Earnings premia of undergraduate single and double majors

Ammar Plumber

University of Pennsylvania

McNeil Building, Rm 438
3718 Locust Walk, Philadelphia, PA 19104

Abstract

This paper quantifies the earnings premia attained by college graduates with different single and double major combinations per year and in aggregate over an eleven-year period. I use data from the National Longitudinal Survey of Youth 1997 cohort to produce a multiple regression model. To attenuate selection bias, I include in my model measures of cognitive and noncognitive ability among other relevant demographic variables. I find that pursuing a second major in almost every case predicts greater earnings than pursuing one of the two majors alone. STEM and Business majors appear to yield the greatest positive income effects, even when ability bias is attenuated.

Keywords: educational economics, college major choice, rate of return, earning differentials, human capital, undergraduate, ability bias
JEL: I20, I21, I23, I26, J24, J31

1. Introduction

The value of a college degree is an oft-discussed topic because of its vast policy and career planning implications. It is well-established that degree-holders enjoy a substantial earnings premium relative to non-graduates[1]. Many studies have examined the differential returns to majors[2, 3]. For instance, STEM majors, on average, earn more per year than any other major[4]. However, another question arises: regarding earnings, is choice of major of causal significance, or are there other characteristics that explain these differences? There may be cognitive and non-cognitive differences between majors, and perhaps these determine earnings more strongly than major choice itself.

Neglected is the question of how double majors influence earnings relative to their single major counterparts. The dearth of literature on this topic is likely attributable to the recency of the double major phenomenon; the number of double majors at UC Davis has increased 50% from 2006 to 2011, and this population has doubled at MIT since 1993[5]. The public and private benefits of double majors are discussed in Del Rossi and Hersh (2016). The authors find that a double major predicts

Email address: ammarp@sas.upenn.edu (Ammar Plumber)

significant monetized benefits[6]. As an example of the potential social benefits of double majors, the American Academy of Arts and Sciences (AAAS) argues that educational breadth can equip STEM-trained researchers with creativity and complex problem-solving abilities, which could enhance product design and development[7]. Noting this possibility and others, policymakers or institutional leaders may conclude that double majors should be more strongly encouraged.

This paper focuses on the private economic payoff to pursuing a double major. Quantifying this benefit may help to clarify whether the private incentive is sufficient to induce students to choose the double major path. Because students with high intellectual ability and motivation are more likely to double major, it is important to control for these features when quantifying earnings premia—a bias that prior literature has neglected to confront. From a policy standpoint, as David Epstein notes in his latest book *Range*, institutional leaders may wish to quantify the relationship between cross-disciplinary education and earnings (arguably a proxy for productivity or value added)[8]. Beyond its policy implications, this study provides important information for students. The opportunity costs of a double major are potentially substantial; an increased courseload may imply foregone social, extracurricular, and professional opportunities, and the challenge may depress one's GPA and, accordingly, one's employment prospects. Upon quantifying the payoffs, independent of ability, one is better equipped to weigh the benefits against the costs. As such, this paper offers a crucial piece to a complex puzzle.

2. Previous Literature

A number of specific questions motivate analyses of this topic. First, what is the general effect of college degree attainment on earnings? Card (1999) is one influential paper that examines this effect by reviewing findings obtained via a number of experimental designs: instrumental variables based on institutional features of universities, differences in education and earnings between twins (to remove family effects), and a regression of returns to education on regional primary and secondary school quality[9]. As the substantial effect of a college degree on earnings is well-established, many have turned their attention to the differences in earnings across majors. How much more do the majors with the greatest returns earn than the majors with the least returns? Altonji, Blom, and Meghir (2012) examine data from the 2009 American Community Survey (ACS) and find that gaps in wage log rates between male electrical engineering and general education majors (0.561) after adjusting for demographics, labor market experience, and graduate education is just slightly less than the gap in wage log rates between college graduates and high school graduates (0.577)[2]. As such, it is at minimum clear that the differences in wage premia across majors are significant and require explanation.

Second, are the differences in returns to college majors a result of differences in skillsets or the disciplinary focuses of the majors themselves? In other words, is an arts major less valuable than a social science major only insofar as the former does not confer a quantitative skillset? Taking seriously this intuition, Grogger & Eide (1995) examine the effect of college skill attainment on lifetime earnings. They regress wages premia for men and women on distributions of major choices and find that the increasing trend toward math-heavy majors explains increases in returns to college degrees and, in particular, narrowing gaps in wage premia between men and women (as women are acquiring quantitative skills at an increasing rate)[10]. Hamermesh & Donald (2008), a more recent study, uses survey data to regress current salaries for each major on SAT scores, GPA, income, working hours, and upper division math and science coursework (credits and grades). They find that students who are identical in every other regard–especially major selection, GPA, and math ability as measured through standardized testing–are apt to earn more if they have taken more upper division STEM classes. This finding suggests that quantitative and scientific skills, beyond their relevance to a major, have intrinsic value in terms of expected earnings[11].

There is also a question of causality, complicated in part by the aforementioned possibility of selection bias. Do certain majors yield greater returns, or do people predisposed to earn more (those with greater ability or motivation) self-select into certain majors? If the latter is the case, one might also wonder whether these choices are driven by expectations of greater earnings or interest. Berger (1988) regresses major selection data from the National Longitudinal Survey (NLS) on entry-level and lifetime earning expectations specific to different majors (acquired through interviews) and finds that expected lifetime earnings better explain college major selection than do entry-level earnings expectations[12]. However, whether major choice is causally significant can only be assessed by accounting for ability bias in major choice. Studies of this kind will be discussed shortly. A related question concerns what other factors might predict differences in disciplinary interests between students. Summarizing a breadth of literature on the topic, Altonji, Blom, and Meghir (2012) suggest that gender, interest, earnings expectations, and cognitive ability all play a significant role[2].

Noting the relevance of skills to earnings premia, it is still clear that major choice exerts an independent effect[11]. However, concerns about selection or ability bias appear salient. As such, the models that appear most promising are those that consider the roles of cognitive and non-cognitive ability. Few papers appear to do this, and, of those that do, Walker and Zhu (2011) is one of the few that examines returns over the life cycle rather than at a specific point in time[13, 3]. The advantage of this life-cycle-based approach is that some specialized majors may have greater entry-level earnings but are eventually surpassed by more generalist or intellectually versatile students later in their careers. As such, it is most helpful to examine longitudinal earnings trajectories (panel data) of individuals over time. Webber (2014), the paper on which my present analysis is based, utilizes this approach but accounts for a key possibility neglected by Walker and Zhu (2011) perhaps due to data constraints: that major choice is endogenous–in part explained by cognitive and non-cognitive abilities[3, 13, 14]. Using data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY) and the American Community Survey (ACS), Webber (2014) formalizes two additional selection effects in his model. First, he includes an ordered logit that estimates the contribution of cognitive ability on likelihood of attaining a college degree. Second, he uses a multinomial logit to account for the effect of cognitive ability on major choice. Note that Webber (2014) utilizes a simulation approach. After estimating the wage premia of different majors, he produces a model that simulates annual earnings given the additional factors of age, ability, and selection bias[14]. Webber's approach has two virtues that are not captured by the pre-existing literature: his focus on earnings trajectories rather than cross-sectional data and his endogenous treatment of major selection. This paper uses a more recent cohort of the same data source (NLSY) and a similar OLS regression model to study the earnings premia of different double major pairings.

The aforementioned versatility, which specialized majors potentially forego, is what David Epstein refers to as cross-disciplinary "range"[8]. Whether this breadth effect is significant in terms of earnings premia merits further study. Only one paper, Hemelt (2010), studies the effect of double majors on earnings, and he does so using the 2003 National Survey of College Graduates (NSCG)[15]. The only wage variable he examines is the earnings from the previous year (2002), and, as such, he cannot account for differential returns over time as Webber (2014) does. Neither does he include measures of cognitive or non-cognitive ability into his linear regression model. For these reasons, I find the pre-existing literature deficient on this question, a problem that this paper purports to address.

3. Data

This analysis uses data from the National Longitudinal Survey of Youth[16]. The NLSY collects annual survey data from a set of respondents over the course of their lifetimes. The surveys focus on sociologically, economically, and politically relevant phenomena such as labor market activities, education, and familial patterns.

Panel or longitudinal data allow a kind of regression analysis that cross-sectional data do not, namely that earnings data can be aggregated for each individual over time. Thus, in addition to modeling incomes in each year, we can model an individual's earnings over ten years, a trajectory that is governed in part by individual characteristics that remain constant over the time period (i.e. race, sex, education, cognitive ability, and noncognitive ability).

As of now, there have been two cohorts of NLSY respondents. The first cohort is comprised of men and women born between 1957 and 1964, and 1979 is the initial survey year. The other cohort, comprised of people born between 1980 and 1984, was first surveyed in 1997, when respondents were ages 12-17.

Unlike Webber (2014), we examine the 1997 cohort for a few reasons. Most importantly, the surveys administered to the 1979 cohort do not ask respondents about a second major. Second, research indicates that double majoring was not a popular choice until recently (the 2000s); students pursuing double majors, even if surveyed in 1979, would lack the sample size necessary to derive precise coefficient estimates[5]. In addition, even if data were available in a large enough sample size, it is tenuous to surmise that conclusions would remain applicable to today's labor market given long-term changes to the economy and what employers value. For instance, it is possible that technological gains have obviated certain niche or hyper-specialized skillsets, a trend that may favor synthetic thinkers or diversely educated students in the present. Conversely, technology creates specialist jobs where none have existed before. For instance, an excellent and singularly focused computer science student may have greater labor market value than a liberal arts and social science major. Given these possibilities, we err towards recency so that inferences made about the modern labor market may be more defensible.

Finally, note that a disadvantage to using the 1997 cohort is that respondents' earnings data are unavailable beyond the age of 37, the age of oldest respondents during the most recent survey year (2017). Thus, this analysis cannot yield conclusions about the effect of double or single major choice on lifetime earnings, as the analysis applies only to earnings spanning little more than a decade after college graduation. Respondents must have received a high school diploma to be included in the regression. In addition to this exclusion, a number of other variables are deemed relevant to the regression analysis: total earnings in the past year, primary and secondary majors, race/ethnicity (Black, Hispanic, White, non-black/non-Hispanic, or mixed race/non-Hispanic), sex (male or female), age, cognitive ability measured using the Armed Services Vocational Aptitude Battery (ASVAB) score, noncognitive ability measured using a so-called Industriousness score, and mother's level of education.

A number of cleaning tasks and adjustments were applied to the data, which are detailed here along with a few other notes about how variables are coded.

Using the Consumer Price Index (CPI) from the US Bureau of Labor Statistics, income values were converted to base year 1997 dollars[17]. A few additional adjustments had to be made to the data. In addition to the values of reported incomes, four additional responses are possible: Refusal, Don't Know, Valid Skip, and Non-Interview. "Refusal," "Don't Know," and "Non-Interview" responses were removed from the dataset. "Valid Skip," for the purposes of this question, means that an individual's response to a previous question indicated that the individual earned no income, and, therefore, no response was needed for this question. As such, a "Valid Skip" response was assigned zero income. Finally, starting in 2011, incomes were reported biannually, so there are missing years. As such, I interpolated values for the years in between (2012, 2014, and 2016). The interpolated incomes were simply the average of the incomes reported in the prior and the subsequent year. If either the prior or subsequent year's income is unreported, then the interpolated variable is assigned a null value, as if it, too, were unreported.

Primary and secondary majors were classified according to the appended table (see Appendix A), almost identical to the classification used in Webber (2014)[14]. The schema used includes four categories of majors: STEM, Business, Social Science, and Art/Humanities.

An additional assumption was needed for the purposes of assigning majors to individuals. Because choice of primary and secondary college major are surveyed each year, the answers given are apt to reflect the respondents' changing plans. Thus, I assumed that the mode of reported majors was ultimately the one pursued by the respondent. If the mode for any given respondent was "None, no major yet", then I assumed that the respondent did not complete the major. If no primary major could be identified under this assumption, then the respondent was categorized as a high school graduate. The NLSY opts to consolidate race and ethnicity into one variable 'KEY!RACE_ETHNICITY' and makes a few additional assumptions. While respondents can be both Black and Hispanic, Hispanic is given priority for respondents who select both options. Most curious is that Asian is not included as an option. Asian, White, American Indian, etc. are all coded as non-black/non-Hispanic.

I removed earnings data for which respondents were below the age of 23. This was to ensure that college graduates in the sample had already completed their degrees and had likely entered the workforce before earnings data are factored into the model. High school graduates, by this point, had likely acquired a few years of work experience, which allows the benefit of experience gained to be compared against the value of a college degree once both college and high school graduates are participants in the workforce.

Webber (2014) uses the NLSY 79 cohort's Armed Forces Qualification Test (AFQT) scores as a proxy for cognitive ability[14]. This score is computed using results from four of nine subtests from the Armed Services Vocational Aptitude Battery (ASVAB). These four subtests are Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Mathematics Knowledge. The NLSY 97 cohort, on the other hand, simply reports the total ASVAB score as an indication of cognitive ability. The five additional subtests on the ASVAB that are not part of the AFQT score are as follows: General Science, Electronics Information, Automotive and Shop Information, Mechanical Comprehension, and Assembling Objects. While it seems that some of these categories seem somewhat extraneous to general cognitive ability, we have no choice but to rely on ASVAB as an indication of cognitive ability, as no other cognitive ability assessments are included in the survey. Perhaps the inclusion of additional subtest scores such as Mechanical Comprehension and Assembling Objects reflect practical and cognitive aptitudes that are not reflected in the AFQT score and would exert a meaningful effect on long-term earnings.

Two measures of noncognitive ability are included in the model produced by Webber (2014): the Rotter Scale and the Rosenberg Score[14]. The Rotter Scale reflects the degree to which an individual believes that he or she has control over life outcomes relative to external factors. The Rosenberg Score measures, as the name indicates, self-esteem, which encompasses beliefs about self-efficacy and self-awareness. Neither of these scores are reported in the 1997 cohort. Instead, I use a mean of four scores taken during two survey years (2008 and 2010) that are intended to capture an individual's industriousness. These questions, drawn from the Chernyshenko Conscientiousness Scales (CCS), measure one's work ethic and ambition—traits that are strongly representative of overall conscientiousness. The questions and possible responses are included in the appendix (see Appendix B). Although the industriousness scores certainly capture different characteristics than Rotter and Rosenberg, they are shown to be highly correlated with career earnings and academic success[18, 19]. While it is inconvenient that industriousness scores are only available in 2008 and 2010, it is worth noting that conscientiousness tends to increase rather than decrease during young adulthood. Therefore, those with low industriousness scores in the survey years must have experienced weak improvement in the trait between 1997 and 2008[20]. While it is unknown whether high industriousness individuals improved or held constant in this respect, both trajectories would portend increased earnings relative to low industriousness individuals. We may also question whether industriousness is a cause or effect of major choice; is the score obtained in 2008 a product of differences in major curriculum, or did the developmental trajectory precede and influence the choice? However sparse, evidence points to the latter proposition being more salient for conscientiousness, providing limited assurance that 2008/2010 scores are useful for attenuating selection bias [21].

Respondents' mother's educational background is included mainly because it is deemed to exert a relevant effect by Webber (2014) [14]. It may reflect a familial factor that is perhaps not directly captured by industriousness or ASVAB. A person whose mother attended college may be more intrinsically motivated or encouraged by family to seek education than a person whose mother did not. Because the variable reports mothers' years of education, I converted the variable to a dummy—zero if below 13 years of education are reported and one if 13 or more years are reported, which I assume to mean that the mother completed her first year of college. One might believe that it would be more useful to set this demarcation at 16 years or four years of college. However, the dataset does not include a variable for whether the mother received a college degree, and the number of years required for a college degree can vary—two for an associate degree and maybe even five or six years for certain degree programs. I surmise that a mother's attendance of college is significant even if the mother drops out before completing the degree; the mother would have some familiarity with the college application process and perhaps places more value on higher education than a mother who never planned to attend college in the first place. I assess the significance of this variable in the regression models that follow.

4. Empirical Methodology

As noted previously, I produce two kinds of regressions. The first kind models disaggregated incomes using the aforementioned regressors (primary and secondary major, race, sex, age, cognitive and noncognitive ability, and mother's education); each reported income is predicted separately without grouping incomes by individual. This disaggregated regression serves to produce more datapoints to predict, which is useful given that the sample size of each major choice category is limited. Moreover, the effect of age/experience on earnings over time can be estimated, whereas an aggregated model entails that the sum total of earnings over a certain timeframe is predicted rather than earnings during each year of an individual's career. The disaggregated regression is specified as follows:

$$\begin{aligned} Y_{it} = \alpha_0 + \alpha_1 Age_{it} + \alpha_2 Sex_i + \alpha_3 Race_i + \alpha_4 MotherEd_i \\ + \alpha_5 ASVAB_i + \alpha_6 Indust_i + \alpha_7 Majors_i \end{aligned}$$

The subscript i indexes individuals, and t denotes different survey years. *Age* is defined as years above the age of 23, the minimum age for an observation to be included in the regression. *Sex* is a dummy variable: 1 if male, 0 if female. *Race* is a vector of racial/ethnic dummy variables. Only those racial variables that appear significant will remain in the final model. As previously detailed, *MotherEd* is a dummy variable indicating whether a mother completed at least 13 years of education, assigned a value of 0 if not. *ASVAB* is defined as an individual's percentile relative to the mean score in the sample. The *Indust* variable is defined as the number of points an individual accrues above or below the mean industriousness score in the sample. The maximum obtainable score is 7, and the minimum is 1 (a 7-point Likert scale). Finally, *Majors* is a vector of dummy variables for each major combination—both single and double majors. The excluded (reference) category, represented by the regression's intercept α_0, is high school graduates without a college degree.

The second kind of regression model aims to predict aggregated incomes over an eleven-year period. I choose eleven years as the longest predictively viable earnings period because any greater duration would exclude a substantial number of younger respondents from the regression. This type of regression is valuable because accounts for persistent individual characteristics that influence earnings in each year. A disaggregated model treats each reported income as distinct values to predict with independent and identically distributed regressors, when, in fact, each individual is responsible for multiple reported incomes. Note that the coefficient estimates under this specification will be less statistically significant than under the disaggregated regression model, as there are substantially fewer datapoints to predict. The aggregated model is as follows:

$$\begin{aligned} Y_i = \alpha_0 &+ \alpha_1 Sex_i + \alpha_3 Race_i + \alpha_4 MotherEd_i + \alpha_5 ASVAB_i \\ &+ \alpha_6 Indust_i + \alpha_7 Majors_i \end{aligned}$$

The key difference between the first and second regression is that Y_i denotes total income accrued by each individual between the ages 23-34. Apart from this difference in regressands and the exclusion of an age variable, the two regressions are identically specified.

5. Results and Discussion

Before proceeding to the regression analysis, I offer a few noteworthy preliminary findings and summary statistics. Below is a table indicating the number of individuals in the sample with each major combination. Note that some individuals of these individuals are excluded from the aggregate regression because they report less than eight incomes over the eleven-year period.

Table 1: Summary Statistics for Each Education Category

Major Combo	**Count**	**% Female**	**Inc**	**ASVAB pctl rel. to mean**	**Industriousness rel. to mean**
HS	2786	44.8	29660.82	-11.48	-0.12
STEM	1123	51.65	49120.53	9.62	0.07
Social Science	1002	63.47	40842.73	5.56	0.09
Business	692	48.99	48278.16	7.27	0.14
Art/Humanities	493	53.35	35883.71	13.21	0.01
SocSci + SocSci	73	60.27	50285.68	14.88	0.11
SocSci + Art/Hum	72	72.22	52026.82	21.47	0.13
Art/Hum + Art/Hum	34	61.76	48975.19	25.05	0.17
STEM + SocSci	31	61.29	51414.44	3.71	0.14
STEM + STEM	31	58.06	61068.31	11.86	-0.02
STEM + Business	26	42.31	58671	8.31	0.31
Business + Business	23	60.87	63153.97	21.02	0.32
Business + Art/Hum	21	57.14	56140.25	16.74	0.38
Business + SocSci	18	44.44	61271.24	17.09	0.36
STEM + Art/Hum	17	29.41	63270.33	22.56	0.06

The largest group in the sample is of high school graduates without college degrees. Double majors make up 5.37% of the overall sample and 9.46% of college graduates in the sample. Two social science subjects appears to be the most popular double major pairing, followed closely by a social science paired with an arts/humanities subject. The least popular major pairings are STEM paired with arts/humanities and business paired with a social science major.

It is clear from the percentages of females in each category that the majority of college graduate survey respondents are female. In the sample, women appear to be especially well-represented in social science fields. While it is possible that the sample does perhaps not reflect the population's sex distribution of college graduates, US Census data shows that by 2014, a greater percentage of women than men in America had completed a college degree[22]. In 1979, 20.4% of men in America had completed a college degree while only 12.9% of women had done so. The impressive reversal of this gap is likely the result of women outnumbering men in college participation in recent decades.

Among double majors, STEM+Art/Humanities majors appear to earn the most, followed by Business+Business, Business+Social Science, STEM+STEM, and STEM+Business. It is clear that business and STEM backgrounds are valued by the labor market. The fact that double Art/Humanities majors earn the least among double majors despite having the greatest average ASVAB scores (and greater Industriousness scores than STEM+Art/Humanities majors) indicates that skillsets are more fundamental than ability when it comes to earning outcomes.

Mean ASVAB scores appear to be higher among double majors than among single majors, and both categories perform significantly better than high school graduates. It is interesting that those who study arts/humanities appear to perform the best, both among single majors and double major pairings.

Also note the substantial differences in industriousness scores betweeen the following groups: high school graduates, single major college graduates, and double major college graduates. High school graduates receive industriousness scores below the sample mean. Single major college graduates appear more industrious than the mean, with business majors receiving the highest scores among single majors. Finally, double majors (with the puzzling exceptions of STEM+STEM and STEM+Art/Humanities) receive very high industriousness scores. It is clear, therefore, that college participation and major choice are subject to selection bias on the basis of motivation and ability.

One possible concern about using these variables is that the model will exhibit multicollinearity; some variables can be predicted using others, which muddles the estimated effects of each regressor. To assess the seriousness of this problem, I produce a correlation matrix showing the relationships between key regressors.

Table 2: Correlation Matrix for Regressors

	Mother Attend Col	**ASVAB**	**Indust.**	**Black**	**Hispanic**	**HS**
Mother Attend Col	1	0.288166	0.018708	-0.04698	-0.18672	-0.1971
ASVAB	0.288166	1	0.100025	-0.31652	-0.16576	-0.36112
Indust.	0.018708	0.100025	1	-0.04733	-0.00838	-0.12368
Black	-0.04698	-0.31652	-0.04733	1	-0.26404	0.030659
Hispanic	-0.18672	-0.16576	-0.00838	-0.26404	1	0.07878
HS	-0.1971	-0.36112	-0.12368	0.030659	0.07878	1

A few correlation coefficients stand out. Whether or not a respondent's mother attended college appears to be somewhat correlated with ASVAB scores. Being placed in the high school category and being Hispanic are negatively correlated with one's mother attending college. ASVAB scores are negatively correlated with being placed in the high school category and being Black or Hispanic. Industriousness is negatively correlated with HS as well. However, none of these correlation coefficients exceed an absolute value of 0.4, and, thus, I do not consider them so large to justify *prima facie* excluding the regressors. Instead, I will produce the regression models and exclude variables if the corresponding coefficient estimates have high p-values or otherwise diminish the models' predictive strength or parsimony.

The disaggregated regression model is presented below. Note that the regression's dependent variable is scaled (Income/1,000) to ensure that coefficient estimates are not too large.

Table 3: Disaggregated Model Results

	coef	std err	t	P> |t|	[0.025	0.975]
Intercept	8.68	0.565	15.36	0.00	7.57	9.79
Male	16.55	0.396	41.84	0.00	15.8	17.3
Age (years above 23)	3.85	0.053	72.35	0.00	3.75	3.96
Mother Attended College	1.76	0.414	4.23	0.00	0.94	2.57
ASVAB Percentile	0.32	8.36	38.30	0.00	0.30	0.34
Industriousness Score	6.34	0.237	26.73	0.00	5.88	6.81
STEM	12.4	0.574	21.61	0.00	11.3	13.5
Business	10.57	0.675	15.67	0.00	9.25	11.9
Social Science	4.92	0.596	8.25	0.00	3.75	6.09
Art/Humanities	-3.97	0.794	-4.99	0.00	-5.52	-2.41
STEM + STEM	31.1	2.718	11.44	0.00	25.8	36.4
STEM + Business	28.84	3.51	8.22	0.00	22.0	35.7
STEM + SocSci	14.22	2.719	5.23	0.00	8.89	19.5
STEM + Art/Hum	16.97	3.31	5.13	0.00	10.5	23.5
Business + Business	36.74	3.344	10.99	0.00	30.2	43.3
Business + SocSci	20.13	3.611	5.57	0.00	13.0	27.2
Business + Art/Hum	17.55	3.345	5.25	0.00	11.0	24.1
SocSci + SocSci	12.11	1.717	7.05	0.00	8.74	15.5
SocSci + Art/Hum	9.76	1.698	5.74	0.00	6.43	13.1
Art/Hum + Art/Hum	0.12	2.459	0.05	0.96	-4.70	4.94
Black	-3.42	0.525	-6.52	0.00	-4.45	-2.39
Hispanic	-0.11	0.552	-0.201	0.841	-1.19	0.97

Dep. Variable:	Income/1000	**R-squared:**	0.189
Model:	OLS	**Adj. R-squared:**	0.189
No. Observations:	52658	**F-statistic:**	583.6
Df Residuals:	52636	**Prob (F-statistic):**	0.00
Df Model:	21	**Log-Likelihood:**	-6.3852e+05
Covariance Type:	nonrobust	**AIC:**	1.277e+06
		BIC:	1.277e+06

For this regression model, the Hispanic dummy variable's coefficient estimate yielded a p-value above 0.8 and a 95% confidence interval from -1.19 to 0.97, which strongly imply the variable's insignificance. Art/Humanities double majors also produced a very high p-value, an indication that having such a degree does not with any reasonable degree of confidence predict earnings above or below high school graduates. I felt that these findings are informative and relevant to this paper's overall aim, so I choose to keep both variables in the model.

The regression's R∧2 value is low—below 0.2. This is typical of models that predict income, as earnings outcomes are difficult to predict using a parsimonious set of variables. Some determinants of income are inherently chancy (i.e. fortuitous social linkages, labor market shocks, etc.). Those that are less chancy and have predictive relevance are perhaps too innumerable to incorporate; growing up in a prosperous neighborhood, having access to a savvy mentor, and many other factors could reliably influence on lifetime earnings. However, our explanatory purposes do not motivate their inclusion. We aim to separate the effect of college major from a particular set of biases enumerated above. As such, the low R∧2 value does not invalidate the general tendencies observed in the model and the interpretations to be gleaned from them.

I now offer interpretations of the models' coefficients. The intercept shows the predicted annual income of an age 23, non-black, female high school graduate with no college degree with the sample mean ASVAB and industriousness scores and whose mother did not attend college. Being male predicts an additional $16,550 in income. A mother's college attendance predicts $1,760 more in earnings, and each additional year past 23 predicts an increase of $3,850. Each ASVAB percentile point predicts an additional $320, and a point above the mean Industriousness score implies $6,340 more. A STEM major predicts the largest increase over high school graduates among single majors, followed by Business, Social Science, and, finally, Art/Humanities, which predicts earning $3,970 *less* than high school graduates. For each possible double major combination, the model implies that having both majors predicts greater earnings than having only one of the two. Business+Business predicts the greatest boost, followed by STEM+STEM and STEM+Business.

Most interesting is that the effect of double major pairings comprised of only STEM and/or Business is greater than the effect of a STEM+Art/Humanities double major, despite the latter category averaging more annual income according to our summary statistics (Table 1). Two observations in the summary statistics together suggest a hypothesis for why this disparity in coefficients is observed. First, STEM+Art/Humanities has the lowest percentage of females among all double major combinations. Second, this category has the second highest mean ASVAB score among all education categories. Therefore, it is possible that the higher earnings among STEM+Art/Humanities double majors is driven by sex distribution (being male predicts greater earnings) as well as cognitive ability. This hypothesis is confirmed upon removing these variables from the regression; the resulting coefficient estimates for wage effects are much closer to one another.

It is clear from both double and single major coefficients that STEM and business majors have the most early career labor market value. This finding supports broad-mindedness and liberal arts training are less profitable early in one's career relative to the acquisition of specialized skillsets with immediate industrial value.

After accounting for differences in cognitive and noncognitive ability, double majors are still seen to command higher earnings than their single major counterparts. There are a few possible reasons for this effect. One is that the skillsets acquired through an additional major offers increased value to firms, and this advantage is compensated accordingly. Another is that double majors are a reliable signal to employers of increased ability or motivation, and, as such, employers are willing to pay a premium for the assurance. Further research is needed to assess which of these explanations is correct or whether another might be more fitting.

Finally, it is clear that race is a significant predictor of annual earnings. Being black predicts a negative effect on income. Many sociological explanations have been offered in the literature for why this negative effect is observed. These include systematic features of the criminal justice system, institutionalized racism, instability in communities, and many other salient factors.

It is also worth examining the second model specification, which predicts incomes aggregated by individual over eleven years. To communicate coefficient estimates most clearly, I again scale the dependent variable, this time by a greater magnitude (Income/100,000); larger coefficient values reflect predicted earnings differentials over an eleven-year period as opposed to annual earnings in the disaggregated model.

Table 4: Aggregated Model Results

	coef	std err	t	P> |t|	[0.025	0.975]
Intercept	2.872	0.119	24.06	0.00	2.64	3.11
Male	1.621	0.115	14.13	0.00	1.4	1.85
ASVAB Percentile	0.031	0.0023	13.14	0.00	0.026	0.035
Industriousness Score	0.625	0.070	8.97	0.00	0.488	0.761
STEM	1.17	0.164	7.14	0.00	0.849	1.49
Business	1.36	0.195	6.99	0.00	0.979	1.74
Social Science	0.331	0.178	1.864	0.06	-0.017	0.68
Art/Humanities	-0.656	0.221	-2.964	0.003	-1.09	-0.222
STEM + STEM	2.906	0.835	3.481	0.001	1.27	4.54
STEM + Business	5.203	1.14	4.57	0.00	2.97	7.44
STEM + SocSci	1.951	0.931	2.095	0.036	0.125	3.78
STEM + Art/Hum	2.574	0.865	2.977	0.003	0.879	4.27
Business + Business	4.449	0.898	4.96	0.00	2.69	6.21
Business + SocSci	1.846	1.02	1.808	0.071	-0.156	3.85
Business + Art/Hum	2.481	0.933	2.66	0.008	0.652	4.31
SocSci + SocSci	1.164	0.497	2.341	0.019	0.189	2.14
SocSci + Art/Hum	1.016	0.493	2.06	0.039	0.049	1.98
Art/Hum + Art/Hum	-0.731	0.68	-1.074	0.283	-2.06	0.603
Black	-0.386	0.171	-2.258	0.024	-0.721	-0.051
Hispanic	-0.15	0.199	-0.754	0.451	-0.54	0.24

Dep. Variable:	Income/100,000	**R-squared:**	0.205
Model:	OLS	**Adj. R-squared:**	0.201
No. Observations:	3227	**F-statistic:**	43.61
Df Residuals:	3207	**Prob (F-statistic):**	2.01e-144
Df Model:	19	**Log-Likelihood:**	-45478.
Covariance Type:	nonrobust	**AIC:**	9.100e+04
		BIC:	9.112e+04

The aggregated model, like the disaggregated one, revealed the Hispanic dummy to be insignificant, with negligibly small coefficient estimate and a p-value exceeding 0.4. Yet, we choose to include the dummy in this model, as its insignificance is relevant. Unlike the disaggregated model, whether or not the respondent's mother attended college initially appeared insignificant in this model. The p-value obtained was greater than 0.4. By removing other variables from the model and observing the changes to the coefficient estimates and p-values, it appeared that ASVAB scores largely captured the effect of mother's college attendance; upon removing ASVAB percentile, the coefficient estimate for mother's college attendance became significant. However, the model performed substantially better when ASVAB was included and mother's college attendance was not. The R∧2 and adjusted R∧2 values were higher, and the Bayesian Information Criterion (BIC) was lower. Therefore, I included ASVAB in the final model and excluded mother's college attendance.

Again, I interpret the coefficients in terms of what the model predicts. The intercept represents the expected earnings over eleven years ($287,200) of a non-black, female high school graduate with no college degree with the sample mean ASVAB and industriousness scores. Being male predicts $162,100 in additional earnings. Each percentile point increase in ASVAB score predicts a $3,100 increase in earnings over the eleven-year period. A point above the mean Industriousness score predicts an additional $62,500. STEM majors are expected to enjoy the largest earnings premia over high school graduates among single majors. Business is next, followed by Social Science, and Art/Humanities, which predicts $65,600 *less* in earnings over the eleven-year period than high school graduates.

As with the disaggregated model, each possible double major predicts greater earnings than having only one of the two. Unlike the disaggregated model, STEM+Business predicts the greatest eleven-year earning premia, followed by Business+Business. STEM+STEM and STEM+Art/Humanities seem to significantly trail the leading two double major pairings in this regard. It is not obvious how this difference between the two models can be explained. Perhaps STEM+Business majors enjoy much bigger increases in annual earnings as they gain experience relative to Business+Business double majors (note that age is excluded from the aggregated model). In other words, perhaps there is an interaction between age and major choice. However, to test this intuition by including age-major interaction terms in the model would require us to double the number of regressors, and doing so substantially decreases the adjusted R∧2 values and increases the BIC and AIC. Let it suffice for now to observe that, while Business+Business predicts the greatest earnings in any individual year, STEM+Business predicts the greatest cumulative earnings over the eleven years after graduation. Perhaps this finding lends credence to the notion that cross-disciplinary range can be add value in over time (even if not immediately), provided that the range of disciplines in question yield diverse skillsets with substantial industrial value. Further research may be needed to confirm or falsify this interpretation.

This model backs a number of insights gleaned from the disaggregated model. For instance, this model, too, shows that the effect of an Art/Humanities double major is not significantly distinguishable ($p>0.25$) from the effect of a high school diploma. However, a single Art/Humanities major again predicts reduced earnings over the eleven-year period. Moreover, major choices that include Business and STEM predict substantially greater earnings over the eleven-year sample period than the others. Likewise, the addition of a second major is seen to predict greater earnings over the sample period. It is clear that the value conferred by STEM/Business skillsets or by a second major persists over a decade at minimum. Whether this balance shifts beyond the sample period cannot be inferred from this model.

The aggregated model confirms the significance of race on earnings. Over eleven years, being black predicts a negative effect on wages of $38,600. Possible reasons for this were noted in the interpretation of the disaggregated model.

6. Conclusion

This study aimed to estimate the effect of major choice, for both single and double majors, on earnings in each year and over an eleven-year period. Two models were used to examine these effects: a disaggregated model that predicts reported incomes in each year separately and an aggregated one that predicts earnings for each individual over a sample period of eleven years.

Both models yielded a number of interesting findings. The value of cross-disciplinary range from an earnings standpoint appears directly related to the labor market value of the corresponding skillset. For instance, the models show that the combination of STEM+Art/Humanities predicts a greater boost to earnings than STEM alone, but STEM+STEM, STEM+Business, and Business+Business have the greatest positive income effects—even though none of these combinations involve a wider cross-disciplinary range than STEM+Art/Humanities. What explains this outcome is that the expertise conferred by a Business or STEM major attracts greater compensation than does an Art/Humanities second major.

It remains possible that a second liberal arts or social science major generates substantial long-term returns beyond the eleven-year sample period. It is also possible that people with broad academic backgrounds have higher earnings *ceilings* than do double STEM or double business majors. The data from the sample period does not lend credence to this possibility; histograms of incomes for each double major category show that those with STEM or Business majors report most of the highest income values in the sample (see Appendix C). Further research is needed to examine whether this tendency remains true beyond eleven years.

Cognitive and noncognitive ability also appear to be partly responsible for the increased earnings premia of double majors. Double majors, on average, report higher ASVAB and Industriousness scores than single majors, and these scores exert significant effects on income according to both models. However, even with these differences accounted for, having a second major appears to have an independent and positive effect on income; for every single major, a second major—no matter the category—predicts greater earnings under both the aggregated and disaggregated models.

These findings are significant from both a policy and a familial/individual decision-making standpoint. While there may be an opportunity cost to having a second major (forgone social, extracurricular, and professional opportunities due to an increased courseload), the effects on income appear to be net positive across the board. It is likely, therefore, that a second major ultimately aids long-term employment prospects and compensation. It may be best for institutions to encourage income-motivated students to pursue a second major and for students to consider this information when choosing their majors.

Appendix A: Categories of Majors

Table A.1: Categories of Majors

Major	**Category**
Agriculture/Natural resources	STEM
Anthropology	Social Science
Archaeology	Social Science
Architecture/Environmental design	Art/Humanities
Area studies	Art/Humanities
Biological sciences	STEM
Business management	Business
Communications	Social Science
Computer/Information science	STEM
Criminology	Social Science
Economics	Social Science
Education	Social Science
Engineering	STEM
English	Art/Humanities
Ethnic studies	Art/Humanities
Fine and applied arts	Art/Humanities
Foreign languages	Art/Humanities
History	Art/Humanities
Home economics	Social Science
Interdisciplinary studies	Art/Humanities
Mathematics	STEM
Nursing	STEM
Other health professions	STEM
Philosophy	Art/Humanities
Physical sciences	STEM
Political science and government	Social Science
Pre-dental	STEM
Pre-law	Art/Humanities
Pre-med	STEM
Pre-vet	STEM
Psychology	Social Science
Sociology	Social Science
Theology/religious studies	Art/Humanities

Appendix B: Industriousness Questions and Answers

Note that the following set of four Industriousness questions were given to respondents twice—once in 2008 and again in 2010. The Industriousness scores used in the regression models are an average of all eight response scores. Questions 1 and 2 are coded in reverse; agreeing strongly implies low industriousness, and disagreeing strongly conveys high industriousness. I modified the values accordingly so that they can be averaged with the values reported in Questions 3 and 4; for Questions 1 and 2, "Disagree strongly" is averaged in as a 7 (high industriousness), and a "Agree strongly" is counted, for our purposes, as a 1 (low industriousness). The questions, as administered in the NLSY surveys, are as follows:

"Now I will read some statements that may or may not apply to you. On the same scale, where 1 means disagree strongly and 7 means agree strongly, please tell me how much you agree or disagree that each statement describes who you are and how you act."

Table A.2: NLSY Possible Responses

Score	**Response**
1	Disagree strongly
2	Disagree moderately
3	Disagree a little
4	Neither agree nor disagree
5	Agree a little
6	Agree moderately
7	Agree strongly

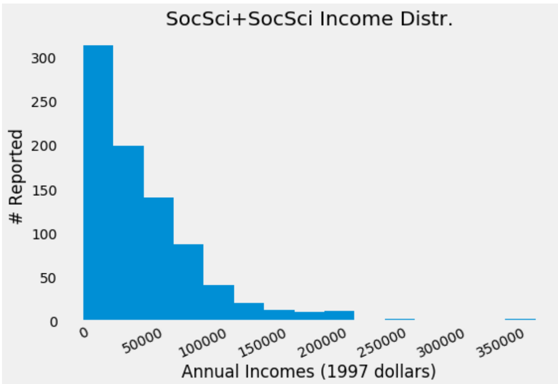
Q1. I do not work as hard as the majority of people around me.

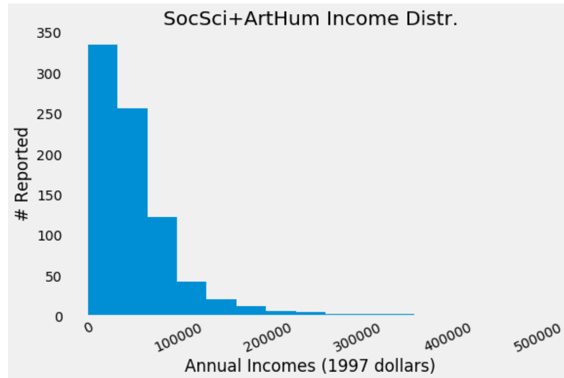
Q2. I do what is required, but rarely anything more.

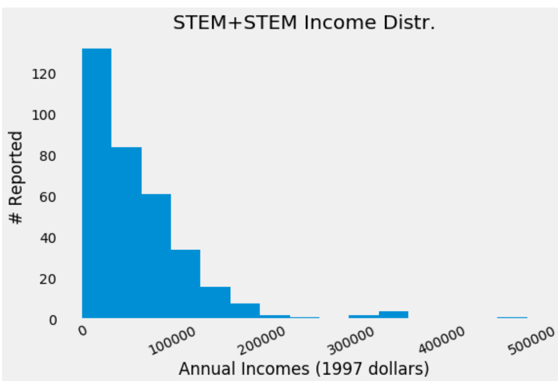
Q3. I have high standards and work toward them.

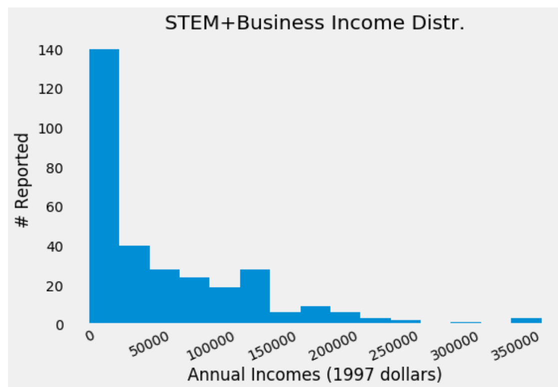
Q4. I make every effort to do more than what is expected of me.

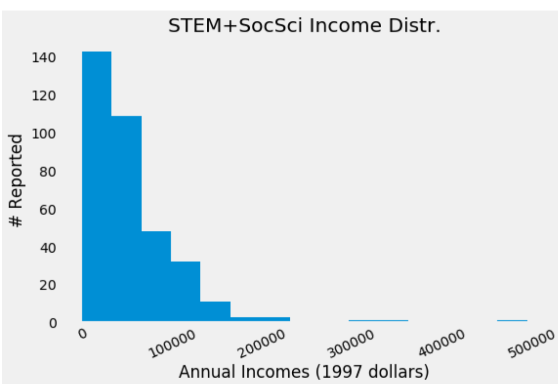
Appendix C: Distributions of Disaggregated Incomes for Each Double Major Pairing

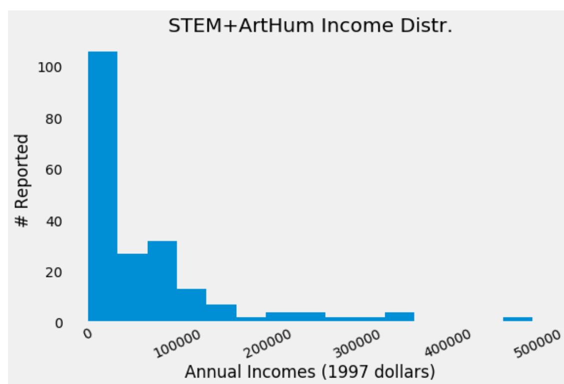


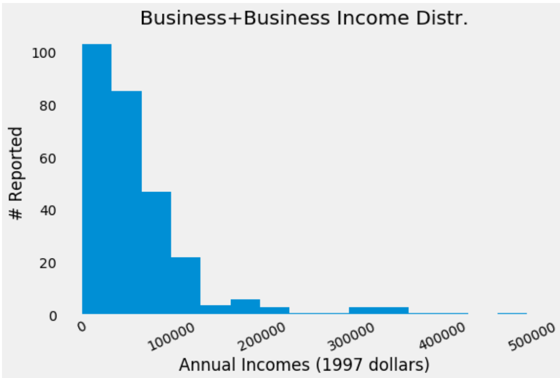


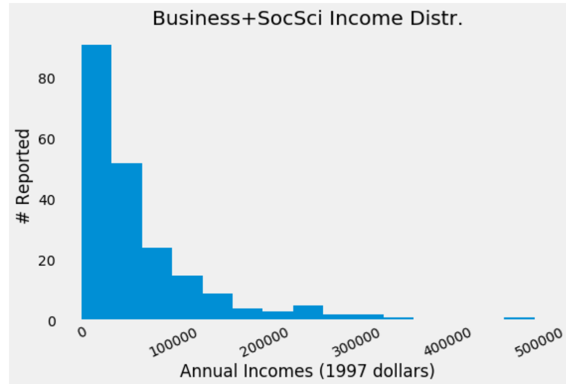


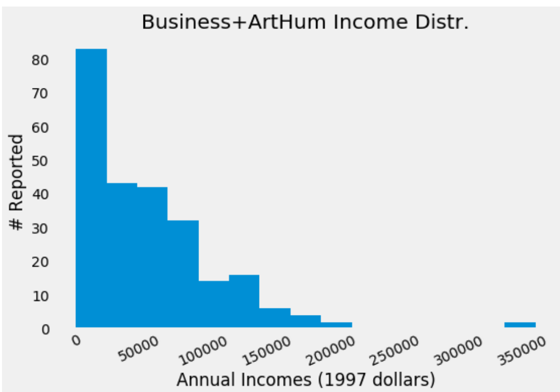


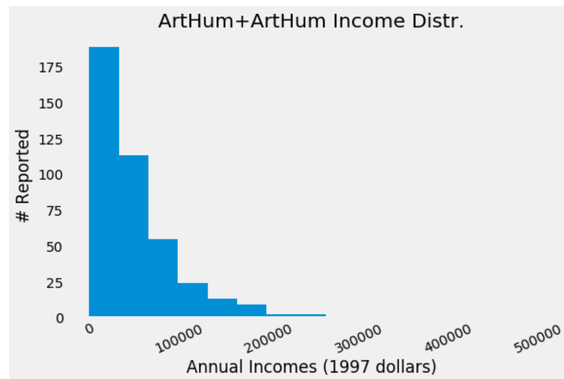












Declaration of Interest

None.

Acknowledgements

I would like to thank Professor Jere Behrman and Dr. Francesco Agostinelli for their invaluable guidance on this paper. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. A. Carnevale, B. Cheah, S. Rose, The College Payoff: Education, Occupations, Lifetime earnings. URL https://cew.georgetown.edu/cew-reports/the-college-payoff/

2. J. Altonji, E. Blom, C. Meghir, Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers, Annual Review of Economics 4 (2012) 185–223. doi: 10.2307/42949936.
URL www.jstor.org/stable/42949936

3. D. Webber, Are college costs worth it? How ability, major, and debt affect the returns to schooling, Economics of Education Review 53 (2016) 296–310. doi:10.1016/j.econedurev.2016.04.007.
URL https://doi.org/10.1016/j.econedurev.2016.04.007

4. A. Carnevale, B. Cheah, A. Hanson, The Economic Value of College Majors.
URL https://cew.georgetown.edu/cew-reports/whats-it-worth-the-economic-value-of-college-majors/

5. R. Pitt, S. Tepper, Double Majors: Influences, Identities, and Impacts.
URL https://www.researchgate.net/publication/279985369_Double_Majors_Influences_Identities_and_Impacts

6. A. D. Rossi, J. Hersch, The Private and Social Benefits of Double Majors, Journal of Benefit-Cost Analysis 7 (2) (2016) 292–325. doi:10.1017/bca.2016.14.

7. American Academy of Arts and Sciences, The Heart of the Matter (2013).
URL https://www.humanitiescommission.org/_pdf/hss_report.pdf

8. D. Epstein, Range: Why Generalists Triumph in a Specialized World, Riverhead Books, 2019.

9. D. Card, The Causal Effect of Education on Earnings, in: Handbook of Labor Economics, Vol. 3A, Elsevier, 1999, pp. 1801–1863.

10. J. Grogger, E. Eide, Changes in College Skills and the Rise in the College Wage Premium, The Journal of Human Resources 30 (2) (1995) 280–310.
URL http://www.jstor.com/stable/146120

11. D. Hamermesh, S. Donald, The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias, Journal of Econometrics 144 (2008) 479–491. doi:10.1016/j.jeconom.2008.04.007.
URL https://doi.org/10.1016/j.jeconom.2008.04.007

12. M. Berger, Predicted Future Earnings and Choice of College Major, Industrial Labor Economics Review 41 (3) (1988) 418–429. doi:10.2307/2523907.
URL https://www.jstor.org/stable/2523907

13. I. Walker, Y. Zhu, Differences by degree: Evidence of the net financial rates of return to undergraduate study for England and Wales, Economics of Education Review 30 (6) (2011) 1177–1186. doi:10.1016/j.econedurev.2011.01.002.
URL https://doi.org/10.1016/j.econedurev.2011.01.002

14. D. Webber, The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors, Labour Economics 28 (2014) 14–23. doi:10.1016/j.labeco.2014.03.009.
URL https://doi.org/10.1016/j.labeco.2014.03.009

15. S. Hemelt, The college double major and subsequent earnings, Education Economics 18 (2) (2010) 167–189. doi:10.1080/09645290802469931.
URL https://doi.org/10.1080/09645290802469931

16. US Bureau of Labor Statistics, National Longitudinal Survey of Youth 1997 Cohort.
URL https://www.nlsinfo.org/content/cohorts/nlsy97

17. US Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items in

U.S. City Average [CPIAUCSL], retrieved from FRED, Federal Reserve Bank of St. Louis.
URL https://fred.stlouisfed.org/series/CPIAUCSL

18. J. A. Green, D. B. O'Connor, N. Gartland, B. W. Roberts, The Chernyshenko Conscientiousness Scales: A New Facet Measure of Conscientiousness, Assessment 23 (3) (2016) 374–385. doi:10.1177/1073191115580639.

19. T. Bogg, B. W. Roberts, The Case for Conscientiousness: Evidence and Implications for a Personality Trait Marker of Health and Longevity, Annals of Behavioral Medicine 45 (3) (2013) 278–288. doi:10.1007/s12160-012-9454-6.

20. M. P. A. Van Dijk, W. W. Hale III, S. T. Hawk, W. Meeus, S. Branje, Personality Development from Age 12 to 25 and its Links with Life Transitions, European Journal of Personality 34 (3) (2020) 322–344. doi:10.1002/per.2251.

21. M. Balsamo, M. Lauriola, A. Saggino, Personality and College Major Choice: Which Come First?, Psychology 3 (5) (2012) 399–405. doi:10.4236/psych.2012.35056.

22. US Census Bureau, Percentage of the U.S. population who have completed four years of college or more from 1940 to 2019, by gender [Graph], In Statista (March 2020).
URL https://www.statista.com/statistics/184272/educational-attainment-of-college-diploma-or-higher-by-gender/