mech} \end{pmatrix} \sim N(\mathbf{0}, \Sigma).$$

Theoretically, the distributions of θ's can be identified nonparametrically, and many previous papers using a factor model assume a mixture of normals instead of a normal distribution. However, since my model has five unobserved factors in total, including major-specific skills discussed below, and it is already computationally intensive, I assume normality.

Skill measurement system

Pre-college skills s_1 are not directly observed by the econometrician, but ASVAB test scores are assumed to be noisy measures of skills. Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, Mathematics Knowledge, and Numerical Operation are used to construct a single cognitive skill measure in many papers. I assume that general cognitive skill affects all of the five test scores. Furthermore, I use Mechanical Comprehension, Auto & Shop information, and Electronics Information as noisy measures of mechanical skill. These three are also used as mechanical tests in Prada and Urzúa (2017) and Speer (2017a). General cognitive skill is assumed to have effects on Mechanical Comprehension and Electronics Information test scores as well. Students need to read and understand questions, and many questions also require basic knowledge on calculation.

I assume the following measurement system:

$$WordKnowledge_i = s^c_{1i} + e_{1i} \tag{3.3}$$
$$ParagraphComprehension_i = \delta_{12} s^c_{1i} + e_{2i} \tag{3.4}$$
$$ArithmeticReasoning_i = \delta_{13} s^c_{1i} + e_{3i} \tag{3.5}$$
$$MathematicsKnowledge_i = \delta_{14} s^c_{1i} + e_{4i} \tag{3.6}$$
$$NumericalOperation_i = \delta_{15} s^c_{1i} + e_{5i} \tag{3.7}$$
$$MechanicalComprehension_i = \delta_{16} s^c_{1i} + \delta_{26} s^{mech}_{1i} + e_{6i} \tag{3.8}$$
$$Auto\&ShopInformation_i = s^{mech}_{1i} + e_{7i} \tag{3.9}$$
$$ElectronicsInformation_i = \delta_{18} s^c_{1i} + \delta_{28} s^{mech}_{1i} + e_{8i}, \tag{3.10}$$

Each measure is standardized to have a mean of 0 and a standard deviation of 1, and measurement error e_s, $s = 1, 2, \cdots, 8$, is idiosyncratic with $E(e_s) = 0$, following a normal distribution. Hence, the population average of each type of skill is normalized to 0. The factor loadings on s^c_1 in Word Knowledge and s^{mech}_1 in Auto & Shop Information are normalized to one. This normalization is necessary for identification. Moreover, one of the mechanical tests has to be assumed to be affected only by the mechanical skill because θ^c and θ^{mech} may be correlated. I choose Auto & Shop

Information following Prada and Urzúa (2017). Details on the identification of the measurement system are described in Appendix B.1.[18]

Skill changes through post-secondary education

Skills will change through post-secondary education. The increment to skills will vary by education level and college major. As mentioned above, courses or credits students have to take depend significantly on their college majors. This means that the amount of time students invest on skill accumulation may vary by their college major. Furthermore, I allow skill changes to depend on pre-college skill. For example, even if students take the same course, those with higher pre-college skill might be able to understand the contents more deeply than those with lower pre-college skill. In addition, those who understand the contents better might take advanced level courses further. In these cases, students with higher pre-college skill increase their skill more than those with lower pre-college skill. On the other hand, if a curriculum focuses on making students achieve a certain common level, then students with lower pre-college skill may have to study harder, and so their skill growth might be larger than those with higher pre-college skill.

Let $Some$ denote some college, H denote Humanities & Social Sciences majors, B denote Business & Economics majors, and S denote STEM majors. For individual i from post-secondary education group $m_+ = Some, H, B, S$, I specify post-college general cognitive skill level as

$$s^c_{2m_+,i} = \lambda^c_{0m_+} + \lambda^c_{1m_+} s^c_{1i}. \tag{3.11}$$

Since s^c_1 is normalized to have a mean of 0 over the population, parameter $\lambda^c_{0m_+}$ represents an average growth in education m_+. On the other hand, $\lambda^c_{1m_+}$ shows the effects of pre-college skill level on skill growth.[19] Skill growth from periods 1 to 2 is written as

$$s^c_{2m_+,i} - s^c_{1i} = \lambda^c_{0m_+} + (\lambda^c_{1m_+} - 1) s^c_{1i}.$$

Hence, if $\lambda^c_{1m_+} > 1$, students with higher pre-college skill will accumulate more. On the other

[18] In order to make the skills intuitive, general cognitive and mechanical skills are defined to be affected by observed characteristics x_s in equations (3.1) and (3.2). Instead, general cognitive skill and mechanical skill could be defined as θ^c and θ^{mech}, respectively, and each type of ASVAB test scores could be assumed to be affected by x_s. The parameter estimates will not change much. This is because θ and the observed characteristics in each decision stage are assumed to be orthogonal.

[19] Although post-education mechanical skill will be specified as $s^{mech}_{2i} = \lambda^{mech}_{0,Some} + \lambda^{mech}_{1,Some} s^{mech}_{1i}$ for some college, mechanical skill growth parameters are not estimated below to focus on general cognitive skill growth. I do not consider mechanical skill for those who go to college graduates. As suggested above, the mechanical dimension does not seem important to college graduates.

hand, if $\lambda^c_{1m_+} < 1$, students with lower pre-college skill will accumulate more.[20]

3.3.2.2 Major-specific skills

If they go to college, individuals can increase a major-specific skill. The increment to skill varies by individual. I assume that individuals have potential to increase their post-college major-specific skill if they choose a relevant type of major. The potential skill level is post-college major-specific skill level if they choose the relevant major. Potential post-college major-specific skills are assumed to be unknown to the individuals until they decide to go to college. Although college students will be sorted into major based on their potential post-college major-specific skill levels, that sorting does not affect education level choice. This assumption makes computation easier. At the same time, this assumption is reasonable given that previous papers show that the sorting occurs as college students learn their ability through GPA (see, e.g., Arcidiacono (2004) and Stinebrickner and Stinebrickner (2014)).

For major $m = H, B, S$, let s^m_{2i} denote individual i's potential post-college major m specific skill, that is, major m specific skill level in period 2 if they choose major m. I assume that potential post-college major-specific skills are orthogonal to s^c_1. I further assume that they are orthogonal to each other and normally distributed:

$$\begin{pmatrix} s^H_2 \\ s^B_2 \\ s^S_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_H & 0 & 0 \\ 0 & \sigma^2_B & 0 \\ 0 & 0 & \sigma^2_S \end{pmatrix}\right).$$

The normality is assumed for simplicity. Each type of potential post-college major-specific skill is standardized to have a mean of 0 over the population. Hence, $s^m_{i2} = 0$ denotes that individual i's potential post-college major m specific skill is the same as the population average. If individual i chooses major m, their post-college major m specific skill is equal to s^m_{2i}, while the other types of major-specific skills do not change.

Skill measurement system

I assume that the last two years of college GPA are noisy measures of post-college major-specific skill. Since around 10% of the college students in my data received the maximum GPA score in at least one of the last two years, I allow for ceiling effects following Hansen et al. (2004), who

[20]Skill change equation (3.11) implies that there is no individual heterogeneity in skill changes after conditioning on pre-college skill. Suppose instead that the skill change equation is given by $s^c_{2m_+,i} = \lambda^c_{0m_+} + \lambda^c_{1m_+} s^c_{1i} + \omega_{im_+}$, where ω_{m_+} is orthogonal with s_1 and $E(\omega_{m_+}) = 0$. As long as ω_{m_+} is unknown to the individual in education choices, λ^c's are consistently estimated in my approach. As explained below, I use task intensity equation to identify λ^c's and the parameters in the equation can be estimated consistently.

examine the effects of schooling on cognitive test scores. For college major m, let GPA_{1i}^{*m} and GPA_{2i}^{*m} denote individual i's latent college GPA in major m in their last year and their second last year. I assume that

$$GPA_{1i}^{*m} = \gamma_{01}^{m} + \gamma_{11}^{m} s_{2mi}^{c} + s_{2i}^{m} + x_{gi}' \beta_{g1}^{m} + e_{1i}^{m} \quad (3.12)$$

$$GPA_{2i}^{*m} = \gamma_{02}^{m} + \gamma_{12}^{m} s_{2mi}^{c} + \gamma_{22}^{m} s_{2i}^{m} + x_{gi}' \beta_{g2}^{m} + e_{2i}^{m}. \quad (3.13)$$

Measurement error e_t^m, $t = 1, 2$, is idiosyncratic with $E(e_t^m) = 0$ and follows a normal distribution.[21] The factor loading on s_2^m in the last year of GPA is set to one for identification. Observed variables x_g include a female dummy and region dummies. A female dummy is added, because females tend to earn better GPA than males. Region dummies are included, because college quality might be different by region and college GPA might be standardized within college. For each year t, observed GPA is assumed to be

$$GPA_{ti}^{m} = \overline{GPA_{t}^{m}} \;\; if \; GPA_{ti}^{*m} \geq \overline{GPA_{t}^{m}} \quad (3.14)$$

$$GPA_{ti}^{m} = GPA_{ti}^{*m} \;\; if \; GPA_{ti}^{*m} < \overline{GPA_{t}^{m}}. \quad (3.15)$$

Each observed GPA is standardized to have a mean of 0 and a standard deviation of 1 over those who actually chose major m and $\overline{GPA_{t}^{m}}$ is the upper limit of GPA.

College GPA GPA_t^m, $t = 1, 2$, are observed in the data only if students actually chose major m, and the GPA equations provide only two noisy measures of each type of major-specific skill. Hence, the GPA equations are not enough to identify the distribution of (s_2^H, s_2^B, s_2^S) . The distribution can be identified by additionally using college major choice, which is described later, as another "noisy measure" of major-specific skills. See Hansen et al. (2004) for details on the identification.

Pre-college major-specific skills

In the beginning of period 1, individual i is endowed with pre-college major-specific skills $(s_{1i}^H, s_{1i}^B, s_{1i}^S)$. For each major m, if they choose high school, some college, or major m', $m' \neq m$, then major m specific skill in period 2 is s_{1i}^m. If they choose major m, then major m specific skill evolves from s_{1i}^m to s_{2i}^m. For each major m, if there is a measure of pre-college major-specific skill s_1^m, an approach of comparing s_{1i}^m and s_{2i}^m to identify major m specific skill growth of individual i from major m

[21] Students may increase the major-specific skill between these two periods. I assume that individual i's major m specific skill in the second last year is expressed in the form of $c_0^m + c_1^m s_{2i}^m$, where c_0^m and c_1^m are constants. With regard to general cognitive skill, I assume either a similar specification holds or the skill does not increase between these periods.

would be relatively straightforward. However, my data do not have good measures of pre-college major-specific skills. In section 3.4.2, I discuss how I make some assumptions to specify s_{1i}^m in the absence of good direct measures. However, for expositional purposes, I assume that I have a good measure of s_{1i}^m in the same scale as s_{2i}^m in the rest of this section and describe the rest of my model.

3.3.2.3 Log wage equations

As mentioned above, skills and tasks are not the same in my model; workers possess skills and apply them in tasks. Occupations involving more intense general cognitive tasks will ask workers to use their general cognitive skill intensely, and workers will produce more or higher quality goods in the occupation. Hence, log wage equations include both general cognitive skill and task intensity. Cognitive jobs are further divided into either related or unrelated to each major. The return to major-specific skill can depend on whether the job is related to the relevant major. Let r denote related and nr denote not related.

Let s_i^c denote individual i's general cognitive skill level in entering the labour market. That is, $s_i^c = s_{1i}^c$ if individual i is a high school graduate, while $s_i^c = s_{2m_+,i,}^c$ if individual i is from post-secondary education $m_+ = Some, H, B, S$. Let τ^c denote general cognitive task intensity. I assume the following log wage function for individual i from education $m_- = HS, Some$, who chooses a cognitive type job:

$$\log w_{m_-,i} = \pi_0 + \pi_1 s_i^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + \sum_{m' \in \{H,B,S\}} \pi_{3,nr}^{m'} s_{1i}^{m'} + x_{wi}'\beta_{m_-}, \tag{3.16}$$

Remember that both high school and some college graduates are assumed to have jobs unrelated to any college major.

College graduates can choose job relatedness regarding their college major. For job relatedness $R = r, nr$, I assume the following log wage equation for college graduate i from major m:

$$\begin{aligned}
\log w_{mRi} &= \pi_0 + \pi_{0r}^m \cdot 1(R = r) + \pi_1 s_{2mi}^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + (\pi_{3,nr}^m + \pi_{3mr} \cdot 1(R = r))s_{2i}^m \\
&\qquad + \sum_{m' \neq m} \pi_{3,nr}^{m'} s_{1i}^{m'} + x_{wi}'\beta_m \\
&= \pi_{0mR} + \pi_1 s_{2mi}^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + \pi_{3,R}^m s_{2i}^m + \sum_{m' \neq m} \pi_{3,nr}^{m'} s_{1i}^{m'} + x_{wi}'\beta_m, \qquad (3.17)
\end{aligned}$$

where $\pi_{0mR} = \pi_0 + \pi_{0r}^m \cdot 1(R = r)$, $\pi_{3R}^m = \pi_{3,nr}^m + \pi_{3mr} \cdot 1(R = r)$, and $1(\cdot)$ is an indicator function. Log wages are observed with an additive idiosyncratic error ε_w, following a normal distribution. Control variables, x_w, include a female dummy, race dummies, region dummies, an urban dummy, and cohort dummies.[22] If $\pi^m_{3,nr} < \pi^m_{3r}$, the log wage return to major m specific skill is larger in related jobs than in unrelated jobs. Still, the major-specific skill is useful in unrelated jobs as long as $\pi^m_{3,nr} > 0$.

3.3.2.4 Job type choice, task intensity choice, and job relatedness choice

As mentioned above, high school graduates and some college graduates decide whether to take a cognitive or mechanical type of job when they enter the labour market. For $m_- = HS, Some$, individual i's latent utility of choosing a cognitive type job is

$$I^c_{m_-,i} = \xi^c_0 + \xi^c_1 s^c_i + \xi^c_2 s^{mech}_i + \sum_{m' \in \{H,B,S\}} \xi^c_{3m'} s^{m'}_{1i} + x'_{ci}\beta^c_{m_-} + \varepsilon^c_{m_-,i}. \tag{3.18}$$

Observed variables x_c include a female dummy, race dummies, region dummies, an urban dummy, household income at age 17, and cohort dummies. The latent utility of choosing a mechanical type of job is normalized to 0. I assume that $\varepsilon^c_{m_-}$ follows a Type-I extreme value distribution for simplicity. This assumption gives a standard logit model.

High school and some college graduates who decide to take a cognitive type of job choose general cognitive task intensity. College graduates who enter the labour market do so as well. As mentioned above, the mechanical dimension does not matter in cognitive type jobs. Thus, they only choose general cognitive task intensity. Workers with higher general cognitive skill will choose higher cognitive task intensity because the return may be higher or they may be able to do the task more easily than those with lower skill.

For individual i who possesses general cognitive skill s^c_i and whose education is $m_{++} = HS, Some, H, B, S$, the optimal general cognitive task intensity is assumed to be:

$$\tau^c_{m_{++},i} = \zeta_0 + \zeta_1 s^c_i + x'_{ci}\beta_{\tau m_{++}} + \varepsilon_{\tau m_{++},i}. \tag{3.19}$$

In addition to the observed variables, there is an idiosyncratic shock, $\varepsilon_{\tau m_{++}}$, that follows a normal distribution with mean of 0.[23] This shock is unknown to the individual when they choose their education. This linear specification can be considered as an approximation to a more general form. Or, although I do not estimate the full model, this linear form can be derived from a utility maximization problem (see Appendix B.2).

[22] I allow the coefficients of a female dummy and race dummies to be different across the education groups. I also assume that the coefficients of cohort dummies are common across the college majors.

[23] The coefficients of region dummies, an urban dummy, and household income at age 17 are assumed to be common across the education groups and the coefficients of cohort dummies are assumed to be common across college majors.

College graduates choose job relatedness to their college major. Both major-specific and general cognitive skills can affect the relatedness choice. Latent utility of choosing a related job for individual i from major m is given by

$$I_{r,m,i} = \xi_{0,r,m} + \xi_{1,r,m} s^{c}_{2mi} + \xi_{2,r,m} s^{m}_{2i} + x'_{wi}\beta_{rm} + \varepsilon_{rmi}, \tag{3.20}$$

where ε_{rm} is an idiosyncratic shock and assumed for simplicity to follow a Type-I extreme value distribution. The latent utility of choosing an unrelated job is normalized to 0.

In this specification, I cannot examine whether individuals are in an unrelated job because they want it or because they cannot find a related job. This distinction however does not matter for skill growth estimation.

3.3.2.5 Education level choice and major choice

High school graduates choose one of three options: work, some college, or college. If they choose to work, they enter the labour market with skills $(s^c_1, s^{mech}_1, s^H_1, s^B_1, s^S_1)$. If they decide to go to college, they then choose a college major. In this two-stage education choice framework, some information, such as potential major-specific skill growth during college and exogenous shocks affecting college major choice, is assumed to be revealed after students decide to go to college.

As suggested in the data section, there is sorting into education level and college major based on pre-college general cognitive skill. The cognitive test scores imply that those with a more advanced degree tend to have higher pre-college general cognitive skill; STEM majors tend to have high pre-college general cognitive skill among college graduates. There are two channels that affect sorting. One is that skill development might depend on pre-college skill level. For example, if the increment to the skill in STEM majors increases with pre-college skill, then students with high pre-college skill will be more likely to choose STEM majors. Another is something other than through skill, such as study cost. Even if low-skilled students know that they will increase their skill more in STEM majors, keeping up with the classes or their peers may require them to work very hard. In this case, students with low pre-college skill may prefer to choose an easier major with a smaller skill increase.

Education level choice

High school graduates have three options: no further education, some college, or college. Let Col denote college. Individual i's latent utility of choosing education level $l = HS, Some, Col$, is

$$I_{li} = \eta_{0l} + \eta_{1l}s^{c}_{1i} + \eta_{2l}s^{mech}_{1i} + \sum_{m' \in \{H,B,S\}} \eta_{3,l,m'}s^{m'}_{1i} + x'_{di}\beta_l + z_{dli}\varphi_{dl} + \varepsilon_{dli}. \tag{3.21}$$

Observed variables x_d include the same variables in x_s and cohort dummies. In addition, a local unemployment rate at age 17 enters as an exclusion variable, z_{dl}.[24] For simplicity, I assume that ε_{dl} follows a Type-I extreme value distribution, which gives a standard multinomial logit model. The probability of choosing education level l given the skills and the observed variables is given by

$$\frac{\exp(\eta_{0l} + \eta_{1l}s^{c}_{1i} + \eta_{2l}s^{mech}_{1i} + \sum_{m'} \eta_{3,l,m'}s^{m'}_{1i} + x'_{di}\beta_l + z_{dli}\varphi_{dl})}{\sum_{L\in\{HS,Some,Col\}} \exp(\eta_{0L} + \eta_{1L}s^{c}_{1i} + \eta_{2L}s^{mech}_{1i} + \sum_{m'} \eta_{3,L,m'}s^{m'}_{1i} + x'_{di}\beta_L + z_{dLi}\varphi_{dL})}.$$

The base group is high school.

College major choice

If high school graduates choose to go to college, they then choose a college major. In addition to pre-college general cognitive and major-specific skills, potential major-specific skill growth, which is revealed after deciding to go to college, can affect college major choice. Since the mechanical dimension does not matter to college graduates, mechanical skill does not appear in the college major choice equation.

As in education level choice, I assume linear latent utility with an idiosyncratic shock following a Type-I extreme value distribution.[25] Individual i's latent utility of choosing major m is given by

$$I_{mi} = \eta_{0m} + \eta_{1m}s^{c}_{1i} + \eta_{2m}s^{m}_{2i} + \eta_{3m}s^{m}_{1i} + x'_{di}\beta_m + z'_{mi}\varphi_m + \varepsilon_{mi}. \tag{3.22}$$

Exclusion variables z_m include a foreign born parents dummy, math test score relative to cognitive test score, and mechanical test score relative to cognitive test score.[26] As shown in the summary statistics above, STEM majors tend to have higher pre-college mechanical test scores

[24] The local unemployment rates by education level are constructed from the CPS. The local unit is defined by the combination of regions and MSA residency.

[25] I assume ε_{rm} in equation (3.20) and ε_m are orthogonal. They could be allowed to be correlated, but, in that case, I want to reduce the number of majors to two because of the computation burden.

[26] A foreign born parents dummy is also used in Kinsler and Pavan (2015). The math test score relative to cognitive test score is defined as residuals from regressing average math test scores on a constant and cognitive test scores among college graduates. The constructed score is included as a preference. I assume that, given general cognitive skill, the composition of the skill does not affect wages or skill growth. Similarly, the mechanical test score relative to cognitive test score is defined as residual from regressing average mechanical test score on a constant and cognitive test score among college graduates.

among college graduates.[27] Also, if cognitive tests are divided into verbal tests and math tests, Humanities & Social Sciences majors, on average, have high verbal scores compared to math scores.[28] The probability of choosing major m given the observed variables, pre-college skills, and potential post-college major-specific skills is given by

$$\frac{\exp(\eta_{0m} + \eta_{1m}s^c_{1i} + \eta_{2m}s^m_{2i} + \eta_{3m}s^m_{1i} + x'_{di}\beta_m + z'_{mi}\varphi_m)}{\sum_{n\in\{H,B,S\}}\exp(\eta_{0n} + \eta_{1n}s^c_{1i} + \eta_{2n}s^n_{2i} + \eta_{3n}s^n_{1i} + x'_{di}\beta_n + z'_{ni}\varphi_n)}.$$

The base group is Humanities & Social Sciences majors.

3.4 Identification of skill growth

3.4.1 General cognitive skill

As mentioned above, the distribution of s^c_1 can be identified from the skill measurement system, equations (3.3) to (3.10). The parameters of general cognitive skill growth are identified from general cognitive task intensity choices. Task intensity choice equation (3.19) is written in terms of general cognitive skill brought to the labour market. For high school graduates in a cognitive type job, the equation is

$$\tau^c_{HS,i} = \zeta_0 + \zeta_1 s^c_{1i} + x'_{ci}\beta_{\tau,HS} + \varepsilon_{\tau,HS,i}. \tag{3.23}$$

Hence, ζ_0, ζ_1, and $\beta_{\tau,HS}$ can be estimated from task intensity choice of high school graduates.

For those choosing some college or college, s^c in task intensity choice equation (3.19) is different from the pre-college skill level. Using skill change equation (3.11), the equation can be rewritten in terms of pre-college general cognitive skill s^c_1. For post-secondary education group $m_+ = Some, H, B, S$, the equation can be rewritten as

$$\begin{aligned}\tau^c_{m_+,i} &= (\zeta_0 + \zeta_1\lambda^c_{0m_+}) + \zeta_1\lambda^c_{1m_+}s^c_{1i} + x'_{ci}\beta_{\tau m_+} + \varepsilon_{\tau m_+,i}\\ &= \tilde{\zeta}_{0m_+} + \tilde{\zeta}_{1m_+}s^c_{1i} + x'_{ci}\beta_{\tau m_+} + \varepsilon_{\tau m_+,i},\end{aligned} \tag{3.24}$$

Hence, $\tilde{\zeta}_{0m_+}$, $\tilde{\zeta}_{1m_+}$, and $\beta_{\tau m_+}$ can be identified. Since $\tilde{\zeta}_{0m_+} = \zeta_0 + \zeta_1\lambda^c_{0m_+}$ and $\tilde{\zeta}_{1m_+} = \zeta_1\lambda^c_{1m_+}$, the

[27] One might think that mechanical skill increases in STEM majors. However, as shown above, the average mechanical task intensity among STEM majors is smaller than that among high school graduates and it does not seem mechanical skill is not important for STEM majors. Hence, pre-college mechanical skill is included in preferences here.

[28] Among Humanities & Social Sciences majors, the mean of the average score of the verbal tests is 0.64 and that of the math tests is 0.53. The means are 0.51 and 0.68 among Business & Economics majors, and 0.73 and 0.89 among STEM majors.

skill growth parameters are written as

$$\lambda^c_{0m_+} = \frac{\tilde{\zeta}_{0m_+} - \zeta_0}{\zeta_1}$$
$$\lambda^c_{1m_+} = \frac{\tilde{\zeta}_{1m_+}}{\zeta_1}.$$

Parameters ζ_0, ζ_1, $\tilde{\zeta}_{0m_+}$, and $\tilde{\zeta}_{1m_+}$ are identified as mentioned above. Hence, $\lambda^c_{0m_+}$ and $\lambda^c_{1m_+}$ in the skill change equation $s^c_{2m_+,i} = \lambda^c_{0m_+} + \lambda^c_{1m_+} s^c_{1i}$ can be identified.

Log wage point metric

Since there is no natural unit of skills, it is difficult to interpret the results without some reference point. With regard to general cognitive skill, multiplying both sides of the skill change equation (3.11) by π_1, which is the coefficient of general cognitive skill in log wage equation (3.16), gives

$$\pi_1 s^c_{2m_+,i} = \pi_1 \lambda^c_{0m_+} + \lambda^c_{1m_+} \pi_1 s^c_{1i},$$

for $m_+ = Some, H, B, S$. By construction, a one unit increase in $\pi_1 s^c_{1i}$ will increase log wage by 1 point. Skill growth from periods 1 to 2 is written as

$$\pi_1 (s^c_{2m_+,i} - s^c_{1i}) = \pi_1 \lambda^c_{0m_+} + (\lambda^c_{1m_+} - 1) \pi_1 s^c_{1i}. \tag{3.25}$$

In this form, skill growth is interpreted in terms of log wage points, that is, by how many points the skill growth will increase log wages. Although general cognitive skill growth can be compared across education groups without this transformation, this log wage point transformation makes interpretation easier.

3.4.2 Major-specific skills

The distributions of potential post-college major-specific skills are identified from GPA equations (3.12) to (3.15) and college major choice equation (3.22). If I had measures of pre-college major-specific skill s^m_1 in the same metric as s^m_2, major m specific skill growth could be measured by $s^m_2 - s^m_1$. However, as mentioned above, I do not have measures of s^m_1, and I need to make some assumptions on s^m_{1i}.

I assume that pre-college major-specific skill levels are the same across individuals. That is, for major m, $s^m_{1i} = s^m_1$ for individual i. Under this assumption, terms of pre-college major-specific skill are absorbed in constant terms in the equations shown in the previous section. Log wage equation for education groups $m_- = HS, Some$, (3.16), is rewritten as

$$\begin{aligned}\log w_{m_-,i} &= \left(\pi_0 + \sum_{m' \in \{H,B,S\}} \pi_{3,nr}^{m'} s_1^{m'}\right) + \pi_1 s_i^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + x_{wi}'\beta_{m_-} \\ &= \tilde{\pi}_0 + \pi_1 s_i^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + x_{wi}'\beta_{m_-}. \end{aligned} \tag{3.26}$$

Log wage equation for major $m = H, E, S$, (3.17), is rewritten as

$$\begin{aligned}\log w_{mRi} &= \left(\pi_{0mR} + \sum_{m' \neq m} \pi_{3,nr}^{m'} s_1^{m'}\right) + \pi_1 s_{2mi}^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + \pi_{3,R}^m s_{2i}^m + x_{wi}'\beta_m \\ &= \tilde{\pi}_{0mR} + \pi_1 s_{2mi}^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2 + \pi_{3,R}^m s_{2i}^m + x_{wi}'\beta_m. \end{aligned} \tag{3.27}$$

Parameters $\pi_{3,nr}^m$, $m = H, B, S$, cannot be identified, but that does not matter to the identification of the skill growth. Job type choice equation (3.18), education level choice equation (3.21), and college major choice equation (3.22) can be rewritten in the same way. The rewritten equations are described in Appendix B.3.

Of course, s_1^m cannot be identified from these equations. I approximate s_1^m to the skill level, with which male high school graduates living in Northeast region and whose pre-college general cognitive skill level is the population average would receive zero GPA in the 0-4.0 scale for courses taken in last year of college.[29]

Log wage point metric

Unlike general cognitive skill, major-specific skill growth cannot be directly compared across college majors because each major increases its own type of skill, which is measured in its own scale. Major-specific skill growth, of course, cannot be directly compared with general cognitive skill growth either. As with general cognitive skill growth, I transform major-specific skill growth into a log wage point metric. However, the coefficient of major-specific skill can be different by job relatedness to the major, and which coefficient is used can significantly affect the result. I transform major-specific skill growth into log wage point metric in three ways.

The first one is $\pi_{3r}^m(s_2^m - s_1^m)$ for each major m, which is log wage contribution of major-specific skill growth in related jobs. In order to see the sensitivity with the assumption on pre-college major-specific skill level, I also calculate $\pi_{3r}^m(s_2^m - s_{2m,10}^m)$ as the second way, where $s_{2m,10}^m$ denotes the tenth percentile of s_2^m of those who choose major m.

[29] Using another group of people, such as females or people with high pre-college skill, does not change my results on the contribution of major-specific skill growth on wage growth.

Although the return to a major-specific skill in related jobs will be larger than that in unrelated jobs, the wage may be low if their major-specific skill level is not large enough. If their major-specific skill level is not high, college graduates can choose an unrelated job to avoid low wage in a related job. On the other side of the coin, evolved major-specific skill can be considered to provide college graduates an option to earn higher wages in related jobs than in unrelated jobs if their skill is high. Based on this perspective, as the third way, I calculate the contribution of major-specific skill growth on log wage given that individuals choose a job relatedness that brings them higher wages. From the point of log wage value of skill growth, this skill growth transformation will be appropriate to compare with general cognitive skill growth.

I transform the skill growth into the log wage point metric as follows. Suppose $\pi^m_{3r} > \pi^m_{3,nr}$ in log wage equation (3.27). In this case, there is $\bar{s}^m$ such that log wage in a related job is larger than that in an unrelated job if and only if $s^m > \bar{s}^m$. Major m specific skill growth in log wage point metric is calculated as

$$0 \;\; if \; s^m_1 > s^m_2 \tag{3.28}$$

$$\pi^m_{3nr}(s^m_2 - s^m_1) \;\; if \; s^m_1 \leq s^m_2 \leq \bar{s}^m \tag{3.29}$$

$$\pi^m_{3,nr}(\bar{s}^m - s^m_1) + \pi^m_{3r}(s^m_2 - \bar{s}^m) \;\; if \; s^m_1 \leq \bar{s}^m < s^m_2 \tag{3.30}$$

$$\pi^m_{3r}(s^m_2 - s^m_1) \;\; if \; \bar{s}^m < s^m_1 \leq s^m_2. \tag{3.31}$$

Equation (3.28) means that, if $s^m_1 > s^m_2$, skill growth is calculated to be zero. Equation (3.29) indicates a situation in which college graduates would take an unrelated job even after graduating from college. Hence, the skill growth is calculated as log wage growth from major-specific skill growth in unrelated jobs. In equation (3.30), college graduates with pre-college level of major-specific skill would take an unrelated job, but, with post-college level of major-specific skill, they would take a related job. Hence, the first term indicates the log wage increase from skill growth up to $\bar{s}^m$ in an unrelated job, while the second term indicates the log wage increase because of skill growth from $\bar{s}^m$ to s^m_2 in a related job. In equation (3.31), even college graduates with pre-college level major-specific skill would take a related job. Hence, the skill growth is calculated as log wage increase from major-specific skill growth in related jobs.

3.5 Estimation

I estimate my model via maximum likelihood. My model has five types of unobserved skills, general cognitive skill, mechanical skill, and three types of major-specific skills. The unobserved skills need to be integrated out in estimation. Post-secondary general cognitive skill $s^c_{2m_+}$, $m_+ = Some, H, E, S$, is unobserved, but can be rewritten in terms of s^c_1 by using skill growth equation

(3.11). The equations I estimate are written in terms of s_1^c in stead of $s_{2m_+}^c$.

For college major m, latent GPA equations (3.12) and (3.13) can be rewritten as

$$
\begin{aligned}
GPA_{1i}^{*m} &= (\gamma_{01}^m + \gamma_{11}^m \lambda_{0m}^c) + \gamma_{11}^m \lambda_{1m}^c s_{1i}^c + s_{2i}^m + x_{gi}' \beta_{g1}^m + e_{1i}^m \\
&= \tilde{\gamma}_{01}^m + \tilde{\gamma}_{11}^m s_{1i}^c + s_{2i}^m + x_{gi}' \beta_{g1}^m + e_{1i}^m && (3.32) \\
GPA_{2i}^{*m} &= (\gamma_{02}^m + \gamma_{12}^m \lambda_{0m}^c) + \gamma_{12}^m \lambda_{1m}^c s_{1i}^c + \gamma_{22}^m s_{2i}^m + x_{gi}' \beta_{g2}^m + e_{2i}^m \\
&= \tilde{\gamma}_{02}^m + \tilde{\gamma}_{12}^m s_{1i}^c + \gamma_{22}^m s_{2i}^m + x_{gi}' \beta_{g2}^m + e_{2i}^m. && (3.33)
\end{aligned}
$$

Log wage equation (3.26) for some college is given by

$$
\begin{aligned}
\log w_{Some,i} &= (\tilde{\pi}_0 + \pi_1 \lambda_{0,Some}^c) + \pi_1 \lambda_{1,Some}^c s_{1i}^c + \pi_{21} \tau_i^c + \pi_{22} (\tau_i^c)^2 + x_{wi}' \beta_{Some} \\
&= \tilde{\pi}_{0,Some} + \pi_{1,Some} s_{1i}^c + \pi_{21} \tau_i^c + \pi_{22} (\tau_i^c)^2 + x_{wi}' \beta_{Some}, && (3.34)
\end{aligned}
$$

and, for major m, log wage equation (3.27) can be written as

$$
\begin{aligned}
\log w_{mRi} &= (\tilde{\pi}_{0mR} + \pi_1 \lambda_{0m}^c) + \pi_1 \lambda_{1m}^c s_{1i}^c + \pi_{21} \tau_i^c + \pi_{22} (\tau_i^c)^2 + \pi_{3,R}^m s_{2i}^m + x_{wi}' \beta_m \\
&= \dot{\pi}_{0mR} + \pi_{1m} s_{1i}^c + \pi_{21} \tau_i^c + \pi_{22} (\tau_i^c)^2 + \pi_{3,R}^m s_{2i}^m + x_{wi}' \beta_m. && (3.35)
\end{aligned}
$$

The other rewritten equations, job type choice equation for some college and job relatedness choice equation, are described in Appendix B.4.[30] Therefore, the unobserved skills that need to be integrated out are s_1^c, s_1^{mech}, s_2^H, s_2^B, and s_2^S.[31]

I estimate the model in three stages.[32] The first stage estimates the ASVAB equations, the education level choice equation, and the job type choice equations. Hence, the distributions of s_1^c and s_1^{mech} are estimated in this stage. Using the parameter estimates in the first stage, the second stage estimates the GPA equations and the college major choice equation. The distributions of s_2^H, s_2^B, and s_2^S are estimated in this stage. Given the parameter estimates in the first and second stages, the general cognitive task intensity choice equations, the job relatedness choice equation, and the log wage equations are estimated in the third stage. This three-stages approach is less efficient than a one-stage approach. However, this approach not only makes computation easier but also makes the identification of the skills more transparent.

[30] In the latent GPA equations, parameters that can be identified are $\tilde{\gamma}_{01}^m$, $\tilde{\gamma}_{11}^m$, $\tilde{\gamma}_{02}^m$, and $\tilde{\gamma}_{12}^m$. Hence, λ_{0m}^c and λ_{1m}^c cannot be identified from these equations. The latent GPA equations are not used to identify the skill growth parameters.

[31] I use Gauss-Hermeite quadrature to numerically evaluate the integral. The order of quadrature is 10.

[32] Appendix B.5 describes the estimated equations concretely.

3.6 Results

The parameter estimates that are not shown below are reported in Appendix B.6.[33]

ASVAB test score equations

Figure 3.5 shows the variance decomposition of ASVAB test scores. Around 70% of the variance of Word Knowledge scores, Paragraph Comprehension scores, Arithmetic Reasoning scores, and Math knowledge scores is explained by the variance of general cognitive skill, which is the sum of the variances of observed and unobserved general cognitive components. Numerical Operation scores are noisier compared to the other cognitive test scores. Less than 40% of its variance can be explained by the variance of general cognitive skill.

With regard to the mechanical test scores, 20% of the variance of Mechanical Comprehension scores and of Electronics Information scores can be explained by the variance of general cognitive skill. The variance of mechanical skill, which is the sum of the observed and unobserved mechanical components, explains 20% to 30% of the variance of each type of test scores.

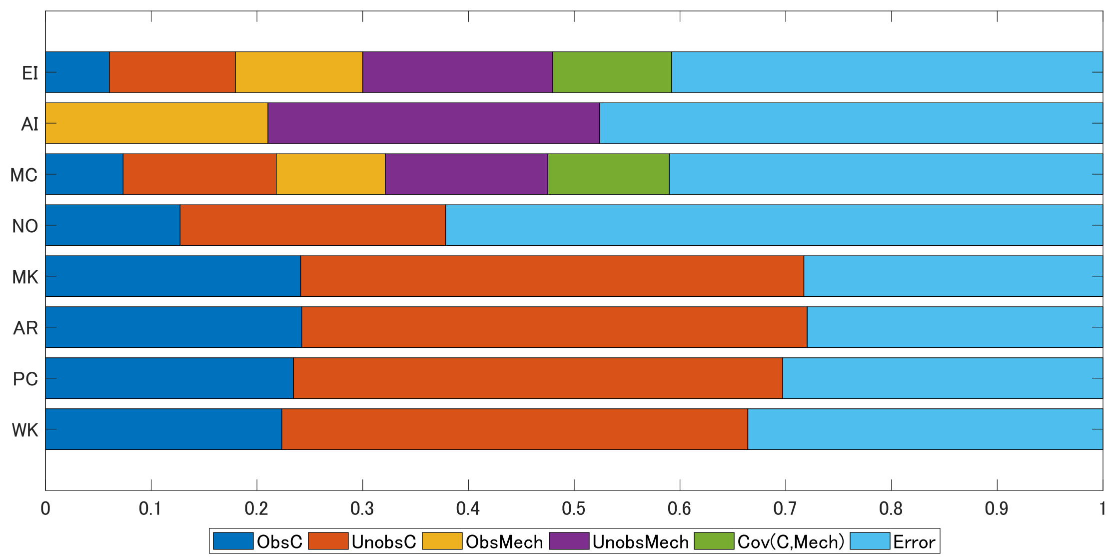

Figure 3.5: Variance decomposition of ASVAB test scores

Notes: Skill equations are $s^c_{1i} = x'_{si}\alpha^c + \theta^c_i$ and $s^{mech}_{1i} = x'_{si}\alpha^{mech} + \theta^{mech}_i$ and the ASVAB equations are equations (3.3) to (3.10). WK: Word Knowledge; PC: Paragraph Comprehension; AR: Arithmetic Reasoning; MK: Mathematics Knowledge; NO: Numerical Operation; MC: Mechanical Comprehension; AI: Auto & Shop Information; EI: Electronics Information. For $s = 1, 2, \cdots, 8$, let $\delta_{11} = 1$, $\delta_{17} = 0$, $\delta_{21} = \delta_{22} = \cdots = \delta_{25} = 0$, and $\delta_{27} = 1$. ObsC: $Var(\delta_{1s}x'_s\alpha^c)$; UnobsC: $Var(\delta_{1s}\theta^c)$; ObsMech: $Var(\delta_{2s}x'_s\alpha^{mech})$; UnobsMech: $Var(\delta_{2s}\theta^{mech})$; Cov(C,Mech): $Cov(\delta_{1s}s^c_1, \delta_{2s}s^{mech}_1)$; Error: $Var(e_s)$.

[33] The brief summary of the job type choice and the job relatedness choice is the following: with regard to job type choice, those who have high pre-college general cognitive skill and who have low pre-college mechanical skill tend to be sorted into cognitive type jobs in both high school and some college graduates. This sorting is stronger among high school graduates, which might suggest that, even though some college graduates work in a mechanical job, some college mainly increases general cognitive skill and mechanical skill becomes less important to them compared to high school graduates.

GPA equations

The last two years of college GPA are used as noisy measures of major-specific skills. Figure 3.6 shows the variance decomposition of the latent college GPA. The last year of latent GPA and the second last year of latent GPA show somewhat different results. The variance of measurement errors is relatively large in the second last year latent GPA of Humanities & Social Sciences and Business & Economics majors. In STEM majors, the two measures are similar.

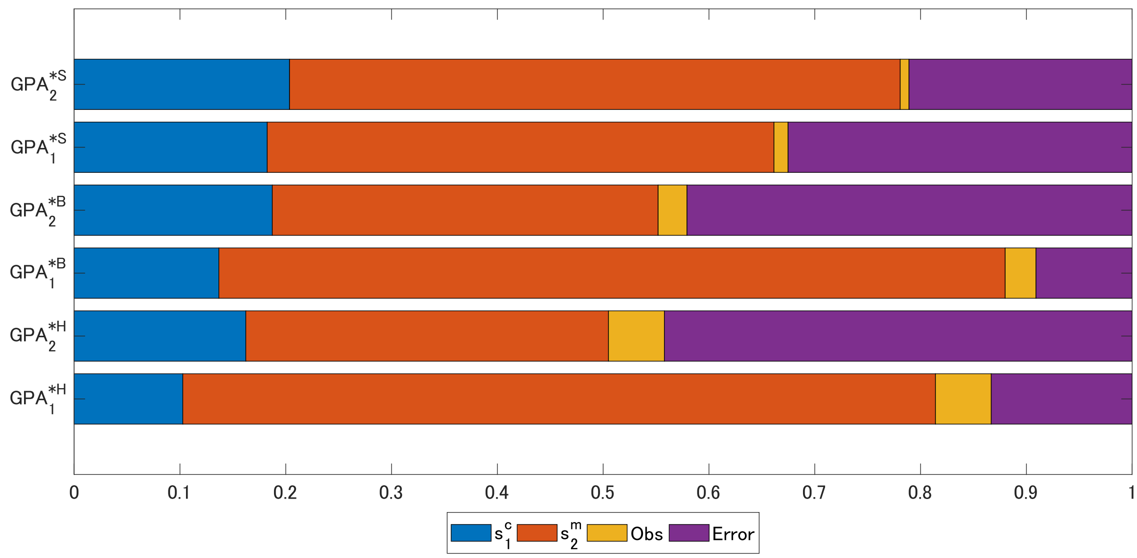

Figure 3.6: Variance decomposition of latent college GPA

Notes: The latent GPA equations are equations (3.32) and (3.33). For $t = 1, 2$, s_1^c: $Var(\gamma_{1t}^m s_1^c)$; s_2^m: $Var(\gamma_{2t}^m s_2^m)$, where $\gamma_{21}^m = 1$; Obs: $Var(x_g' \beta_{gt}^m)$, Error: $Var(e_t^m)$.

In any major, a large part of the variance of GPA is explained by the variance of major-specific skill. This implies that college students learn something that cannot be captured by general cognitive skill and that the skills are heavily weighed in grading. Students with higher pre-college general cognitive skill tend to earn higher GPA in any college major, but the impacts are smaller than those of major-specific skill.

Sorting into education level and major based on skills

Figure 3.7 shows average pre-college skill levels conditional on education level. Each type of skill is standardized to have a mean of 0 and a standard deviation of 1 in this figure. Students who choose more education tend to have higher pre-college cognitive and mechanical skills. However, given pre-college general cognitive skill, those who have higher pre-college mechanical skill tend to choose high school. This is consistent with Prada and Urzúa (2017).

Figure 3.8 shows the average pre-college general cognitive skill levels and the average levels of potential post-college major-specific skills by college major. In this figure, each type of skill is

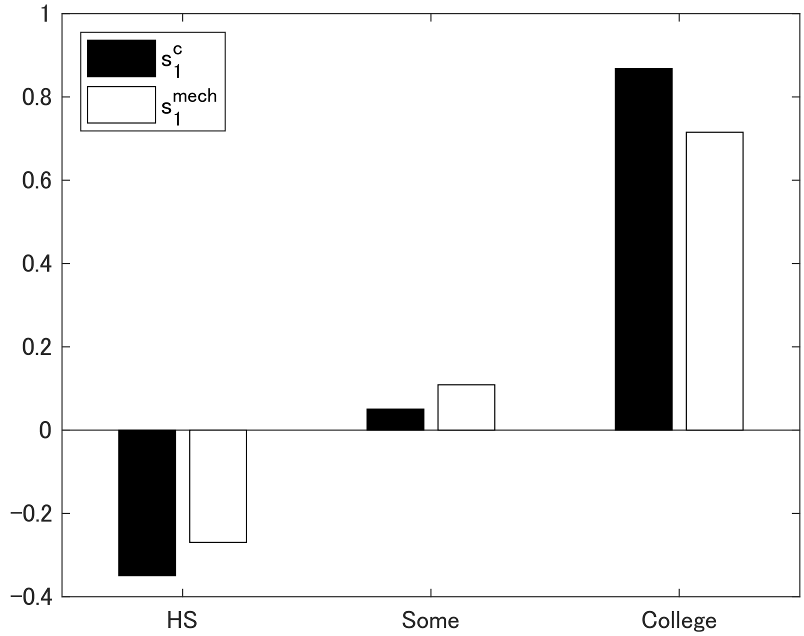

Figure 3.7: Sorting into education level based on pre-college skills

Note: Each type of skill is standardized to have a mean of 0 and a standard deviation of 1 in this figure. HS: High school; Some: Some college.

standardized to have a mean of 0 and a standard deviation of 1 over the population. STEM majors tend to have higher pre-college general cognitive skill than the other two majors. With regard to potential post-college major-specific skills, students are positively sorted into Humanities & Social Sciences and STEM majors based on their respective major-specific skill. Especially, students choosing STEM majors tend to have much higher potential post-college STEM major specific skill than the average. In contrast, students are negatively sorted into Business & Economics majors. They tend to be below average in all types of potential post-college major-specific skills. This negative selection on major-specific skill into Business & Economics majors seems counter intuitive. However, a counterfactual analysis in Kinsler and Pavan (2015) shows that the average return to business major is smallest for those who choose business among college graduates. My result appears consistent with that.

Task intensity choice

Table 3.3 reports parameter estimates of the general cognitive task intensity choice equations, (3.23) for high school graduates and (3.24) for the other education groups. Constant terms represent the average level of cognitive task intensity taken by those with the population average level of pre-college general cognitive skill. The estimates of the constant are substantially different across the education groups. People with a higher level of education tend to take a job involving more intense general cognitive tasks. This suggests that a higher level of education will increase general

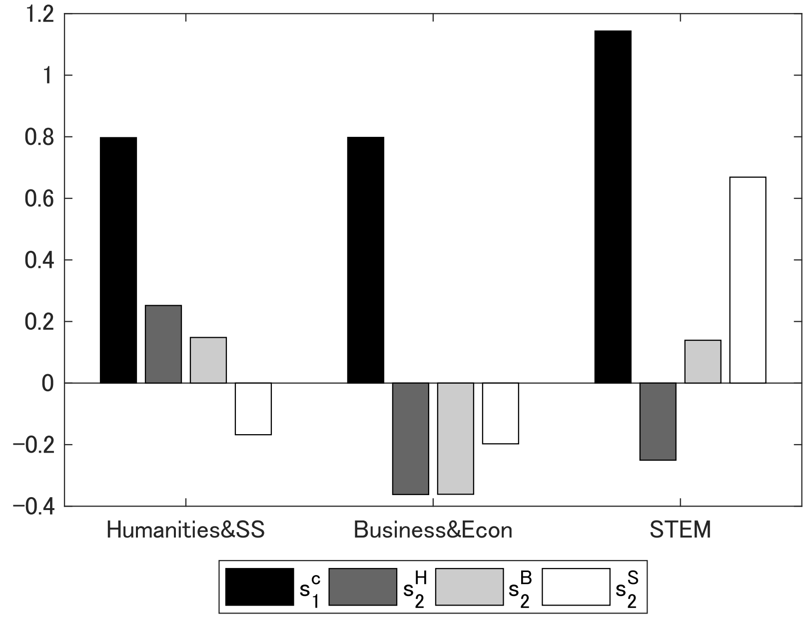

Figure 3.8: Sorting into major based on skills

Note: Each type of skill is standardized to have mean 0 and standard deviation 1 in this figure.

cognitive skill more than a lower level of education. Also, the largest constant for STEM majors implies that general cognitive skill increases the most in STEM majors.

The estimate of ζ_1 is positive. This means that general cognitive skill has positive effects on general cognitive task choice. In every education group, the sign of pre-college general cognitive skill is positive. This implies that those with higher pre-college general cognitive skill are more likely to end up in an occupation involving more intense cognitive tasks. There are some differences in size across the education groups, and these differences reflect the differences in the effects of pre-college skill level on skill growth.

Log wage equations

The parameter estimates of log wage equations, (3.26) for high school graduates, (3.34) for some college graduates, and (3.35) for college graduates, are reported in Table 3.4. Parameter π_1 is estimated to be positive, which means that general cognitive skill has a positive effect on log wages. The positive coefficients of task intensity terms show that wages are higher in occupations involving more intense cognitive tasks.

In all three majors, major-specific skills are unrewarded in jobs unrelated to their respective majors. In Humanities & Social Sciences majors, major-specific skill does not have a positive effect on wages even in jobs related to their majors. Skills that are specifically acquired in Humanities & Social Sciences majors do not appear to be rewarded in the labour market. In Business & Economics and STEM majors, although they are insignificant at the 5% significant level because

Table 3.3: Cognitive task intensity choice parameter estimates

	Const		s_1^c	
High school	(ζ_0)	-0.5865	(ζ_1)	0.1694
		(0.1107)		(0.0333)
Some college	($\tilde{\zeta}_{0Some}$)	-0.1024	($\tilde{\zeta}_{1Some}$)	0.1664
		(0.1121)		(0.0577)
Humanities&SS	($\tilde{\zeta}_{0H}$)	0.4507	($\tilde{\zeta}_{1H}$)	0.0085
		(0.1350)		(0.0462)
Business&Econ	($\tilde{\zeta}_{0B}$)	0.7450	($\tilde{\zeta}_{1B}$)	0.1716
		(0.2314)		(0.0805)
STEM	($\tilde{\zeta}_{0S}$)	1.1397	($\tilde{\zeta}_{1S}$)	0.2614
		(0.3070)		(0.1048)

Notes: Equation (3.23) for high school and equation (3.24) for the other education groups. Other parameter estimates are omitted in this table.
Standard errors are in parentheses.
Standard errors are calculated by bootstrap. The number of replication is 200.

of the sample size, the estimates of the coefficient of the major-specific skill are positive in related jobs.

3.6.1 Skill growth

3.6.1.1 General cognitive skill growth

As explained in Section 5, the general cognitive task intensity choices are used to estimate the skill growth parameters in equation (3.25). Figure 3.9 shows general cognitive skill growth across pre-college skill levels. As can be expected, every major shows higher skill growth than some college. Among college majors, STEM majors show substantially larger skill growth than the other two majors regardless of pre-college skill levels. At the population average pre-college skill level, the difference between STEM majors and Humanities & Social Sciences majors is 16 log wage points. There are some differences in how much pre-college skill levels matter to skill growth. However, even though students with lower pre-college skill levels have a weaker monetary incentive to choose STEM majors, majoring in STEM will still bring them higher growth in general cognitive skill. Table 3.5 shows that, for any education group of people, average general cognitive skill growth will be the highest in STEM majors.

Although I cannot identify this from my model, the observed strong sorting based on pre-college general cognitive skill might be due to study cost differences by college major. Students may be able to accumulate a larger amount of general cognitive skill in STEM majors. However, this might mean that students have to study much harder than in the other college majors. Despite

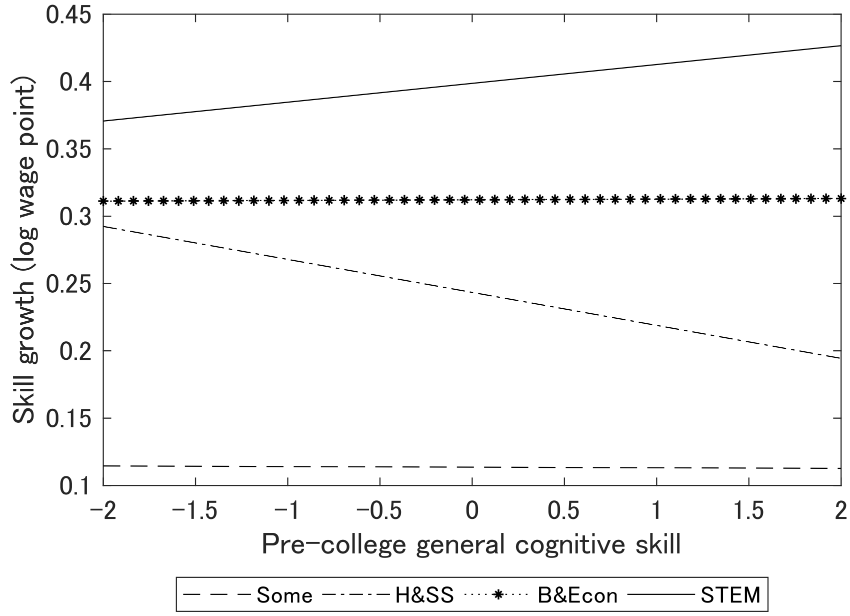

Figure 3.9: General cognitive skill growth by education group

Notes: Pre-college general cognitive skill in the x-axis is standardized to have mean 0 and standard deviation 1. General cognitive skill growth of individual i from education groups m_+, $m_+ = Some, H, B, S$, is given by $\pi_1(s^c_{2m_+,i} - s^c_{1i}) = \pi_1\lambda_{0m_+} + (\lambda^c_{1m_+} - 1)\pi_1 s^c_{1i}$.

the fact that STEM majors have higher pre-college cognitive test scores on average, they tend to spend more time in studying than other majors (see, e.g., Brint et al. (2012) and Ahn et al. (2018)). Keeping up in classes and with peers may be very difficult for those who do not have enough background or preparation.

3.6.1.2 Occupation choice and wages

General cognitive skill growth in the log wage point metric indicates the direct contribution of general cognitive skill growth on log wages. There is also an indirect contribution of skill growth through occupation choice. As seen above, workers with higher general cognitive skill tend to choose a job involving more intense general cognitive tasks. Also, wages in jobs involving more intense general cognitive tasks tend to be higher.

In order to examine the size of the indirect effects, I estimate the log wage equations excluding task intensity terms. The parameter estimates are reported in Table B.9 in Appendix B.6. Now the estimate of π_1 in equation (3.26) without the task intensity terms, $\hat{\pi}_{1,-\tau}$, means the effects of general cognitive skill, including the indirect effects through task intensity choice. Since $\hat{\pi}_1$, which is the estimate of π_1 in equation (3.26) with the task intensity terms, only includes the direct effects of general cognitive skill, $(\hat{\pi}_{1,-\tau} - \hat{\pi}_1)(s^c_2 - s^c_1)$ measures the indirect effects of general cognitive skill. Figure 3.10 shows the indirect effects of general cognitive skill growth on log wages. The indirect

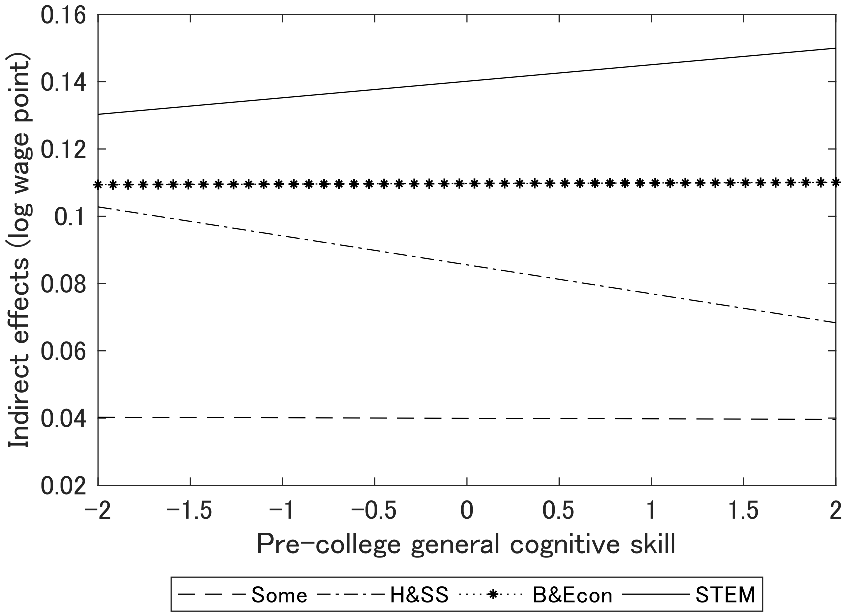

Figure 3.10: Indirect effects of general cognitive skill growth on wages through task intensity choice

Note: Pre-college general cognitive skill in the x-axis is standardized to have mean 0 and standard deviation 1. The effects for individual i from education groups m_+, $m_+ = Some, H, B, S$, is calculated by $\hat{\pi}_{1,\tau}(s^c_{2m_+,i} - s^c_{1i}) - \pi_1(s^c_{2m_+,i} - s^c_{1i})$, where $\hat{\pi}_{1\tau}$ is the estimate of π_1 in equation (3.26) excluding task intensity terms.

effects are smaller than the direct effects. The direct effects are about three times larger. STEM majors show the largest indirect effects among the majors.

3.6.1.3 Major-specific skill growth

Using the estimate of $\pi^m_{3,r}$ and the estimated standard deviation of s^m_2, the log wage effect of one standard deviation increase in s^m_2 in related jobs can be calculated for each major m.[34] Since the estimate of π^H_{3r} is almost zero, I set it to 0. Then, the effect is 0 for Humanities & Social Sciences majors, 0.0390 for Business & Economics majors, and 0.1292 for STEM majors. The log wage effect of a major-specific skill in related jobs is the largest for STEM majors. It is more than three times larger than for Business & Economics majors.

In symmetry with general cognitive skill, I have measures of post-college major-specific skill levels, but do not have a measure of pre-college skill levels. In the absence of such a measure, I approximate pre-college major-specific skills to levels, with which an average male high school graduate living in Northeast region who has the population average general cognitive skill would fail all courses taken in the last year of college.[35] Table 3.6 shows the calculated pre-college levels

[34]Note that the standard deviations of potential major-specific skills are not set to 1. They can be different by type of major.

[35]Using another group of people, such as females, people with high pre-college skill, or different regions, does not change my results on the contribution of major-specific skill growth on log wage growth because the calculated

of each major-specific skill. The metric of s_1^m is the same as that of s_2^m. For each major m, potential post-college major-specific skill s_2^m is standardized to have a mean of 0 and a standard deviation of 1 over the population.

Table 3.7 reports major-specific skill growth measured in related jobs. It is evaluated at the population average level of potential post-college major-specific skill, that is, $s_2^m = 0$ for each major m. As a reference point, pre-college major-specific skill s_1^m is used in the first row, while the tenth percentile of post-college skill level among those who choose the major $s_{2m,10}^m$ is used in the second row. The skill growth is calculated at the population mean of s_2^m for each major m. Business & Economics and STEM majors show huge differences between the two cases. With regard to STEM majors, the major-specific skill growth is larger than that of general cognitive skill growth in the first case. On the other hand, it is much smaller in the second case and is even smaller than that of Business & Economics majors. This large reduction is because, as seen in Figure 3.8, students who choose STEM majors tend to have high potential post-college STEM major specific skill and $s_{2S,10}^S$ is much larger than s_1^S.

I then calculate major-specific skill growth following equations (3.28) to (3.31). Although I call it major-specific skill growth, this growth represents the contribution of major-specific skill growth on log wage growth given that individuals choose the higher-paying job relatedness. Hence, even if skill itself increases, that is, $s_2^m > s_1^m$, the calculated skill growth measure can be zero if s_2^m is not rewarded in the chosen job. Figure 3.11 shows major-specific skill growth across potential post-college major-specific skill levels. As shown in Table 3.4, the returns to major-specific skills are almost zero in unrelated jobs for all majors and in related jobs for Humanities & Social Sciences majors. Hence, I set them to 0 for the figure. In order to make it easy to compare the growth across majors, the x-axis is standardized to mean 0 and standard deviation 1. Since a large part of people would choose unrelated jobs, their major-specific skills are not utilized and the contribution on wage growth is zero among them. Therefore, the average major-specific skill growth is much smaller than general cognitive skill growth (see Table 3.8).

Major-specific skill growth has small positive effects on wages among Business & Economics majors. STEM majors show larger effects. If the skill level is 1.5 standard deviations above the population mean, then the growth in STEM major specific skill increases wages by about 10%. Although major-specific skill growth brings positive wage returns to some people, the returns are small relative to those from general cognitive skill growth. For example, even if STEM major specific skill is 1.5 standard deviations above the population average, growth in general cognitive skill contributes three to four times more than growth in major-specific skill. The skill growth estimates are robust with s_1^m because they do not change as long as $s_1^m \leq \bar{s}^m$ and $\bar{s}^m$ is more than 0.5 standard deviation above the average of s_2^m in any major m.

pre-college levels in any group are very small compared to post-college levels.

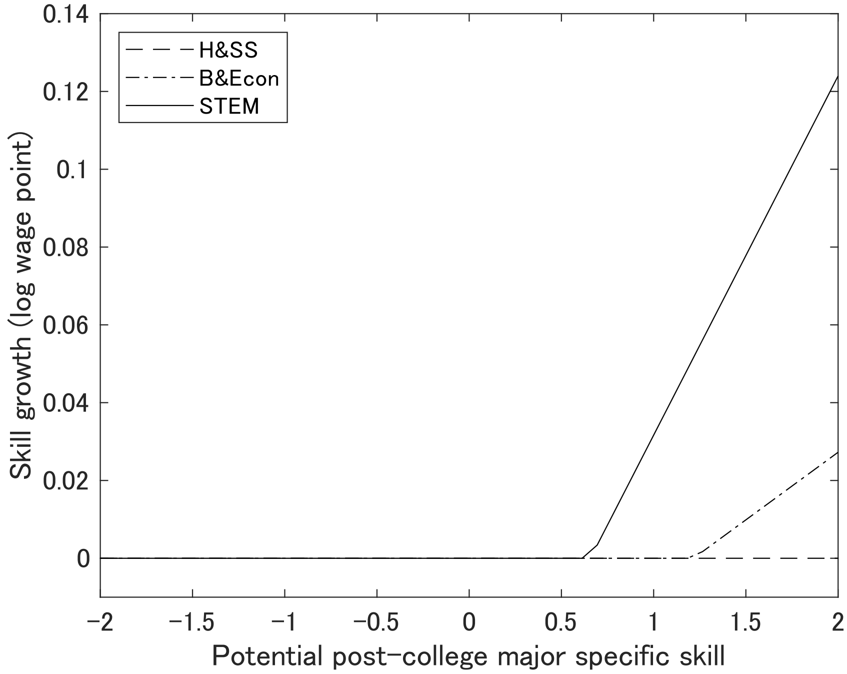

Figure 3.11: Major-specific skill growth across post-college major specific skill

Note: Potential post-college major specific skill in the x-axis is standardized to have mean 0 and standard deviation 1. See equations (3.28) to (3.31) for the definition of the skill growth.

Since I take a different approach from Lemieux (2014), who takes a wage decomposition approach, my results cannot be directly compared to his results. Still, my estimates on major-specific skill growth are small compared to his. There are several reasons. First, Lemieux (2014) uses college workers aged 25 to 64 and high school workers aged 29 to 59. It is plausible that college workers increase their major-specific skill on the job. Second, he does not consider individual heterogeneity in major-specific skill growth, but people with a high level of major-specific skill may have a job related to their major. If this is true, major-specific skill growth is overestimated in his paper. A third reason is the coarse occupation classification in his paper. Related and unrelated jobs may be different in general cognitive task intensity even if they are both categorized in the same occupation in his paper. If unrelated jobs tend to involve less intensive general cognitive task, his estimated major-specific skill growth may partially reflect general cognitive skill growth and task intensity choice effects.[36]

Major-specific skill growth in log wage point metric depends on their estimated wage returns. Since skill price and skill quantity cannot be separately identified, I cannot conclude whether Humanities & Social Sciences majors learn something that is not rewarded in the labour market or that they learn little other than general cognitive skill. However, given that the variance decomposition

[36]Other reasons are the different countries, Canada and US, and the different types of job relatedness measures, worker's self-assessed measure and job analysis measure.

of GPA is not much different across college majors, I suspect that Humanities & Social Sciences majors learn something other than general cognitive skill as the other majors.[37] If students mainly develop general cognitive skill in Humanities & Social Sciences majors, a large part of their GPA variances should be explained by the variance of general cognitive skill.

3.7 Conclusion

Large income differences across college majors are attracting a lot of attention, and many studies examine whether there exist wage differentials across majors even when controlling for selection or sorting into majors based on ability or skills. This paper examines where wage differentials across college majors comes from, by estimating skill growth by major in a multi-dimensional skill framework. Credits or courses college students have to take substantially depend on their major. Hence, college students may accumulate different types and amounts of skills by college major, and examining the similarity and the uniqueness of the accumulated skills in different majors will be helpful for understanding differences across majors deeply.

I assume that each major increases two types of skills: general cognitive skill and a major-specific skill. General cognitive skill can increase in any major. This skill captures the similarity of the skills accumulated across college majors. On the other hand, major-specific skill can increase only in its relevant major. This skill captures the uniqueness of the skills acquired in the major. I allow for individual heterogeneity in growth in both types of skills.

I use the NLSY97 for individual data and the O*NET for occupation data. As in most datasets, the NLSY97 does not contain test scores of each type of skill at both pre- and post-college periods. Therefore, I utilize ASVAB test scores, which were taken in pre-college period, and post-education occupation choice, which reflects post-education skill level, for general cognitive skill. For major-specific skills, I make some assumptions on pre-college skill levels and use college GPA for post-college skill levels. I assume that ASVAB test scores and college GPA are noisy skill measures and use a dynamic factor model to deal with the measurement errors.

In exploiting post-education occupation choice, I take a task-based approach to relate occupations each other. One type of the task intensity constructed in this paper is general cognitive task intensity, which corresponds to general cognitive skill. The correspondence between the skill levels and task intensity choice in the labour market can be estimated using high school graduates, who are assumed to enter the labour market with the skill level measured by ASVAB test scores. Given pre-college level of general cognitive skill, a major who increases the skill more will take a job involving more intense general cognitive tasks. Hence, general cognitive skill growth during

[37] This type of skill, which is learned in college but is not rewarded in the labour market, is called an academic skill in some papers such as Kinsler and Pavan (2015) and Arcidiacono et al. (2016a).

college by college major is implied by the differences in choice of general cognitive task intensity from high school graduates.

The results show that all majors substantially increase students' general cognitive skill, but with large differences across majors. For students with a population average level of pre-college general cognitive skill, STEM majors increase general cognitive skill by 16 log wage points more than Humanities & Social Sciences majors. The effects of pre-college general cognitive skill levels on skill growth are somewhat different by college major. Skill growth increases with pre-college skill levels in STEM majors, while it decreases in Humanities & Social Sciences. Still, the effects are not large and people who choose another major would increase their skill more in STEM majors. Indirect effects of general cognitive skill growth on wages through task intensity choice are smaller than the direct effects. The direct effects are about three times larger than the indirect effects.

Wage returns to major-specific skills depend on whether the job is related to the corresponding major or not. The returns are zero in unrelated jobs for all majors, but they are positive in related jobs for Business & Economics and STEM majors. STEM majors show larger returns. However, the wage growth effects of major-specific skill growth are only about one quarter of those from general cognitive skill growth even for STEM majors whose potential post-college major-specific skill is 1.5 standard deviation above the population average.

These results suggest that growth in the general cognitive skill is the main contributor to the wage increase over high school graduation. The quantity of skill growth is different across majors, and that is a big factor in the observed large differences in wages across majors. In terms of productivity in the labour market, major-specific skills do not characterize college majors that much, and the main role of college majors is to produce workers with different levels of general cognitive skill.

My results also suggest that majoring in STEM will bring students a large monetary return in the labour market. I cannot identify from my paper, but its large skill increase might mean a large study cost in STEM majors. In fact, it is documented that STEM majors tend to study more hours than other majors despite their high pre-college cognitive test scores. Also, although many students switch out STEM majors to another major during college, only a small number of students switch into STEM majors from another major (see, e.g., Arcidiacono (2004) and Stinebrickner and Stinebrickner (2014)). This suggests that STEM majors are more difficult than other majors. Improving general cognitive skill and academic preparation in the pre-college stage may be an effective strategy to increase the number of STEM majors.

I show that college students increase a large amount of general cognitive skill, but what else they learn during college remains unclear. As shown above, the growth in major-specific skill does not contribute to wages for Humanities & Social Sciences majors although the acquired major-specific skill is an important factor of GPA. Why this is so remains an interesting question for future research. My paper examines only skill growth during college, but on-the-job skill accumulation might be different by major as well. This is another interesting topic for the future.

3.8 Bibliography

Aaronson, D., and B. Mazumder, 2011, The Impact of Rosenwald Schools on Black Achievement: Journal of Political Economy, **119**, 821–888.

Acemoglu, D., and D. Autor, 2011, Skills, Tasks and Technologies: Implications for Employment and Earnings, *in* Handbook of Labor Economics: Elsevier B.V., **4**, 1043–1171.

Ahn, T., P. Arcidiacono, A. Hopson, and J. Thomas, 2018, Grade Inflation in General Equilibrium with Implications for Female Interest in STEM Majors: mimeo.

Altonji, J. G., P. Bharadwaj, and F. Lange, 2012a, Changes in the Characteristics of American Youth: Implications for Adult Outcomes: Journal of Labor Economics, **30**, 783–828.

Altonji, J. G., E. Blom, and C. Meghir, 2012b, Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers: Annual Review of Economics, **4**, 185–223.

Arcidiacono, P., 2004, Ability Sorting and the Returns to College Major: Journal of Econometrics, **121**, 343–375.

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom, 2016a, College Attrition and the Dynamics of Information Revelation: NBER Working Paper.

Arcidiacono, P., E. M. Aucejo, and V. J. Hotz, 2016b, University Differences in the Graduation of Minorities in STEM Fields: Evidence from California: American Economic Review, **106**, 525–562.

Arcidiacono, P., J. Hotz, A. Maurel, and T. Romano, 2017, Ex Ante Returns and Occupational Choice: mimeo.

ASVAB Career Exploration Program, 2011, The ASVAB Career Exploration Program: Theoretical and Technical Underpinnings of the Revised Skill Composites and OCCU-Find: Technical report.

Aucejo, E. M., and J. James, 2016, The Path to College Education : Are Verbal Skills More Important than Math Skills?: mimeo.

Brint, S., A. M. Cantwell, and P. Saxena, 2012, Disciplinary Categories, Majors, and Undergraduate Academic Experiences: Rethinking Bok's "Underachieving Colleges" Thesis: Research in Higher Education, **53**, 1–25.

Carneiro, P., K. Hansen, and J. Heckman, 2003, Estimating Distribution of Counterfactuals with an Application to the Returns to Schooling and Measurement of the Effect of Uncertainty on Schooling Choice: International Economic Review, **44**, 361–422.

Carnevale, A. P., B. Cheah, and A. R. Hanson, 2015, The Economic Value of College Majors: Gerogetown University.

Cascio, E. U., and E. G. Lewis, 2006, Schooling and the Armed Forces Qualifying Test Evidence from School-Entry Laws: Journal of Human Resources, **41**, 294–318.

College Board, 2015, Total Group Profile Report: Technical report.

Deming, D. J., 2017, The Growing Importance of Social Skills in the Labor Market: Quarterly Journal of Economics.

Deming, D. J., and K. L. Noray, 2018, STEM Careers and Techonological Change: NBER Working Paper.

Fitzpatrick, M. D., D. Grissmer, and S. Hastedt, 2011, What a Difference a Day Makes: Estimating Daily Learning Gains during Kindergarten and First Grade Using a Natural Experiment: Economics of Education Review, **30**, 269–279.

Freeman, J. A., and B. T. Hirsch, 2008, College Majors and the Knowledge Content of Jobs: Economics of Education Review, **27**, 517–535.

Fruehwirth, J. C., S. Navarro, and Y. Takahashi, 2016, How the Timing of Grade Retention Affects Outcomes : Identification and Estimation of Time-Varying Treatment Effects: Jornal of Labor Economics, **34**, 979–1021.

Gormley, W. T., and T. Gayer, 2005, Promoting School Readiness in Oklahoma: An Evaluation of Tulsa's Pre-K Program: Journal of Human Resources, **40**, 533–558.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer, 2016, Multidimensional Skill Mismatch: mimeo.

Hamermesh, D. S., and S. G. Donald, 2008, The Effect of College Curriculum on Earnings: An Affinity Identifier for Non-Ignorable Non-Response Bias: Journal of Econometrics, **144**, 479–491.

Hansen, K. T., J. J. Heckman, and K. J. Mullen, 2004, The Effect of Schooling and Ability on Achievement Test Scores: Journal of Econometrics, **121**, 39–98.

Heckman, J., J. E. Humphries, and G. Veramendi, 2016a, Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking: NBER Working Paper.

Heckman, J. J., J. E. Humphries, and G. Veramendi, 2016b, Dynamic Treatment Effects: Journal of Econometrics, **191**, 276–292.

Kinsler, J., and R. Pavan, 2015, The Specificity of General Human Capital: Evidence from College Major Choice: Journal of Labor Economics, **33**, 933–972.

Kirkeboen, L. J., E. Leuven, and M. Mogstad, 2016, Field of Study, Earnings, and Self-Selection: Quarterly Journal of Economics, **131**, 1057–1111.

Lemieux, T., 2014, Occupations, Fields of Study and Returns to Education: Canadian Journal of Economics, **47**, 1047–1077.

Nordin, M., I. Persson, and D. O. Rooth, 2010, Education-Occupation Mismatch: Is There an Income Penalty?: Economics of Education Review, **29**, 1047–1059.

Poletaev, M., and C. Robinson, 2008, Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000: Journal of Labor Economics, **26**, 387–420.

Prada, M. F., and S. S. Urzúa, 2017, One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance and Wages: Journal of Labor Economics, **55**, 953–991.

Ransom, M. R., 2014, The Changing Occupational Distribution by College Major: mimeo.

Robst, J., 2007, Education and Job Match: The Relatedness of College Major and Work: Economics of Education Review, **26**, 397–407.

Speer, J. D., 2017a, Pre-Market Skills, Occupational Choice, and Career Progression: Journal of Human Resources, **52**, 187–246.

Stinebrickner, R., and T. R. Stinebrickner, 2014, A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout: Review of Economic Studies, **81**, 426–472.

Wiswall, M., and B. Zafar, 2015, Determinants of College Major Choice: Identification Using an Information Experiment: Review of Economic Studies, **82**, 791–824.

Yamaguchi, S., 2012, Tasks and Heterogeneous Human Capital: Journal of Labor Economics, **30**, 1–53.

Table 3.4: Log wage equation estimates

τ^c	(π_{21})	0.0991			
		(0.0088)	Business&Econ		
$(\tau^c)^2$	(π_{22})	0.0388	s_1^c	(π_{1E})	0.0148
		(0.0064)			(0.0441)
High school			Unrelated		
Const	$(\tilde{\pi}_0)$	1.9211	Const	$(\dot{\pi}_{0E,nr})$	2.3333
		(0.0937)			(0.1127)
s_1^c	(π_1)	0.0396	s_2^E	$(\pi_{3,nr}^E)$	-0.0178
		(0.0125)			(0.0329)
Some college			Related		
Const	$(\tilde{\pi}_{0Some})$	2.1844	Const	$(\dot{\pi}_{0E,r})$	2.2911
		(0.0870)			(0.1214)
s_1^c	(π_{1Some})	0.0411	s_2^E	$(\pi_{3,r}^E)$	0.0369
		(0.0271)			(0.0353)
Humanities&SS			STEM		
s_1^c	(π_{1H})	0.0025	s_1^c	(π_{1S})	0.0674
		(0.0381)			(0.0661)
Unrelated			Unrelated		
Const	$(\dot{\pi}_{0H,nr})$	2.2768	Const	$(\dot{\pi}_{0S,nr})$	2.3343
		(0.0961)			(0.2039)
s_2^H	$(\pi_{3,nr}^H)$	-0.0019	s_2^S	$(\pi_{3,nr}^S)$	-0.0186
		(0.0269)			(0.0690)
Related			Related		
Const	$(\dot{\pi}_{0H,r})$	2.1494	Const	$(\dot{\pi}_{0S,r})$	2.2737
		(0.1024)			(0.2045)
s_2^H	$(\pi_{3,r}^H)$	-0.0099	s_2^S	$(\pi_{3,r}^S)$	0.1093
		(0.0263)			(0.0682)

Notes: Equation (3.26) for high school, equation (3.34) for some college, and equation (3.35) for major $m = H, B, S$.

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Table 3.5: Average treatment effects; general cognitive skill growth

	Treatment			
	Some college	Humanities & SS	Business & Econ	STEM
High school	0.1138	0.2541	0.3119	0.3925
Some college	0.1136	0.2419	0.3122	0.3995
Humanities & SS	0.1132	0.2189	0.3126	0.4126
Business & Econ	0.1132	0.2189	0.3126	0.4126
STEM	0.1130	0.2083	0.3129	0.4187

Table 3.6: Pre-college levels of major-specific skills

	Humanities&SS	Business&Econ	STEM
s_1^m	-6.94	-7.50	-5.02

Note: $s_1^m = \underline{GPA}_1^m - \hat{\gamma}_{01}^m$, where $\underline{GPA}_1^m$ corresponds to zero GPA in major m in the 0-4.0 scale and $\hat{\gamma}_{01}^m$ is the estimate of γ_{01}^m.

Table 3.7: Major-specific skill growth at average; related jobs

	Humanities&SS	Business&Econ	STEM
$\pi_{3r}^m(0 - s_1^m)$	0	0.2762	0.5488
$\pi_{3r}^m(0 - s_{2m,10}^m)$	0	0.0666	0.0295

Table 3.8: Average treatment effects; major-specific skill growth

	Treatment		
	Humanities & SS	Business & Econ	STEM
Humanities & SS	0	0.0025	0.0091
Business & Econ	0	0.0008	0.0084
STEM	0	0.0024	0.0339

Chapter 4

Social Skill and Management

4.1 Introduction

As college workers accumulate experience in the labour market, some of them become managers. This paper examines what skills are most closely associated with male college workers becoming managers from the perspective of human capital theory.[1] Economists have traditionally paid a lot of attention to cognitive skill. Managers are located at the top of a hierarchy of workers and typically have a high level of cognitive skill. Nevertheless, managers have many duties and talents, which include leading a team of workers, communicating to their subordinates, and assigning tasks to them (Lazear (2012)). Hence, a high level of cognitive skill is unlikely to be the only requirement for becoming a manager.

This paper considers social skill in addition to cognitive skill. I define social skill as the capacity to work with people to achieve goals. The literature on leadership argues that social skill is important for management, and this skill has also been called interpersonal skill, human skill, or communication skill in different papers in the literature.[2] In organizational behaviour, Katz's three-skill approach considers social skill to be one of the basic skills for effective administration (see Katz (1974)).[3] Furthermore, emotional and social intelligence are identified as key determinants of leadership effectiveness in the literature because leadership is embedded in a social context (see, e.g., Kerr et al. (2006)). In the Economics literature, Lazear (2012) states that social skill is

[1] I exclude females because I want to focus on skill effects on workers' transitions in the labour market. I do not consider labour market participation decision.

[2] To be exact, leaders and managers can be differentiated (see, e.g., House and Aditya (1997)). Still, these two are largely overlapping.

[3] The other two skills are technical skills and conceptual skills. Technical skills are defined as knowledge about and proficiency in a specific type of work or activity, and conceptual skills are defined as the ability to work with ideas and concepts. Katz argues that technical skills are the most important at lower and middle levels of management and less important in upper management while the opposite is true for conceptual skills. He thinks that social skills, or, human skills in his paper, are important at any level of management.

likely to be important for leaders to convince others that they have leadership ability. Kuhn and Weinberger (2005) show that people who held leadership positions in high school are more likely to become managers and that their leadership variable captures some sort of social skill. McCann et al. (2016) develop a multi-sector matching model, in which individuals are heterogeneous in cognitive and social skill levels. Those with a high level of social skill become managers in teams, who specialize in social tasks, rather than workers in teams, who specialize in cognitive tasks.

Social skill seems to be important for managers, but there are not many empirical studies on this connection from the perspective of human capital theory. Therefore, this paper examines transition patterns into management jobs from the perspective of the theory and provides new support for the argument that social skill is important for becoming a manager. First, I demonstrate the link between management jobs and social tasks. The literature states that social tasks are an important factor of management jobs. Moreover, person-oriented behaviour is identified as one of the two broad types of behaviour in the leadership literature analyzing leader's actual behaviour (see, e.g., House and Aditya (1997)).[4] However, no paper actually constructs a data-based social task measure and empirically supports the statement that social tasks are important in management jobs. Therefore, the first part of my analysis constructs a social task intensity measure using the American Community Survey (ACS) Public Use Microdata Sample (PUMS) 2010-2017 and the O*NET and empirically shows that management jobs tend to have a high level of social task intensity as well as a high level of cognitive task intensity.

Then, using the constructed task intensity measures and panel data from the National Longitudinal Survey of Youth 1997 (NLSY97), I investigate transition patterns into management jobs. The NLSY97 contains cognitive test scores obtained in the pre-college period and self-assessed scales of personality traits. I examine what kinds of early career job characteristics are associated with becoming a manager and who tends to start in jobs with these characteristics. I demonstrate that most workers start their career in non-management jobs. Non-management jobs tend to have a lower level of social task intensity than management jobs, but there is considerable variation in social task intensity in non-management jobs. Workers who become managers by age 30 tend to start at a job with a high level of social task intensity. Conversely, cognitive task intensity in the first job does not increase the probability to become a manager. These findings suggest that increasing social skill in non-management but high social task jobs may be essential for becoming a manager. The workers in jobs with a high level of social task intensity at an early stage of their careers tend to be those with a high score in the cognitive test and with high social personality traits.

This paper also looks at the effect of college majors on workers' transition patterns into management jobs. Business majors are more likely to become managers, which may stem from social skill that is acquired during college. Furthermore, since skill growth during college substantially

[4] Another type is task-oriented behaviour.

depends on college major (see, e.g., Onozuka (2019)), the patterns of becoming a manager are likely to differ by college major. I aggregate college majors into three types: Science, Technology, Engineering, and Mathematics (STEM), Business & Economics, and Humanities & Social Sciences majors. I show that even after conditioning on social and cognitive task intensities in the first job, Business & Economics majors are more likely to be managers at age 30 than the other majors. Conversely, STEM majors, who are likely to have very high levels of cognitive skill compared to social skill at the college graduation point, tend to hold low social task intensity jobs in the early stage of their careers even after conditioning on their cognitive test score and social personality traits.

Management or leadership has been extensively studied, and there are many studies other than those listed above. In particular, there is a large literature dealing with management on the subject in organizational behaviour. Many theories have been proposed in that field, and summaries are provided by, for example, House and Aditya (1997); Lazear (2012), and Northouse (2018). One of the traditional approaches in this literature is to search for individual characteristics or traits associated with leadership positions. Many previous papers show that extraversion is positively correlated with leadership emergence and with leadership effectiveness (Judge et al. (2002)). According to Barrick and Mount (1991), extraversion increases job performance, particularly in management and sales jobs. Extraversion is considered to capture some sociability. My paper uses extraversion scales to construct a measure of social personality traits. Thus, my paper provides a possible explanation for these positive effects of extraversion on leadership emergence and on job performance in management jobs.

There are also papers on managers or leaders from an Economics perspective. For example, Lazear (2012) proposes a theory that implies that leaders are likely to be generalists, to have higher capabilities, and to place themselves in visible decision-making situations more frequently, and he tests the theory using data on Stanford alumni. Tong et al. (2019) examine the effects of various types of cognitive and non-cognitive skills, which do not include social skill, on an individual's potential of being a leader and find the strongest effect in problem-solving skill and in perseverance. There are also papers that show the effects of measures of youth sociability on adult job choice (see Borghans et al. (2008b) and Borghans et al. (2014)).

This paper is also related to studies on non-cognitive skills. Non-cognitive skills or personality traits, which are traditionally studied in Psychology, have more recently been studied by economists, and a growing number of papers examine the effects of non-cognitive skills on various labour market outcomes (see, e.g., Borghans et al. (2008a)). Gensowski (2018) estimates the effects of personality traits and IQ on lifetime earnings using a sample of high-IQ individuals. She finds large positive effects of conscientiousness and extraversion and large negative effects of agreeableness. She additionally finds that these effects are largest between ages 40 and 60 and for highly educated men. Social skill is also attracting attention these days in Economics, but many papers place particularly emphasis on the increasing complementarity between cognitive and social skills in the context of Skill-Biased Technological Change (see, e.g., Weinberger (2014) and Deming (2017)).

4.2 ACS and O*NET analysis

In this section, I provide descriptive evidence that links social task intensity and management jobs. As mentioned in the introduction, the literature supposes that management jobs are likely to involve intense social tasks. As far as I know, however, no paper actually shows empirical evidence by constructing a measure of social task intensity associated with a worker's job. Hence, I use data from the O*NET and construct a social task intensity measure. I show that management jobs have a high level of social task intensity compared to non-management jobs.

4.2.1 Data

4.2.1.1 O*NET

I use the O*NET to construct cognitive and social task intensity in this paper.[5] The O*NET is sponsored by the US Department of Labor/Employment and Training Administration and started as a successor to the Dictionary of Occupational Titles (DOT). Both the O*NET and the DOT characterize occupations using standardized measures and have been used in many papers taking a task-based approach (see, e.g., Poletaev and Robinson (2008) and Guvenen et al. (2016)). The O*NET contains a number of standardized measures describing the day-to-day aspects of the job, the qualifications and the interests of the typical workers in the occupations. I combine occupation data from the O*NET with individual data from the ACS matching on job codes. The ACS is described in the next subsection.

I select O*NET measures for cognitive and social task intensity, listed in Table 4.1. The selection for cognitive task intensity is following Onozuka (2019),[6] and the O*NET measures for social task intensity are selected following Deming (2017). Following previous papers taking a task-based approach, I employ a Principal Component Analysis (PCA) to the selected O*NET measures for cognitive and social task intensity, separately. For the PCA, I use the job information from the ACS PUMS 2010-2017, for US-born full-time full-year male workers aged between 25 and 54, who have a Bachelor's or higher degree. I take the first component after the PCA as cog-

[5]This paper uses O*NET 16.0, which was released in July 2011.

[6]The selection is mainly based on a technical report by the ASVAB Career Exploration Program (ASVAB Career Exploration Program (2011)).

nitive and social task intensity, respectively. Each type of task intensity measure is normalized to have a mean of 0 and a standard deviation of 1 over the selected sample from the ACS.[7]

Table 4.1: O*NET measures used to construct task intensity

	Cognitive	
Oral Comprehension	Oral Expression	Written Comprehension
Written Expression	English Language	Reading Comprehension
Speaking	Writing	Mathematical Reasoning
Number Facility	Mathematics	Mathematics Skill
Deductive Reasoning	Inductive Reasoning	Analyzing Data or Information
	Social	
Social Perceptiveness	Coordination	Persuasion
Negotiation		

4.2.1.2 ACS

I use the ACS PUMS 2010-2017. The ACS is a cross-sectional survey, which started in 2005. It is a replacement for the decennial census long form and asks various questions about individuals and households. The PUMS is a sample of responses to the ACS, and PUMS files are released every year. For an individual year, PUMS files contain data on approximately 1% of the US population.

The survey asks respondents their undergraduate major field of study. I divide the majors into three types: STEM, Business & Economics, and Humanities & Social Sciences. The definition of STEM majors follows Carnevale et al. (2015).[8] Although data on undergraduate major field of study are available from 2009 onward, I do not use 2009 because of a large change in occupation classification that occurred in 2010.

I restrict my analysis sample to US-born full-time full-year male workers, whose highest degree is a Bachelor's degree or higher.[9] After these restrictions are imposed, the total number of observations is 953,349. I use the provided person's weighting in the following analysis.

[7]There are other ways to construct task intensity measures. For example, Deming (2017) defines social task intensity as the average of the four selected O*NET measurements. As with Deming (2017), I could follow Autor et al. (2003) and convert average scores to a scale ranging from 0 to 10 that reflects a weighted percentile rank in the ACS sample. The results presented below are essentially the same as those obtained by using this approach.

[8]STEM majors contain architecture, engineering, biology, life sciences, computers, statistics, mathematics, and physical sciences. Business & Economics majors contain business and Economics. Humanities & Social Sciences majors contain arts, communications and journalism, education, humanities and liberal arts, industrial arts, consumer services, and recreation, law and public policy, psychology and social work, and social sciences except for Economics.

[9]Full-time full-year workers are defined here as those who work at least 30 hours per week and who work at least 40 weeks in a year.

4.2.2 Analysis

Table 4.2 reports the top ten occupations that have the highest social task intensity. Seven out of the ten occupations are management jobs.[10] This suggests that management jobs tend to have high social task intensity. Furthermore, given that previous papers, such as Yamaguchi (2012), show a positive relationship between worker's skill levels and their choice of task intensity, the table suggests that managers tend to have a high level of social skill. Column 4 in the table indicates that management jobs also have a higher level of cognitive task intensity than the average jobs held by college workers. As shown next, the management jobs that college workers tend to hold have a higher level of cognitive task intensity than non-management jobs.

Table 4.2: Top ten most social task intense occupations

Rank	Occupation	Social task intensity	Cognitive task intensity
1	Chief Executives	2.73	1.55
2	Clergy	2.47	0.67
3	Purchasing Managers	1.77	0.50
4	Public Relations and Fundraising Managers	1.65	0.73
5	Social and Community Service Managers	1.62	0.73
6	Child Family and School Social Workers	1.49	0.08
7	Sales Engineers	1.44	1.35
8	Marketing Managers	1.42	0.45
9	Human Resources Managers	1.39	0.57
10	Construction Managers	1.33	0.35

To present the differences in task intensities between management and non-management jobs in more detail, Tables 4.3 to 4.5 report the average task intensities of non-management and management jobs by 5-year age group and by college major. Management jobs tend to have a higher level of both cognitive and social task intensity than non-management jobs across all age groups and college majors. The differences in social task intensity between management and non-management jobs are particularly large in any college major. This supports the idea that a high level of social task intensity is a key feature of management jobs.

Workers' transition patterns into management jobs are examined in the next chapter using data from panel survey the NLSY97, but Tables 4.3 to 4.5 suggest some tentative patterns. Column 2 reveals that the fraction of workers in management jobs substantially increases with age regardless of college major. Columns 3 to 6 show that cognitive and social task intensities increase with age within management and within non-management jobs. This is consistent with the idea that college workers increase their cognitive and social skills, move up to jobs involving more intense cognitive

[10] Management jobs are the major group of Management Occupations in the occupation classification codes.

and social tasks, and become managers if their skill levels become high enough for management positions.

Table 4.3: Fraction of workers in management jobs, and mean task intensity of management and non-management jobs across 5-year age groups; STEM

Age	Fraction of managers	Non-management		Management	
		Cognitive	Social	Cognitive	Social
25-29	0.0945	0.2394	-0.4792	0.2681	0.7713
30-34	0.1479	0.2903	-0.3765	0.3200	0.7941
35-39	0.2051	0.2764	-0.3334	0.3542	0.8219
40-44	0.2423	0.2645	-0.3255	0.3856	0.8773
45-49	0.2649	0.3019	-0.3034	0.4077	0.8917
50-54	0.2714	0.3735	-0.2883	0.4242	0.9114

Table 4.4: Fraction of workers in management jobs, and mean task intensity of management and non-management jobs across 5-year age groups; Business & Economics

Age	Fraction of managers	Non-management		Management	
		Cognitive	Social	Cognitive	Social
25-29	0.1823	-0.2592	-0.3815	0.2887	0.8272
30-34	0.2498	-0.2192	-0.2886	0.3533	0.8671
35-39	0.2911	-0.2152	-0.2554	0.3933	0.9187
40-44	0.3119	-0.2171	-0.2259	0.4441	0.9968
45-49	0.3297	-0.2129	-0.2078	0.4692	1.0440
50-54	0.3334	-0.2111	-0.2228	0.4959	1.0921

Some important differences are observed across majors. The fraction of managers in the youngest age group (25-29) is the highest in Business & Economics majors and is the lowest in STEM majors. Business majors are more likely to learn leadership techniques or coordination of people and might, therefore, have received some training during college on how to be a manager. Given that STEM majors have a very high level of cognitive skill at the end of college (Onozuka (2019)), it is plausible that they focus on using their technical cognitive skill in non-management positions before possibly moving into management. As seen in columns 3 and 5 in Table 4.3, the average level of cognitive task intensity in STEM majors' non-management jobs is similar to the average level of cognitive task intensity in their management jobs. They may, therefore, be able to move from non-management jobs to management jobs relatively easily because they already meet the high cognitive task level. The fraction of managers increases the fastest for STEM majors with age.

Table 4.5: Fraction of workers in management jobs, and mean task intensity of management and non-management jobs across 5-year age groups; Humanities & Social Sciences

Age	Fraction of managers	Non-management		Management	
		Cognitive	Social	Cognitive	Social
25-29	0.1222	-0.5831	-0.3898	0.2633	0.8835
30-34	0.1563	-0.3721	-0.2321	0.3140	0.9103
35-39	0.1930	-0.2899	-0.1667	0.3654	0.9394
40-44	0.2129	-0.2724	-0.1484	0.3977	0.9690
45-49	0.2205	-0.2567	-0.1355	0.4166	1.0082
50-54	0.2235	-0.2314	-0.1135	0.4217	1.0217

4.3 NLSY97 analysis

So far, I have investigated the link between management jobs and social tasks using the ACS. Although the ACS has an advantage in its large sample size, it is a cross-sectional survey, and transition patterns cannot be studied with data from the ACS. To examine transition patterns, I combine individual data from the NLSY97 with the task intensity measures constructed above using the ACS. I examine what job characteristics are associated with a transition into management jobs and who tends to hold jobs with these characteristics. I also look at the effects of college majors on the transition.

4.3.1 Data

The NLSY97 is a panel survey conducted by the US Bureau of Labor Statistics. It started in 1997, was conducted annually until 2011, and has been conducted biennially since then. This section uses data from 1997 to 2015. In the first round, 8,984 males and females, who were between 12 and 17 years old at that time, were interviewed. I restrict the sample to males who have a Bachelor's degree or higher and whose bachelor's degree field falls into STEM, Business & Economics, or Humanities & Social Sciences.[11]

To analyze transition patterns in the labour market, I mainly look at job type at two particular points of the worker's career: their first job and the job held at age 30.[12] First job is defined as the worker's first full-time job after obtaining a Bachelor's degree.[13] This is also restricted to jobs

[11]The definition of college majors correspond to that using the ACS. STEM majors include biological sciences, architecture/environmental design, computer/information science, engineering, mathematics, and physical sciences. Business & Economics majors include business management, Economics, and hotel/hospitality management. Humanities & Social Sciences majors include anthropology, archaeology, area studies, communications, criminology, education, English, ethnic studies, fine and applied arts, foreign languages, history, home economics, interdisciplinary studies, philosophy, political science and government, psychology, sociology, and theology/religious studies.

[12]Age 30 is selected because the youngest cohort was 30 years old in 2015.

[13]Full-time workers here are defined as individuals who worked at least 30 hours per week.

taken before age 27.[14] Job at age 30 is defined as a full-time job held at age 30. Due to the biennial nature of the sample, some individuals did not take a survey at age 30. For these individuals, I use job information at age 29 or 31 as a proxy for the job at age 30.[15]

Most of the NLSY97 respondents took an Armed Services Vocational Aptitude Battery (ASVAB) test during the first round. The ASVAB test is composed of many parts, five of which are used in this section: Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, Mathematics Knowledge, and Numerical Operation. Since there may be an age effect, I adjust the test scores for age using the method of Altonji et al. (2012).[16] I use the average of the five test scores as a cognitive skill measure.[17] I standardize the score to have a mean of 0 and a standard deviation of 1 over the males who have a Bachelor's degree or higher in the NLSY97.

The NLSY97 asked the respondents to assess their personality traits in 2008, when the respondents were between 23 and 28 years old. Following Deming (2017), I construct a measure of social personality traits by taking the average of a scale of Extraversion and the opposite of a scale of Quietness.[18] I consider the constructed measure to represent personality traits that are stable with age but can help to increase social skill. According to the psychological literature on the Big Five, extraversion exhibits a flat trend from emerging adulthood through middle age (see, e.g., Soto et al. (2011)). I standardize the social personality traits measure to have a mean of 0 and a standard deviation of 1 over the males who have a Bachelor's degree or higher in the NLSY97.

Table 4.6 reports the means and the standard deviations of the cognitive test scores and the social personality traits measure by college major. The cognitive test score is the highest in STEM majors and the lowest in Humanities & Social Sciences majors. This ranking of majors is the same as that in cognitive task intensity shown in the ACS analysis and in post-college cognitive skill level estimated in Onozuka (2019). Since the ASVAB test was taken in the pre-college period, post-college cognitive skill level will be higher. According to Onozuka (2019), the differences across majors in cognitive skill level are larger at graduation than in the pre-college period.

For social personality traits, STEM majors show a much lower level than the other majors. This is consistent with the lowest task intensity and the smallest fraction of managers among college majors in their late 20s that are observed using the ACS. Business & Economics majors show a

[14]In my analysis sample, 89% of workers obtain their first job within a year after college graduation. Adding age dummies in the below analysis does not change the results.

[15]In order to keep the sample size as large as possible, I also use job information at age 29 or 31 as a proxy for those who did not answer the survey at age 30. If job information is available at both ages 29 and 31, I use the one at age 29.

[16]In this method, ASVAB score at the qth percentile in the distribution of the taker's age is assigned to the ASVAB score at the qth percentile in the age 16 distribution. This method implicitly assumes that age effects do not change the rank in ASVAB scores although age effects can be nonconstant.

[17]I standardize these five test scores over the males who have a Bachelor's degree or higher in the NLSY97. I take the average and then re-standardize them.

[18]Like I did with the cognitive test score, I standardize the two scales over the males who have a Bachelor's degree or higher in the NLSY97. I take the average and then re-standardize them.

Table 4.6: Means and standard deviations of cognitive test scores and social personality traits

	STEM	Business & Economics	Humanities & Social Sciences
Cognitive test score	0.2320	0.1458	0.0580
	(0.9927)	(0.8409)	(0.9908)
Social personality traits	-0.1450	0.0188	0.1263
	(0.8842)	(0.8252)	(0.8500)
N	149	151	182

lower level of social personality traits than Humanities & Social Sciences majors, although the above ACS results show that the average social task intensity and the fraction of managers are larger in Business & Economics majors in their late 20s than in Humanities & Social Sciences majors in their late 20s. Since Business & Economics majors may increase their social skill as part of managerial training during college, they may be able to take a job involving intense social tasks after college despite their lower innate social personality traits.

4.3.2 Analysis

In order to examine the transition between non-management and management jobs with age, Table 4.7 reports frequencies by management/non-management jobs in the first jobs and jobs held at age 30.[19] Over 92% (445/482) of workers start their careers in non-management jobs; less than 8% (37/482) start in management jobs. The fraction of managers increases to 18% (89/482) at age 30. This increase stems from a large fraction of those who started their careers as non-managers, becoming managers as they accumulate experience.[20] In addition to the results using the total individuals, the following analysis shows the results using only those who started their careers in non-management jobs to focus on the transition from non-management to management jobs.

Table 4.7: Cross tabulation of managers/non-managers in first jobs and jobs held at age 30

		Age 30		
		Non-management	Management	Total
First job	Non-management	375	70	445
	Management	18	19	37
	Total	393	89	482

What college majors and what job characteristics are associated with the transition to management? To answer this question, I run a set of probit regressions, reported in Table 4.8.[21] Columns

[19] Management jobs are the major group of Management Occupations in the occupation classification codes.
[20] The results are similar when separating by college major.
[21] Coefficient estimates are reported in Table C.1 in the appendix.

1 and 3 show that Business & Economics majors are more likely to be managers at age 30 than Humanities & Social Sciences majors.

Table 4.8: Probability of being a manager at age 30: average marginal effects from Probit regressions

	All		First job=Non management	
Business & Economics	0.1047**	0.1227***	0.1057**	0.1200***
	(0.0475)	(0.0484)	(0.0472)	(0.0487)
STEM	0.0359	0.0709	0.0138	0.0361
	(0.0466)	(0.0514)	(0.0443)	(0.0494)
Social task intensity in first job		0.0949***		0.0510**
		(0.0228)		(0.0237)
Cognitive task intensity in first job		-0.0385*		-0.0224
		(0.0219)		(0.0207)
N	482	482	445	445

Dependent variable is a dummy of being a manager at age 30.

Standard errors are in parentheses.

* p<0.1, ** p<0.05, *** p<0.01.

Estimates of a constant and year dummies are omitted.

As shown above, management jobs tend to be high in social and cognitive task intensity. Most college workers are not in management jobs at the beginning of their careers. Nevertheless, there is considerable variation in both social and cognitive task intensity in these non-management jobs. If workers increase their skills more in high task intensity jobs as in Yamaguchi (2012), individuals who start in high task intensity jobs will be more likely to be managers at age 30. Columns 2 and 4 show that this is true for social task intensity. Even though the coefficient of social task intensity in the first job decreases when the sample is restricted to individuals whose first job is in non-management, it is still positive and significant at the 5% level. This descriptive analysis does not exclude the possibility that workers with high levels of skills for management are likely to hold jobs involving intense social tasks, but workers seem to gain some experience in non-management jobs to perform a high level of social tasks before shifting into management jobs.

Interestingly, the same story does not apply to cognitive task intensity. Accumulating experience in high cognitive task intensity jobs is not be very helpful for becoming a manager after conditioning on social task intensity. Working in high cognitive task intensity jobs may be helpful for increasing cognitive skill. However, according to Tables 4.3 to 4.5 in the previous section, college workers seem to possess much lower levels of social skill than that would be required in a management job when they first enter the labour market.[22] Hence, increasing social skill in high

[22]Moreover, according to Chronicle of Higher Education (2012) and Association of American College and Universities (2018), employers identify oral and written communication skills and knowledge as the most important skills when hiring, but they think that recent college graduates are not well-prepared in these skills.

social task intensity jobs may be essential for workers to transition into management, and being good at only cognitive tasks is unlikely to be sufficient.

Furthermore, inclusion of the task intensity leaves the significant effect of being a Business & Economics major largely unchanged. Even after controlling for social task intensity in their first jobs, they are more likely to be managers at age 30. This suggests that Business & Economics majors have acquired some skills other than social skill that are useful for management, during college.

Table 4.8 shows the importance of a high level of social task intensity in the first job for the possibility of becoming a manager. What determines social task intensity on the first job? To provide some answers to this, I regress social task intensity on a constant, college major dummies, a social personality traits measure, cognitive test scores, and year dummies. Columns 1, 2, 5, and 6 in Table 4.9 use only first jobs. To expand the sample size, columns 3, 4, 7, and 8 use jobs held before age 27.[23]

Workers with high social personality traits and high cognitive test scores tend to take jobs that have a high level of social task intensity. Individuals with high social personality traits might experience little stress in communicating and working with people. It is also possible that they have held leadership positions in the past and, thus, have a high level of social skill compared to those with low social personality traits (see, e.g., Kuhn and Weinberger (2005)). Cognitive test score has a positive effect on social task intensity choice. This may be because cognitive skill is required to understand other people's opinions or to construct arguments when coordinating or negotiating with people. In addition, the cognitive and social task intensity measures are positively correlated with each other. Some knowledge or skills on work will be required in coordinating or negotiating with people. The positive effect of cognitive test score also implies that not only social skill but also cognitive skill might increase in high social task intensity jobs, and a high social task intensity job might be enough to increase cognitive skill to the level required for management. This can be another reason why cognitive task intensity in the first job is insignificant in Table 4.8.

Overall, STEM majors tend to take jobs that have a low level of social task intensity compared to the other two types of majors, even when controlling for their social personality traits and cognitive test scores. As mentioned above, since they will have a very high level of cognitive skill at the college graduation point, as estimated in Onozuka (2019), they might want to focus on utilizing their cognitive skill in non-management jobs in their early career. Furthermore, if their STEM skills become obsolete relatively quickly as argued in Deming and Noray (2018), they might want to take full advantage of their technical skill in non-management jobs before the skill becomes less valuable.

One might think that Humanities & Social Sciences majors have a stronger preference for

[23] Age dummies are added in the regressions in this case.

Table 4.9: Social task intensity equation

	All				First job=Non management			
	First job		Age≤26		First job		Age≤26	
Business & Economics	-0.0437	-0.0397	-0.0779	-0.0739	-0.1183	-0.1195	-0.1227	-0.1251
	(0.1221)	(0.1224)	(0.1050)	(0.1037)	(0.1218)	(0.1221)	(0.1072)	(0.1057)
STEM	-0.0833	-0.0753	-0.2442**	-0.2372**	-0.1184	-0.1244	-0.3018***	-0.3072***
	(0.1228)	(0.1243)	(0.1036)	(0.1024)	(0.1215)	(0.1232)	(0.1022)	(0.1001)
Social personality traits		0.0638		0.1008**		0.0442		0.0922*
		(0.0601)		(0.0489)		(0.0591)		(0.0484)
Cognitive test score		0.0599		0.0971**		0.0798		0.1116**
		(0.0552)		(0.0461)		(0.0548)		(0.0461)
N	482	482	1531	1531	445	445	1411	1411

Dependent variable is social task intensity.

Standard errors are in parentheses. Standard errors in columns 5 to 8 are clustered at the individual level.

* p<0.1, ** p<0.05, *** p<0.01.

Age dummies are also included from columns 5 to 8.

Estimates of a constant, year dummies, and age dummies (columns 5 to 8 only) are omitted.

high social task intensity jobs than the other college majors. If this is the case, the insignificant coefficient of a Business & Economics major dummy will be consistent with that they have a higher level of social skill than the other college majors. However, Table 4.8 shows that inclusion of social task intensity in the first job does not affect the probability of Business & Economics majors being a manager at age 30. This suggests that their social skill levels will not be very high compared to those of other majors.

4.4 Conclusions and future work

Some college workers become managers as they accumulate experience in the labour market. The literature argues that management jobs involve intense social tasks, such as leading a team of workers or communicating to their subordinates, and that both cognitive and social skills are likely to be important skills for management jobs. However, there are not many papers that empirically study the relationship between social skill and management jobs from the perspective of human capital theory.

This paper examined male college workers' transition patterns into management with a focus on social and cognitive skills. The results suggest the importance of social skill. Management positions tend to involve intense social and cognitive tasks, and their very high level of social task intensity is a key feature of management jobs. Conversely, new college workers are likely to have a lower level of social and cognitive skills than the level that would be required for management. In particular, their social skill levels tend to be much lower. Given a large part of college education is performed through classroom lecture compared to group work, college students may not increase their social skill as much as their cognitive skill. My results are consistent with that increasing social skill is essential for college workers to becoming a manager. Although cognitive skill has been heavily emphasized in the Economics literature, being good only at cognitive tasks is unlikely to be enough for management. Furthermore, college teaching styles might need to be changed in order to foster social skill during college.

Business & Economics majors are more likely to become managers compared to the other types of college majors, which is consistent with the idea that they learn some skills that are useful for management during college. According to the results, however, their high probability of being a manager does not seem to stem mainly from their social skill. Further research on the reason why they are more likely to become managers is important to develop a better leadership curriculum.

The method in this paper is descriptive, and more detailed analysis which addresses endogenous choice of college major and occupation remains for future work. This paper cannot exclude the possibility that workers hold high social task intensity jobs because of their preference for social tasks or their high social skill level and that skill growth does not differ across jobs. It is also possible that the positive effect of being a Business & Economics major on the probability of being in a management job shown in this paper is simply because they have a strong preference for management jobs.

Due to the data limitation, I looked at only the early stage of a worker's career, but the fraction of managers significantly increases by age, particularly until the mid 40s. Transition patterns in the middle and later stages of a worker's career might be different from that in the early stage. A longer panel dataset will be helpful for understanding transition patterns over a worker's whole career. Furthermore, it will be interesting to examine the relationships between timing of becoming a manager and management effectiveness. If workers shift into management positions relatively early, they might not increase their technical cognitive skill much although they will accumulate management experience a lot in the management jobs. However, it is possible that top-level management jobs need a higher level of technical cognitive skill than lower levels of management jobs. If the skill becomes more difficult to increase as workers get older, becoming a manager earlier will not necessarily be better.

4.5 Bibliography

Altonji, J. G., P. Bharadwaj, and F. Lange, 2012, Changes in the Characteristics of American Youth: Implications for Adult Outcomes: Journal of Labor Economics, **30**, 783–828.

Association of American College and Universities, 2018, Fulfilling the American Dream: Liberal Education and the Future of Work: Technical Report July.

ASVAB Career Exploration Program, 2011, The ASVAB Career Exploration Program: Theoretical and Technical Underpinnings of the Revised Skill Composites and OCCU-Find: Technical report.

Autor, D. H., F. Levy, and R. J. Murnane, 2003, The Skill Content of Recent Technological Change: An Empirical Exploration: Quarterly Journal of Economics, 1279–1333.

Barrick, M. R., and M. K. Mount, 1991, The Big Five Personality Dimensions and Job Performance: A Meta-Analysis: Personnel Psychology, **44**, 1–26.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel, 2008a, The Economics and Psychology of Personality Traits: Journal of Human Resources, **43**, 972–1059.

Borghans, L., B. ter Weel, and B. A. Weinberg, 2008b, Interpersonal Styles and Labor Market Outcomes: Journal of Human Resources, **43**, 815–858.

Borghans, L., B. T. Weel, and B. A. Weinberg, 2014, People Skills and the Labor-Market Outcomes of Underrepresented Groups: ILR Review, **67**, 287–334.

Carnevale, A. P., B. Cheah, and A. R. Hanson, 2015, The Economic Value of College Majors: Georgetown University.

Chronicle of Higher Education, 2012, The Role of Higher Education in Career Development: Employer Perceptions: Technical report, Washington.

Deming, D. J., 2017, The Growing Importance of Social Skills in the Labor Market: Quarterly Journal of Economics.

Deming, D. J., and K. L. Noray, 2018, STEM Careers and Technological Change: NBER Working Paper.

Gensowski, M., 2018, Personality, IQ, and Lifetime Earnings: Labour Economics, **51**, 170–183.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer, 2016, Multidimensional Skill Mismatch: mimeo.

House, R. J., and R. N. Aditya, 1997, The Social Scientific Study of Leadership: Quo Vadis?: Journal of Management, **23**, 409–473.

Judge, T. A., J. E. Bono, R. Ilies, and M. W. Gerhardt, 2002, Personality and Leadership: A Qualitative and Quantitative Review: Journal of Applied Psychology, **87**, 765–780.

Katz, R., 1974, Skills of an Effective Administrator: Harvard Business Review, **52**, 90–102.

Kerr, R., J. Garvin, N. Heaton, and E. Boyle, 2006, Emotional Intelligence and Leadership Effectiveness: Leadership and Organization Development Journal, **27**, 265–279.

Kuhn, P., and C. Weinberger, 2005, Leadership Skills and Wages: Journal of Labor Economics, **23**, 395–436.

Lazear, E. P., 2012, Leadership: A Personnel Economics Approach: Labour Economics, **19**, 92–101.

McCann, R. J., X. Shi, A. Siow, and R. Wolthoff, 2016, Becker Meets Ricardo: Multisector Matching with Communication and Cognitive Skills: Journal of Law, Economics, and Organization, **31**, 690–720.

Northouse, P. G., 2018, Leadership: Theory and Practice, eighth ed.: SAGE Publications.

Onozuka, Y., 2019, Heterogeneous Skill Growth Across College Majors: mimeo.

Poletaev, M., and C. Robinson, 2008, Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys, 1984–2000: Journal of Labor Economics, **26**, 387–420.

Soto, C. J., O. P. John, S. D. Gosling, and J. Potter, 2011, Age Differences in Personality Traits From 10 to 65: Big Five Domains and Facets in a Large Cross-Sectional Sample: Journal of Personality and Social Psychology, **100**, 330–348.

Tong, T., H. Li, and S. Greiff, 2019, Human Capital and Leadership: The Impact of Cognitive and Noncognitive Abilities: Applied Economics, **51**, 5741–5752.

Weinberger, C. J., 2014, The Increasing Complementarity Between Cognitive and Social Skills: Review of Economics and Statstics, **96**, 849–861.

Yamaguchi, S., 2012, Tasks and Heterogeneous Human Capital: Journal of Labor Economics, **30**,

1–53.

Appendix A

Chapter 2 Appendix

Table A.1: Component loadings based on PCA after varimax rotation

	Component 1	Component 2	Unexplained
Speaking	**0.3198**	-0.1494	0.0999
Active listening	**0.2975**	-0.1116	0.1634
Writing	**0.2827**	-0.0457	0.0715
Written expression	**0.2729**	-0.0399	0.1203
Oral expression	**0.2681**	-0.0239	0.0972
Critical thinking	**0.2326**	0.0318	0.1258
Reading comprehension	**0.2238**	0.0595	0.0650
Written comprehension	**0.2186**	0.0643	0.0795
Oral comprehension	**0.2164**	0.0628	0.1013
Memorization	**0.2134**	0.0109	0.3320
Active learning	**0.2119**	0.0803	0.0505
Learning strategy	**0.2112**	-0.0300	0.4710
Monitoring	**0.1840**	0.0727	0.2714
Fluency of ideas	**0.1805**	0.0870	0.2305
Inductive reasoning	**0.1704**	0.1083	0.1952
Originality	**0.1616**	0.1064	0.2596
Deductive reasoning	**0.1550**	0.1513	0.0767
Problem sensitivity	0.1119	**0.1588**	0.3013
Category flexibility	0.0790	**0.2259**	0.1301
Information ordering	0.0681	**0.2459**	0.0754
Speed of closure	0.0585	**0.2253**	0.2438
Science	0.0495	**0.2181**	0.3260
Mathematical reasoning	0.0355	**0.2348**	0.3056
Number facility	0.0156	**0.2379**	0.3763
Mathematics	0.0119	**0.2569**	0.2945
Time sharing	-0.0048	**0.1002**	0.9073
Flexibility of closure	-0.0111	**0.2974**	0.1697
Selective attention	-0.0463	**0.2911**	0.3361
Perceptual speed	-0.0796	**0.3440**	0.1591
Visualization	-0.1100	**0.3561**	0.1772
Spatial orientation	-0.2276	**0.2019**	0.5876

Table A.2: Basic statistics for the other variables: males

	Closely related	Somewhat related	Not related
Years full-time experience	15.74	16.17	15.30
	(8.45)	(8.52)	(8.85)
Years part-time experience	1.23	1.26	1.37
	(3.32)	(3.37)	(3.48)
Disabled	0.06	0.06	0.07
	(0.23)	(0.24)	(0.25)
Hispanic	0.06	0.05	0.07
	(0.24)	(0.23)	(0.25)
White	0.76	0.79	0.74
	(0.43)	(0.41)	(0.44)
Black	0.06	0.06	0.08
	(0.24)	(0.24)	(0.28)
Asian	0.11	0.09	0.10
	(0.31)	(0.29)	(0.30)
Native	0.01	0.01	0.01
	(0.09)	(0.08)	(0.09)
Native born US citizen	0.81	0.85	0.84
	(0.39)	(0.36)	(0.36)
Naturalized US citizen	0.12	0.10	0.10
	(0.33)	(0.30)	(0.30)
Non-US citizen	0.07	0.05	0.06
	(0.25)	(0.21)	(0.23)
Never married	0.14	0.16	0.18
	(0.35)	(0.36)	(0.38)
Management training	0.27	0.31	0.23
	(0.44)	(0.46)	(0.42)
Technical training	0.59	0.56	0.47
	(0.49)	(0.50)	(0.50)
General professional training	0.23	0.23	0.18
	(0.42)	(0.42)	(0.38)
Other work-related training	0.10	0.09	0.08
	(0.30)	(0.28)	(0.27)

Standard deviations are in parentheses.

Table A.3: Basic statistics for the other variables: females

	Closely related	Somewhat related	Not related
Years full-time experience	13.58	12.58	12.17
	(7.73)	(7.38)	(7.72)
Years part-time experience	1.51	1.60	1.67
	(3.39)	(3.46)	(3.74)
Disabled	0.06	0.07	0.07
	(0.24)	(0.25)	(0.26)
Hispanic	0.08	0.07	0.07
	(0.27)	(0.26)	(0.26)
White	0.67	0.67	0.64
	(0.47)	(0.47)	(0.48)
Black	0.15	0.14	0.15
	(0.35)	(0.35)	(0.35)
Asian	0.10	0.11	0.13
	(0.30)	(0.31)	(0.33)
Native	0.01	0.01	0.01
	(0.10)	(0.10)	(0.10)
Native born US citizen	0.85	0.85	0.82
	(0.35)	(0.36)	(0.38)
Naturalized US citizen	0.10	0.11	0.12
	(0.30)	(0.31)	(0.32)
Non-US citizen	0.04	0.05	0.06
	(0.20)	(0.21)	(0.24)
Never married	0.21	0.26	0.25
	(0.41)	(0.44)	(0.43)
Management training	0.23	0.29	0.20
	(0.42)	(0.45)	(0.40)
Technical training	0.60	0.55	0.46
	(0.49)	(0.50)	(0.50)
General professional training	0.29	0.28	0.21
	(0.45)	(0.45)	(0.41)
Other work-related training	0.17	0.14	0.11
	(0.38)	(0.35)	(0.31)

Standard deviations are in parentheses.

Table A.4: Correspondence table for categories of fields of study

Major	Minor	Fraction
Computer & Math Sciences	Computer & information sciences	0.62
	Mathematical sciences	0.38
Life & Related Sciences	Biological sciences	0.77
	Agricultural & food sciences	0.15
	Environmental life sciences	0.08
Physical & Related Sciences	Chemistry, except biochemistry	0.44
	Earth science, geology & oceanography	0.24
	Physics & astronomy	0.25
	Physical & related sciences	0.03
	Other physical sciences	0.04
Social & Related Sciences	Psychology	0.42
	Sociology & related engineering	0.19
	Political & related sciences	0.15
	Economics	0.14
	Other social sciences	0.11
Engineering	Electrical & related engineering	0.32
	Mechanical engineering	0.20
	Civil & architectural engineering	0.18
	Chemical engineering	0.08
	Aerospace & related engineering	0.04
	Industrial engineering	0.04
	Other engineering	0.14
Non-S&E degrees	Management & administration	0.26
	Teaching - except S&E & postsecondary	0.22
	Health & related	0.15
	Art, humanities & related	0.12
	Social service & related	0.05
	Technology & technical	0.03
	Sales & marketing	0.03
	Other non-S&E	0.14

Appendix B

Chapter 3 Appendix

B.1 Identification of measurement system of general cognitive skill

I specify $\theta^{mech} = a\theta^c + \theta^{omech}$ and assume θ^c and θ^{mech} are orthogonal to each other. Using this specification and skill equations (3.1) and (3.2), the skill measurement system, equations (3.3) to (3.10), can be rewritten as follows:

$$\begin{aligned}
WordKnowledge_i &= x'_{si}\alpha^c + \theta_i^c + e_{1i}\\
ParagraphComprehension_i &= \delta_{12}(x'_{si}\alpha^c + \theta_i^c) + e_{2i}\\
ArithmeticReasoning_i &= \delta_{13}(x'_{si}\alpha^c + \theta_i^c) + e_{3i}\\
MathematicsKnowledge_i &= \delta_{14}(x'_{si}\alpha^c + \theta_i^c) + e_{4i}\\
NumericalOperation_i &= \delta_{15}(x'_{si}\alpha^c + \theta_i^c) + e_{5i}\\
MechanicalComprehension_i &= \delta_{16}(x'_{si}\alpha^c + \theta_i^c) + \delta_{26}(x'_{si}\alpha^{mech} + a\theta_i^c + \theta_i^{omech}) + e_{6i}\\
Auto\&ShopInformation_i &= x'_{si}\alpha^{mech} + a\theta_i^c + \theta_i^{omech} + e_{7i}\\
ElectronicsInformation_i &= \delta_{18}(x'_{si}\alpha^c + \theta_i^c) + \delta_{28}(x'_{si}\alpha^{mech} + a\theta_i^c + \theta_i^{omech}) + e_{8i}.
\end{aligned}$$

Since x_s is orthogonal with θ^c, θ^{omech}, and e's, the parameters on x_s, that is, α^c, $\delta_{1s}\alpha^c$, α^{mech}, and $\delta_{2s}\alpha^{mech}$, $s = 1, 2, \cdots, 8$, can be identified. Moving the identified terms into the left hand side gives

$$\begin{aligned}
\widetilde{WordKnowledge}_i &= \theta_i^c + e_{1i} && \text{(B.1)}\\
\widetilde{ParagraphComprehension}_i &= \delta_{12}\theta_i^c + e_{2i}\\
\widetilde{ArithmeticReasoning}_i &= \delta_{13}\theta_i^c + e_{3i}\\
\widetilde{MathematicsKnowledge}_i &= \delta_{14}\theta_i^c + e_{4i}\\
\widetilde{NumericalOperation}_i &= \delta_{15}\theta_i^c + e_{5i}\\
\widetilde{MechanicalComprehension}_i &= (\delta_{16} + a\delta_{26})\theta_i^c + \delta_{26}\theta_i^{omech} + e_{6i}\\
\widetilde{Auto\&ShopInformation}_i &= a\theta_i^c + \theta_i^{omech} + e_{7i} && \text{(B.2)}\\
\widetilde{ElectronicsInformation}_i &= (\delta_{18} + a\delta_{28})\theta_i^c + \delta_{28}\theta_i^{omech} + e_{8i}.
\end{aligned}$$

This is a triangular form explained in Carneiro et al. (2003). There are five measures affected only by θ^c and three measures affected by θ^c and θ^{omech}, and the factor loadings on θ^c in equation (B.1) and on θ^{omech} in equation (B.2) are normalized to 1. Therefore, the factor loadings on θ^c and on θ^{omech} and the distributions of θ^c, θ^{omech}, and e's can be identified. Since a, δ_{26}, δ_{28}, $\delta_{16} + a\delta_{26}$, and $\delta_{18} + a\delta_{28}$ are identified, δ_{16} and δ_{18} can be identified.

B.2 Derivation of linear task intensity choice

Individual i's problem in period 2 is

$$\max_{\tau} u(\tau^c, s_i^c, \varepsilon_{hi}^c),$$

where ε_h^c is a working cost shock. Utility in period 2 is modeled as

$$u(\tau_i^c, s_i^c, \varepsilon_{hi}^c) = \log w(\tau_i^c, s_i^c) - h(\tau_i^c, s_i^c, \varepsilon_{hi}^c),$$

where $h(\cdot)$ is a working cost function. For simplicity, suppose that log wage equation is

$$\log w_i = \pi_0 + \pi_1 s_{i2}^c + \pi_{21}\tau_i^c + \pi_{22}(\tau_i^c)^2.$$

Suppose working cost is

$$h(\tau_i^c, s_i^c, \varepsilon_{hi}) = h_1 s_{i2}^c + h_2\tau_i^c + h_3(\tau_i^c)^2 + h_4(s_i^c\tau_i^c) + \tau_i^c\varepsilon_{hi}^c$$

The FOC for τ^c is

$$\pi_{21} + 2\pi_{22}\tau_i^c = h_2 + 2h_3\tau_i^c + h_4 s_i^c + \varepsilon_{hi}^c.$$

Hence, the optimal math task intensity given s_2^c is

$$\tau_i^* = \frac{\pi_{21} - h_2 - h_4 s_i^c - \varepsilon_{hi}^c}{2(h_3 - \pi_{22})}.$$

which is linear in s^c.

B.3 Equations rewritten using the assumptions on pre-college major-specific skills

Job type choice equation (3.18) is given by

$$\begin{aligned} I_{m_-,i}^c &= \left(\xi_0^c + \sum_{m' \in \{H,E,S\}} \xi_{3m'}^c s_1^{m'}\right) + \xi_1^c s_i^c + \xi_2^c s_i^{mech} + x_{ci}'\beta_{m_-}^c + \varepsilon_{m_-,i}^c \\ &= \tilde{\xi}_0^c + \xi_1^c s_i^c + \xi_2^c s_i^{mech} + x_{ci}'\beta_{m_-}^c + \varepsilon_{m_-,i}^c. \end{aligned} \tag{B.3}$$

Education level choice equation (3.21) is rewritten as

$$\begin{aligned} I_{li} &= \left(\eta_{0l} + \sum_{m' \in \{H,E,S\}} \eta_{3,l,m'} s_1^{m'}\right) + \eta_{1l} s_{1i}^c + \eta_{2l} s_{1i}^{mech} + x_{di}'\beta_l + z_{dli}\varphi_{dl} + \varepsilon_{dli} \\ &= \tilde{\eta}_{0l} + \eta_{1l} s_{1i}^c + \eta_{2l} s_{1i}^{mech} + x_{di}'\beta_l + z_{dli}\varphi_{dl} + \varepsilon_{dli}. \end{aligned} \tag{B.4}$$

College major choice equation (3.22) is rewritten as

$$\begin{aligned} I_{mi} &= (\eta_{0m} + \eta_{3m} s_1^m) + \eta_{1m} s_{1i}^c + \eta_{2m} s_{2i}^m + x_{di}'\beta_m + z_{mi}'\varphi_m + \varepsilon_{mi} \\ &= \tilde{\eta}_{0m} + \eta_{1m} s_{1i}^c + \eta_{2m} s_{2i}^m + x_{di}'\beta_m + z_{mi}'\varphi_m + \varepsilon_{mi}. \end{aligned} \tag{B.5}$$

B.4 Equations rewritten in terms of pre-college general cognitive skill

I suppose $s^{mech}_{2,Some,i} = \lambda^{mech}_{0,Some} + \lambda^{mech}_{1Some} s^{mech}_{1i}$ similar to general cognitive skill change. The job type choice equation (B.3) for some college is rewritten as

$$\begin{aligned} I^c_{Some,i} &= (\xi^c_0 + \xi^c_1 \lambda^c_{0,Some} + \xi^c_2 \lambda^{mech}_{0,Some}) + \xi^c_1 \lambda^c_{1,Some} s^c_{1i} + \xi^c_2 \lambda^{mech}_{1,Some} s^{mech}_{1i} + x'_{ci} \beta^c_{Some} + \varepsilon^c_{Some,i}. \\ &= \tilde{\xi}^c_{0,Some} + \tilde{\xi}^c_1 s^c_{1i} + \tilde{\xi}^{mech}_2 s^{mech}_{1i} + x'_{ci} \beta^c_{Some} + \varepsilon^c_{Some,i}. \end{aligned} \tag{B.6}$$

Job relatedness choice equation (3.20) can be rewritten as

$$\begin{aligned} I_{rmi} &= (\xi_{0rm} + \xi_{1rm} \lambda^c_{0m}) + \xi_{1rm} \lambda^c_{1m} s^c_{1i} + \xi_{2rm} s^m_{2i} + x'_{wi} \beta_{rm} + \varepsilon_{rmi} \\ &= \tilde{\xi}_{0rm} + \tilde{\xi}_{1rm} s^c_{1i} + \xi_{2rm} s^m_{2i} + x'_{wi} \beta_{rm} + \varepsilon_{rmi}. \end{aligned} \tag{B.7}$$

B.5 Estimated equations

The first stage estimates the ASVAB equations (3.3) to (3.10), the education level choice equation (B.4), and the job type choice, equation (B.3) for high school and equation (B.6) for some college. The second stage estimates the GPA equations (3.32), (3.33), (3.14) and (3.15) and the college major choice (B.5). In the third stage, the general cognitive task intensity choice, equation (3.23) for high school and equation (3.24) for the other education groups, job relatedness choice (B.7), and log wage equations, equation (3.26) for high school, equation (3.34) for some college, and equation (3.35) for major $m = H, E, S$, are estimated.

B.6 Parameter estimates

Table B.1: Skill equation parameter estimates

Variables	Cognitive (α_c)		Mechanical (α_{mech})	
Constant	-0.8231	(0.0542)	-0.1389	(0.0526)
Female	0.0944	(0.0262)	-0.5163	(0.0232)
Hispanic/Mixed race	0.3493	(0.0412)	0.2782	(0.0321)
Non-Black/Non-Hispanic	0.5917	(0.0329)	0.6341	(0.0294)
North central 1997	-0.0553	(0.0400)	0.1063	(0.0367)
South 1997	-0.1203	(0.0375)	0.0003	(0.0310)
West 1997	-0.1864	(0.0469)	0.0010	(0.0378)
Urban 1997	0.0464	(0.0316)	-0.1983	(0.0264)
Broken home 1997	-0.0839	(0.0264)	-0.0185	(0.0251)
Father's education	0.0869	(0.0148)	0.0149	(0.0114)
Mother's education	0.1229	(0.0170)	0.0394	(0.0133)
Household income (Quartile groups)	0.1091	(0.0116)	0.0447	(0.0109)
Number of siblings	-0.0360	(0.0071)	-0.0302	(0.0068)

Notes: General cognitive skill: $s^c_{1i} = x'_{si}\alpha^c + \theta^{mech}_i$ and mechanical skill:
$s^{mech}_{1i} = x'_{si}\alpha^{mech} + \theta^{mech}_i$, where x_s is a vector of observed variables.
Standard errors are in parentheses.
Standard errors are calculated by bootstrap. The number of replication is 200.

Table B.2: ASVAB test score equations estimates

	s_1^c		s_1^c		Error	
WK		1		0	$\log(Var(e_1))$	-1.1338
						(0.0247)
PC	(δ_{12})	1.0233		0	$\log(Var(e_2))$	-1.2388
		(0.0138)				(0.0285)
AR	(δ_{13})	1.0390		0	$\log(Var(e_3))$	-1.3212
		(0.0155)				(0.0295)
MK	(δ_{14})	1.0373		0	$\log(Var(e_4))$	-1.3090
		(0.0160)				(0.0269)
NO	(δ_{15})	0.7615		0	$\log(Var(e_5))$	-0.5003
		(0.0178)				(0.0230)
MC	(δ_{16})	0.5335	(δ_{26})	0.6397	$\log(Var(e_6))$	-1.0773
		(0.0218)		(0.0287)		(0.0285)
AI		1		1	$\log(Var(e_7))$	-0.7482
						(0.0319)
EI	(δ_{18})	0.4850	(δ_{28})	0.6929	$\log(Var(e_8))$	-1.0787
		(0.0209)		(0.0281)		(0.0294)
	$\log(Var(\theta^c))$	-0.8618	$\log(Var(\theta^{omech}))$	-1.5863	$Corr(\theta^c, \theta^{mech})$ (a)	0.5038
		(0.0278)		(0.0720)		(0.0179)

Notes: Equations (3.3) to (3.10).

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

WK: Word Knowledge; PC: Paragraph Comprehension; AR: Arithmetic Reasoning; MK: Mathematics Knowledge;

NO: Numerical Operation; MC: Mechanical Comprehension; AI: Auto & Shop Information; EI: Electronics Information.

Table B.3: Parameter estimates of education level choice equation

	High school		Some college		College	
Const	$(\tilde{\eta}_{0HS})$	0	$(\tilde{\eta}_{0Some})$	-1.0765	$(\tilde{\eta}_{0College})$	-0.9228
				(0.3237)		(0.4128)
s_1^c	(η_{1HS})	0	(η_{1Some})	0.7823	$(\eta_{1College})$	2.3326
				(0.1020)		(0.1509)
s_1^{mech}	(η_{2HS})	0	(η_{2Some})	-0.0791	$(\eta_{2College})$	-0.9353
				(0.1389)		(0.1574)
LUR	$(\varphi_{LUR,HS})$	-0.0077	$(\varphi_{LUR,Some})$	-0.0398	$(\varphi_{LUR,College})$	-0.4577
		(0.0316)		(0.0514)		(0.1523)

Notes: Base group is high school. LUR: Local unemployment rates. Equation (B.4).

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Table B.4: College major choice equation estimates

	Humanities&SS		Business&Econ		STEM	
Parameters						
Const	$(\tilde{\eta}_{0H})$	0	$(\tilde{\eta}_{0E})$	-1.5149	$(\tilde{\eta}_{0S})$	-2.2176
				(0.6983)		(1.0105)
s_1^c	(η_{1H})	0	(η_{1E})	-0.0928	(η_{1S})	1.0917
				(0.2076)		(0.4065)
s_2^m	(η_{2H})	0.8398	(η_{2E})	-0.6813	(η_{2S})	1.9520
		(0.3701)		(0.5811)		(1.6242)
Math relative score	$(\varphi_{M,H})$	0	$(\varphi_{M,E})$	2.1878	$(\varphi_{M,S})$	2.6633
				(0.4279)		(0.5266)
Mechanical relative score	$(\varphi_{M,echH})$	0	$(\varphi_{Mech,H})$	0.2210	$(\varphi_{Mech,H})$	0.7569
				(0.1690)		(0.2143)
Marginal effects at means*						
s_1^c		-0.1434		-0.0164		0.1598
s_2^m		0.1512		-0.0376		0.2841

Notes: Base group is Humanities&SS. Base group is Humanities&SS. Equation (B.5).

*θ^c is set to one standard deviation above 0 and θ^m is set to 0.

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Table B.5: Job type choice

	High school		Some college	
s_1^c	(ξ_1^c)	0.8991	$(\tilde{\xi}_1^c)$	0.6876
		(0.1026)		(0.2363)
s_1^{mech}	(ξ_2^c)	-0.8371	$(\tilde{\xi}_2^c)$	-0.6065
		(0.1542)		(0.2588)
Const	$(\tilde{\xi}_0^c)$	0.6258	$(\tilde{\xi}_{0,Some}^c)$	0.8658
		(0.2728)		(0.4219)

Notes: Dependent variable = 0: Mechanical job; =1: Cognitive job.

Equation (B.3) for high school and equation (B.6) for some college.

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Table B.6: Parameter estimates: latent GPA equations

		Humanities&SS		Business&Econ		STEM	
		$\log Var(s_2^H)$	-0.0366	$\log Var(s_2^E)$	-0.1144	$\log Var(s_2^S)$	-0.3363
			(0.1754)		(0.1599)		(0.6662)
GPA_1^{*m}	Const	$(\tilde{\gamma}_{01}^H)$	-0.8868	$(\tilde{\gamma}_{01}^E)$	0.0272	$(\tilde{\gamma}_{01}^S)$	-1.4647
			(0.1652)		(0.2911)		(0.6314)
	s_1^c	$(\tilde{\gamma}_{11}^H)$	0.4675	$(\tilde{\gamma}_{11}^E)$	0.5077	$(\tilde{\gamma}_{11}^S)$	0.6538
			(0.1121)		(0.1154)		(0.2022)
	s_2^m		1		1		1
	Error	$(\log Var(e_1^H))$	-1.7119	$(\log Var(e_1^E))$	-2.2157	$(\log Var(e_1^S))$	-0.7235
			(0.3886)		(0.3568)		(0.8152)
GPA_2^{*m}	Const	$(\tilde{\gamma}_{02}^H)$	-0.9830	$(\tilde{\gamma}_{02}^E)$	-0.2183	$(\tilde{\gamma}_{02}^S)$	-1.2845
			(0.1417)		(0.2478)		(0.5596)
	s_1^c	$(\tilde{\gamma}_{12}^H)$	0.5599	$(\tilde{\gamma}_{12}^E)$	0.5842	$(\tilde{\gamma}_{12}^S)$	0.6864
			(0.0820)		(0.1297)		(0.1911)
	s_2^m	(γ_{22}^H)	0.6613	(γ_{22}^E)	0.6888	(γ_{22}^S)	1.0911
			(0.1205)		(0.0891)		(0.4968)
	Error	$(\log Var(e_2^H))$	-0.6090	$(\log Var(e_2^E))$	-0.7170	$(\log Var(e_2^S))$	-1.1686
			(0.1973)		(0.1765)		(0.8949)

Notes: Equations (3.32) and (3.33).

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Table B.7: Job relatedness choice

	Humanities&SS		Business&Econ		STEM	
s_1^c	$(\tilde{\xi}_{1rH})$	-0.0950	$(\tilde{\xi}_{1rB})$	0.1419	$(\tilde{\xi}_{1rS})$	0.1402
		(0.1406)		(0.2649)		(0.3012)
s_2^m	(ξ_{2rH})	0.3108	(ξ_{2rB})	0.3588	(ξ_{2rS})	0.1678
		(0.1087)		(0.1595)		(0.2351)
Const	$(\tilde{\xi}_{0rH})$	-0.1280	$(\tilde{\xi}_{0rB})$	0.9856	$(\tilde{\xi}_{0rS})$	-2.1851
		(0.3904)		(0.7267)		(0.8945)

Notes: Dependent variable = 0: Unrelated job; =1: Related job.

Equation (B.7).

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Table B.8: Cognitive skill growth parameter estimates; log wage point metric

Some college			Business&Economics		
Const	$(\pi_1\lambda^c_{0Some})$	0.1136	Const	$(\pi_1\lambda^c_{0B})$	0.3122
$\pi_1 s^c_1$	(λ^c_{1Some})	0.9827	$\pi_1 s^c_1$	(λ^c_{1B})	1.0188
Humanities&Social Sciences			**STEM**		
Const	$(\pi_1\lambda^c_{0H})$	0.2434	Const	$(\pi_1\lambda^c_{0S})$	0.3986
$\pi_1 s^c_1$	(λ^c_{1H})	0.0497	$\pi_1 s^c_1$	(λ^c_{1S})	1.5426

Notes: For $m_+ = Some, H, E, S$, post-college general cognitive skill in the log wage metric is $\pi_1 s_{2m_+,i} = \pi_1\lambda^c_{0m_+} + \lambda^c_{1m_+}\pi_1 s^c_{1i}$.

Table B.9: Log wage equation estimates excluding task intensity

τ^c	(π_{21})	0	Business & Econ		
$(\tau^c)^2$	(π_{22})	0	s^c_1	(π_{1B})	0.0463
					(0.0468)
High school			Unrelated		
Const	$(\tilde{\pi}_0)$	1.8819	Const	$(\dot{\pi}_{0B,nr})$	2.4369
		(0.0947)			(0.1257)
s^c_1	(π_1)	0.0536	s^E_2	$(\pi^B_{3,nr})$	0.0101
		(0.0127)			(0.0374)
Some college			Related		
Const	$(\dot{\pi}_{0Some})$	2.1943	Const	$(\dot{\pi}_{0B,r})$	2.4957
		(0.0960)			(0.1364)
s^c_1	(π_{1Some})	0.0608	s^E_2	$(\pi^B_{3,r})$	0.0440
		(0.0286)			(0.0362)
Humanities & SS			STEM		
s^c_1	(π_{1H})	0.0059	s^c_1	(π_{1S})	0.1030
		(0.0414)			(0.0637)
Unrelated			Unrelated		
Const	$(\dot{\pi}_{0H,nr})$	2.3602	Const	$(\dot{\pi}_{0S,nr})$	2.4481
		(0.1003)			(0.2185)
s^H_2	$(\pi^H_{3,nr})$	-0.0036	s^S_2	$(\pi^S_{3,nr})$	-0.0047
		(0.0282)			(0.0647)
Related			Related		
Const	$(\dot{\pi}_{0H,r})$	2.2911	Const	$(\dot{\pi}_{0S,r})$	2.5723
		(0.1066)			(0.2194)
s^H_2	$(\pi^H_{3,r})$	0.0003	s^S_2	$(\pi^S_{3,r})$	0.1076
		(0.0261)			(0.0706)

Notes: Equation (3.26) for high school, equation (3.34) for some college, and equation (3.35) for major m, $m = H, B, S$.

Other parameter estimates are omitted in this table.

Standard errors are in parentheses.

Standard errors are calculated by bootstrap. The number of replication is 200.

Appendix C

Chapter 4 Appendix

Table C.1: Probability of being a manager at age 30: coefficient estimates from Probit regressions

	All		First job=Non management	
Business & Economics	0.3816**	0.4597***	0.4289**	0.4889***
	(0.1644)	(0.1704)	(0.1781)	(0.1832)
STEM	0.1360	0.2733	0.0606	0.1577
	(0.1727)	(0.1895)	(0.1916)	(0.2090)
Social task intensity in first job		0.3845***		0.2298**
		(0.0938)		(0.1076)
Cognitive task intensity in first job		-0.1560*		-0.1007
		(0.0888)		(0.0935)
N	482	482	445	445

Dependent variable is a dummy of being a manager at age 30.

Standard errors are in parentheses.

* p<0.1, ** p<0.05, *** p<0.01.

Estimates of a constant and year dummies are omitted.

Curriculum Vitae

Name:	Yuki Onozuka
Post-Secondary Education and Degrees:	University of Western Ontario London, ON, Canada 2019 Ph.D. in Economics (expected)
	Hitotsubashi University Kunitachi, Tokyo, Japan 2012 M.A. in Economics
	Hitotsubashi University Kunitachi, Tokyo, Japan 2010 B.A. in Economics
Honours and Awards:	Summer Paper Prize University of Western Ontario 2015
	Study Abroad Fellowship Japan Student Service Organization (JASSO) 2013-2018
Related Work Experience:	Research Assistant Research Institute of Economy, Trade and Industry (RIETI) 2012-2016
Publications:	"University Prestige, Performance Evaluation, and Promotion: Estimating the Employer Learning Model Using Personnel Datasets" (with Shota Araki and Daiji Kawaguchi), 2016, *Labour Economics*, 41, pp. 135-148.
	"The Gender Wage Gap and Sample Selection in Japan" 2016, *Journal of the Japanese and International Economies*, 39, pp. 53-72.