Res High Educ (2017) 58:365–394
DOI 10.1007/s11162-016-9431-1

Choice of Academic Major at a Public Research University: The Role of Gender and Self-Efficacy

Iryna Y. Johnson[1] · William B. Muse[2]

Received: 22 June 2015 / Published online: 30 July 2016

Abstract Females are underrepresented in certain disciplines, which translates into their having less promising career outlooks and lower earnings. This study examines the effects of socio-economic status, academic performance, high school curriculum and involvement in extra-curricular activities, as well as self-efficacy for academic achievement on choices of academic disciplines by males and females. Disciplines are classified based on Holland's theory of personality-based career development. Different models for categorical outcome variables are compared including: multinomial logit, nested logit, and mixed logit. Based on the findings presented here, first generation status leads to a greater likelihood of choosing engineering careers for males but not for females. Financial difficulties have a greater effect on selecting scientific fields than engineering fields by females. The opposite is true for males. Passing grades in calculus, quantitative test scores, and years of mathematics in high school as well as self-ratings of abilities to analyze quantitative problems and to use computing are positively associated with choice of engineering fields.

Keywords Holland's theory of vocational choices · Social cognitive theory · Multinomial logit · Nested logit · Mixed logit · Gender

✉ Iryna Y. Johnson
iyj0001@auburn.edu

William B. Muse
Muse_William@columbusstate.edu

1 Office of Institutional Research, Auburn University, 203 Samford Hall, Auburn, AL 36849-5111, USA

2 Department of Mathematics, Columbus State University, 218 University Hall, Columbus, GA 31907, USA

Introduction

Choice of major is an important career decision. First, it affects one's odds of finding employment, since career outlook varies greatly by a disciplinary area. According to the Occupational Outlook Handbook of the Bureau of Labor Statistics (2015), employment of petroleum engineers is projected to grow 26 percent; while employment of reporters, correspondents, and broadcast news analysts is projected to decline 13 percent from 2012 to 2022. Secondly, it influences future earnings to a greater extent than the type of college one chooses to attend—more or less selective (James et al. 1989; Ma and Savas 2014). The 2012 median pay of petroleum engineers with Bachelor's degree—$130,280 a year—is about 3.5 times the median pay of reporters, correspondents, and broadcast news analysts with bachelor's degree—$37,090 a year (Bureau of Labor Statistics 2015). Additionally, choice of major is consequential for the required level of training and education. A successful career in science might require more years of study than a successful career in engineering. Lastly, choice of discipline affects the type of work one is likely to engage in.

Choice of discipline by females is puzzling as they seem to frequently overlook such important factors as future employment prospects or wages by discipline. While females outperform males in college enrollment, persistence, and degree attainment (Snyder and Dillow 2012; cited in DiPrete and Buchmann 2013), they tend to choose less lucrative fields with less favorable career outlooks. Based on National Science Foundation (2014), in 2010 female representation was the highest in social science occupations (58.1 %); followed by biological, agricultural, and environmental life science (48.2 %); physical science (30.0 %); computer and mathematical science (25.1 %); and engineering (12.7 %). Thus, gender segregation is most significant in engineering, where participation of females is roughly half of what it is in computer and mathematical science. Underrepresentation of women in certain disciplines translates into their lower earnings and poses questions about choices they make when they enter college.

Multiple empirical studies analyzed the determinants of college major choices. With respect to differences in choice of major of females and males, researchers (e.g., Davies and Guppy 1997; Song and Glick 2004) suggest that females pick majors with lower earning potentials. Ma (2009) indicates that females place greater emphasis on "intrinsic, altruistic, and social job rewards," while males place greater emphasis on "extrinsic rewards," such as money and prestige. Some researchers (e.g., Paglin and Rufolo 1990; Oakes 1990; cited in Ma and Savas 2014) explain female underrepresentation in certain fields of study by differences in abilities or academic preparation in math and science. However, others (e.g., Adelman 2003; cited in Ma and Savas 2014) found that gender differences in high school academic preparation are declining. According to Jacobs (1986), Lackland (2001), and Solnick (1995; cited in Porter and Umbach 2006), women tend to choose disciplines like education or nursing because of their gender role orientation.

Several researchers explored the association between socio-economic status (SES) and choice of major. However, evidence from these studies is not conclusive. According to Davies and Guppy (1997), SES does not affect choice of fields with higher economic returns. Contrary to this finding, Goyette and Mullen (2006) suggest that low SES students are more likely to choose vocational majors, while high-SES students choose arts and sciences. Similarly, Ma (2009) found that family SES has a significant influence on choice of major. "[L]ower SES families tend to choose technical, life/health science, and business majors—those higher paying fields upon graduation—over humanities and social science/education majors" (ibid, p. 277). Ma (2009) also suggests that the effect of family SES on college major choice might be asymmetric for men and women. "Men from a high SES family may still be expected to choose a lucrative career, whereas women from comparable backgrounds may not" (ibid, p. 215). For example, males from wealthy families become business majors more often than females from wealthy families (Green 1992; cited by Ma 2009).

Eide and Waehrer (1998) suggest that choices of students who intend to complete graduate studies might be different from choices of students who intend to enter the workforce after completing a bachelor's degree. Furthermore, the effect of the highest intended degree might differ by gender. For women the intent to complete a graduate degree is correlated with greater odds of choosing liberal arts and science majors. For men the intent to complete a graduate degree is associated with greater chances of choice of liberal arts and science majors as well as other majors except computer science/engineering.

Several studies focused on the association between self-rated abilities and self-efficacy on choice of major. Betz and Hackett (1983) suggest that mathematics self-efficacy is significantly related to the extent to which students selected science-based college majors. At that, males tend to have stronger mathematics self-efficacy than females. Smart et al. (2000) provided support for the association between student self-rated abilities and academic fields they choose. "[S]tudents who intended to major in a particular academic environment generally perceived themselves as having decidedly stronger abilities and interests commensurate with those that their chosen environment tended to reinforce and reward" (ibid, p. 116). Interestingly, the association between a choice of discipline and self-ratings of one's abilities and interests is similar for males and females across all disciplines with the exception of social disciplines. The magnitude of differences in self-rated abilities and interests of students who chose social and non-social disciplines was greater among males than among females (ibid), thus indicating that factors other than self-rates abilities and interests might lead females to choose social disciplines.

While previous studies contributed greatly to our understanding of major choices, most of these studies did not analyze the choices of males and females separately to account for possible gender differences in effects. For example, based on a previous study (Betz and Hackett 1983), mathematics self-efficacy is significantly related to the choice of a college major and is stronger for males. But does the effect of mathematics self-efficacy on the choice of a college major differ for males and females?

Our study is aimed at exploring the effects of different characteristics—i.e., socio-economic status, prior academic achievement, high school curriculum, extra-curricular involvement, and self-efficacy for academic achievement—on choice of discipline by females and males. Apart from substantive contribution, our study seeks to overcome the independence of irrelevant alternatives (IIA) assumption inherent in the multinomial logit model used in many existing studies of major choice. Since the Hausman–McFadden (HM) test rejects the IIA assumption in our study, we followed the multinomial logit analysis with a nested model and a mixed model. We also compared the consistency of average expected probabilities based on multinomial logit and mixed logit models.

Theoretical Framework

Holland's Theory of Vocational Choices

Our classification of academic disciplines is based on the Holland's (1966, 1973, 1987, 1997) theory of vocational choices—artistic, social, enterprising, conventional, realistic, and investigative. According to Holland, congruency between personality and career leads to greater occupational satisfaction and success. Hence, individuals have a tendency of choosing an occupation that fits their skills, abilities, attitudes and values. Thus, realistic occupations attract individuals who enjoy concrete and practical activities. Investigative professions are chosen by those who are fond of analytical and intellectual work. Artistic fields are for those who possess imagination, unconventional ideas, and aesthetic values. People in social occupations tend to be tactful, understanding, empathetic, responsible, and helpful. Enterprising jobs reward confidence, social skills, and ambition. Individuals in conventional occupations are good at following plans and paying attention to detail.

Nauta (2010) provides an overview of empirical studies that validate the existence of personality and occupational types suggested by Holland. While Holland (1997; cited in Nauta 2010) recognized that having six categories is a simplification of reality, he still abided by his model, since having a greater complexity would be less practical. For instance, in career counseling having fewer categories "help provide clients with families of occupational titles" (Nauta 2010, p. 17) and a greater number of possibilities to explore as opposed to having a single occupation to discuss.

Holland (1973, 1997; cited in Nauta 2010) suggests that occupational and personality types—artistic, social, enterprising, conventional, realistic, and investigative—are broadly applicable, yet acknowledged that gender impacts career development. Empirical studies support this idea. For example, Fouad (2002) indicates that males scored higher on realistic measures, while females scored higher on artistic and social measures. According to Betz and Gwilliam (2002), males report higher confidence on realistic, investigative, enterprising, and conventional inventories, while females report higher confidence on social inventory. Research also suggests that Holland's model fits data across gender (see Nauta 2010).

Porter and Umbach (2006) used Holland's theory to study the effects of political orientations and student personalities on choice of an academic major. They found that political orientations and student personalities have a significant effect on choice of a discipline, while effects of test scores disappear after control for student personalities. Porter and Umbach (2006) conclude that "Holland categories provide an excellent framework for the study of student major choice" (p. 445).

Smart et al. (2000) used Holland's classification in their study of characteristics of students entering academic fields. However, they excluded two out of Holland's six categories—conventional and realistic—because their focus was on artistic, social, enterprising, and investigative fields. Furthermore, very few college students and faculty fit into these categories (ibid, cited in Jones 2011). Following Smart et al. (2000), we also excluded the conventional category from the analysis as the data set for the study does not include college majors that fit into this category. (Future accounting majors enter a study university as pre-business majors.) However, we did include the realistic category, since our dataset includes a significant number of students in engineering and architecture who represent the realistic class (see, for example, Holland and Lutz 1967). See Table 1 for lists

Table 1 Variable descriptions and descriptive statistics

Variable description	Mean (SD)	
	Females	Males
Vocational choice		
Artistic (arts; communications; english; foreign language; music; theatre)	0.13 (0.34)	0.06 (0.23)
Enterprising (business; industrial and systems engineering; economics)	0.10 (0.30)	0.19 (0.39)
Investigative (agriculture, biosystems engineering; sciences; mathematics)	0.32 (0.47)	0.22 (0.41)
Realistic (architecture; building science; engineering; computer science)	0.11 (0.31)	0.45 (0.50)
Social (education; psychology; social sciences; health sciences; nursing)	0.33 (0.47)	0.08 (0.27)
Race, social class and highest degree expectation		
Caucasian (1 if caucasian, 0 otherwise)	0.90 (0.31)	0.86 (0.35)
First generation status (1 for first generation, 0 otherwise)	0.14 (0.34)	0.12 (0.32)
Financial stress (factor score)[a]	−0.01 (0.91)	0.01 (0.88)
Highest degree a student intends to obtain (1-associate's to 4-doctoral)	3.04 (0.71)	2.99 (0.69)
Prior achievement and high school curriculum		
High school GPA	3.83 (0.43)	3.69 (0.47)
ACT verbal score or SAT equivalent	27.32 (4.44)	27.02 (4.53)
ACT quantitative score or SAT equivalent	24.96 (4.11)	26.76 (4.20)
A passing grade in Calculus (1 for yes, 0 otherwise)	0.46 (0.47)	0.56 (0.47)
During high school, how many years of the following subjects did you complete? (1-none to 6-five or more)		
English/Literature	4.03 (0.21)	4.02 (0.20)
Math	4.08 (0.41)	4.12 (0.43)
Science	3.98 (0.46)	3.99 (0.46)
History/social sciences	3.92 (0.44)	3.93 (0.44)
Foreign language	2.69 (0.95)	2.54 (0.93)
Involvement in activities: during your high school years, how involved were you in the following activities at your school or elsewhere? (1-not involved to 6-highly involved)		
Performing or visual arts	3.05 (2.03)	2.42 (1.86)
Athletic teams	3.86 (2.04)	4.24 (1.93)
Student government	2.31 (1.81)	1.80 (1.42)
Publications	2.05 (1.71)	1.49 (1.13)
Academic honor societies	3.71 (1.92)	2.93 (1.81)
Academic clubs	2.37 (1.70)	2.18 (1.54)
Vocational clubs	1.79 (1.40)	1.73 (1.29)
Religious youth groups	3.44 (2.00)	2.95 (1.93)
Community service or volunteer work	4.35 (1.32)	3.68 (1.43)
Self-efficacy for academic achievement: How prepared are you to do the following in your academic work at this college? (1-not at all prepared to 6-very prepared)		
Write clearly and effectively	4.84 (1.07)	4.52 (1.15)
Speak clearly and effectively	4.66 (1.13)	4.59 (1.13)
Think critically and analytically	4.72 (1.03)	4.89 (0.99)
Analyze math or quantitative problems	4.22 (1.30)	4.59 (1.21)
Use computing or information technology	4.33 (1.16)	4.61 (1.12)
Work effectively with others	5.16 (0.90)	4.88 (0.97)

Table 1 continued

Variable description	Mean (SD)	
	Females	Males
Learn effectively on your own	4.94 (0.98)	4.84 (1.02)

To test the significance of mean differences between males and females, independent samples t test was conducted. Due to large sample sizes all differences are significant at the 5 % alpha level with the exception of years of history/social sciences in high school

[a] See Table 2 for variable description and descriptive statistics of observed variables used to calculate Financial stress

of disciplines and shares of female and male students in each of the five categories included in the analysis.

Gottfredson (1981; cited in Trusty et al. 2000) suggests an alignment between socio-economic status and Holland's occupational types. "I[nvestigative] occupations are the most prestigious. E[nterprising], A[rtistic], and S[ocial] occupations have roughly average levels of prestige, and R[ealistic] and C[onventional] occupations have the lowest level of prestige" (ibid, p. 464). Hence we expect to find an association between first generation status and a measure of financial stress, on the one hand, and choice of academic discipline, on the other hand.

The influences of demographics and academic performance on career choices have been accepted by career theorists (Trusty et al. 2000). With respect to ethnicity, some researchers (e.g., Swanson 1992; cited in Ryan et al. 1996) suggest that Holland's structure fits white students better. Yet Ryan et al. (1996) indicate that this lack of fit might be associated to socio-economic status rather than ethnicity. While we include ethnicity in our study, the vast majority of students in our sample are white and the number of students from other ethnic groups is rather small. Hence, our study does not analyze each ethnic group separately and compares white and non-white students.

According to Schneider and Overton (1983), educational achievement, measured by GPA and SAT scores, can be linked to Holland's personality types. The association between achievement and personality types also varies for males and females (ibid). Hence we include high school GPA and test scores as predictors in our study.

Skills and interests change as a result of learning experiences (Krumboltz 1996; cited in Trusty et al. 2000). Spending more time on certain subjects will result in improved achievement and might lead to change in interests (ibid; see also Görlitz and Gravert 2015). Therefore, we hypothesize that years of English/Literature, Math, Science, History/Social Sciences, and Foreign Language in high school affect choice of a discipline in college.

"The way students may validate congruency between themselves and their college major is through their involvement in high school and college extracurricular activities" (Patrick et al. 1993, p. 28). Participation in extracurricular activities—performing or visual arts, athletic teams, student government, publications, academic honor societies, academic clubs, vocational clubs, and religious youth groups—helps students acquire self-understanding and confidence for effective career decisions (ibid). We, therefore, hypothesize that high school extracurricular involvement affects a choice of a college major.

Table 2 Principal component analysis: financial stress

	Mean (SD)	Factor loading	Factor score
About how much of your college expenses this year will be provided by each of the following sources? (1 = none; 2 = less than half; 3 = half or more; 4 = all or nearly all)			
Student loans	1.53 (0.86)	0.69	0.34
Self (work on-campus or off-campus, savings)	1.57 (0.67)	0.74	0.36
During the coming school year, about how many hours do you think you will spend in a typical 7-day week doing each of the following? (1 = 0–8 = more than 30)			
Working for pay on- or off-campus	2.59 (1.70)	0.68	0.33
During the coming school year, how difficult do you expect the following to be? (1 = not at all difficult to 6 = very difficult)			
Paying college (university) expenses	3.25 (1.67)	0.74	0.36

Cronbach's alpha is .61; eigen value is 2.05; percentage of variance explained is 51.20 %

Prior research supports relations between interests, skills, and abilities within the same Holland type, but these associations tend to be rather small (Ackerman and Heggestad 1997; Randahl 1991; Swanson 1993, cited in Nauta 2010). At the same time associations between Holland type and self-efficacy tend to be more substantial (Betz et al. 1996; cited in Nauta 2010). Hence we proceed to discuss self-efficacy and social cognitive theory.

Social Cognitive Theory

According to social cognitive theory, self-efficacy is a person's belief in his or her ability to succeed in a particular situation (Bandura 1977, 1982, 1986, 1995, 1997, 2006). This belief in one's ability to succeed affects one's behavior. Efficacy varies by domain or "distinct realms of functioning" (Bandura 2006). Choice of academic major is expected to correlate with a belief in one's ability to perform different academic tasks—such as effective writing and speaking, critical thinking, analysis of quantitative problems, using computing and information technology, working independently, or working in team environment.

Huss et al. (2002) suggest that students' self-rated preparedness can be viewed from the perspective of self-efficacy or beliefs about one's ability to succeed at a given task. Bubany and Hansen (2010) compare self-efficacy with ability self-estimate scores and suggest that empirical differences may be due to measurement error or scale content, rather than due to meaningful reasons. Given these prior findings, we use self-rated preparation, abilities, and self-efficacy interchangeably. We expect that, controlling for other characteristics, a student's self-rated preparedness to analyze math and quantitative problems will have a positive effect on her odds of choosing an investigative or realistic field. A student's self-rated preparedness to write or speak clearly and effectively and to think critically and analytically is presumed to have a positive effect on her odds of choosing an artistic field. Self-rated preparedness to use computing or information technology might be associated with greater odds of choosing a realistic field. Students who are certain in their ability to work effectively with others might be more comfortable with social or enterprising fields, while those who feel prepared to learn effectively on their own might be inclined to choose investigative fields.

Given the varying expected returns by academic field and by degree level (Arcidiacono 2004; Berger 1988; Eide and Waehrer 1998), we also included expected earnings and intended degree in our study.

The dataset is based on student records and responses to the Beginning College Survey of Student Engagement. Variable descriptions as well as descriptive statistics of our study are presented in Table 1.

Data and Method

The institution studied is a Research University (high research activity) with about 4000 first-time freshmen enrolled each fall. Five fall cohorts of first-time freshmen (starting from fall of 2008) are included in this study. Each year, about 95 % of first-time freshmen at a study institution complete the Beginning College Survey of Student Engagement (BCSSE). Only those students who completed BCSSE are included. The data contain information on choices of one of five disciplinary categories—artistic, social, enterprising, realistic, and investigative—by 9918 females and 8939 males (see Table 1).

Consistent with Davies and Guppy (1997), Song and Glick (2004) and Ma (2009), females at a study institution are more likely to self-select into artistic, social, and investigative fields; while males tend to self-select into more lucrative careers—business, engineering, and architecture. One should note here that careers in fields selected by females might require more years of study; which corresponds to a greater mean time to intended degree for females. Females have, on average, higher high school grade point averages and ACT verbal scores, but lower ACT quantitative scores. The percentage of those who have passing grades in Calculus is also lower for females, compared to males. With respect to involvement in extra-curricular activities, females report greater involvement in performing or visual arts, student government, publications, academic honor societies, academic and vocational clubs, religious youth groups, and community service and volunteer work. Males report greater involvement in athletic teams. Females self-report greater preparation to write, speak clearly and effectively, work effectively with others, and learn effectively on their own. Males, on the other hand, feel more prepared to analyze math and quantitative problems, use computing and information technology, and think critically and analytically.

Note that one variable included in the analysis—financial stress—is a latent variable that was measured using the principal component analysis. Following Cole (2012), we included reliance on student loans and self to pay college expenses, number of hours a student plans to spend in a typical week working on- or off-campus, and perceived difficulty of paying college expenses as indicators of financial stress. Table 2 contains means and standard deviations of observed variables as well as factor loadings and factor scores for the model. The percentage of variance explained is 51.20 %. Cronbach's Alpha for this set of observed variables is 0.61. The factor loadings range from 0.68 to 0.74. Thus, the greater the index of financial stress, the more difficulty a student expects to encounter paying college expenses.

Selection into majors depends on the monetary returns (e.g., Arcidiacono 2004). An alternative specific variable included in our study—natural logarithm of salary—is based on the Salary Survey of National Association of Colleges and Employers (2014). The salary for enterprising disciplines is based on the average salary for business ($54,234). Most of artistic and social disciplines are assigned salaries for humanities and social sciences ($37,058). Exceptions are: communications with the average salary of $43,145; education with the average salary of $40,480; and health with the average salary of $49,713. Aside from computer science with the average salary of $59,977, realistic fields were assigned a salary of engineering ($62,535). Investigative disciplines were assigned a salary of math and sciences ($42,724).

The outcome variable—a discipline chosen by a first-time freshman—is categorical and calls for a multinomial model. The most commonly used model for categorical outcomes is multinomial logit model (MNLM), according to which the probability of observing an outcome i for an individual n is (Train 2009, p. 37):

$$P_{ni} = \frac{\exp(x_{ni}\beta)}{\sum_{j=1}^{J} \exp(x_{nj}\beta)} \quad (1)$$

where the dependent variable has J categories numbered from 1 to J, x is a vector of independent variables and β is a vector of regression coefficients.

The multinomial logit model assumes the independence of irrelevant alternatives (IIA). This assumption implies that adding or removing categories of the outcome variable does not change the relative risks of the remaining categories. Let's consider a scenario in which we could separate accounting majors and add a conventional discipline to our model. If IIA holds, student odds of choosing an enterprising discipline over artistic, investigative, realistic, or social discipline should not depend on adding a conventional discipline to our model. In practice, however, the odds of choosing an enterprising discipline would likely go down, since some of the pre-business majors can be expected to choose accounting if this option was available to them.

To test for the IIA assumption, we use the Hausman–McFadden (HM) test, which involves the following steps (Long and Freeze 2014; pp. 408–409): (1) estimate the full model with all J alternatives included; (2) estimate a model with one or more alternatives excluded; and (3) conduct HM test that compares parameter estimates of the full set with those of the subset. If the parameters of the full set and parameters of the subset are inconsistent, the IIA assumption is rejected.

Because the IIA assumption does not hold for our models, we estimated nested logit and mixed logit models. Given the number of parameters we estimate, the multinomial probit model (MNP) produced a difficult computational problem relative to nested or mixed logit and did not reach the point of saturation. A similar observation was made by Dow and Endersby (2004) who stated that "[t]he MNP presents a difficult maximum likelihood optimization problem that sometimes fails to converge at a global optimum or produces parameter estimates that are sufficiently imprecise as to make statistical inferences suspect" (p. 109). Thus, we will limit our discussion to three models—multinomial logit, nested logit, and mixed logit.

By allowing correlation between some choices, the nested logit model partially relaxes the IIA assumption. While the IIA holds within nests, it does not hold for alternatives in different nests (Train 2009). For example, if we combine realistic and investigative disciplines into one nest, the ratio of probability of choosing realistic over investigative disciplines will remain the same regardless of removing artistic, social, or enterprising disciplines. At the same time, the ratio of probabilities of choosing realistic over social disciplines—i.e., disciplines that do not belong to one nest—might change if one of the remaining alternatives is removed. The nested logit probability is the product of two standard logit probabilities—the probability of choosing nest B_k and the probability of choosing an alternative i within the nest B_k (ibid, p. 82):

$$P_{ni} = P_{ni|B_k} P_{nB_k} \tag{2}$$

where $P_{ni|B_k}$ is the conditional probability of choosing the alternative i given an alternative in nest B_k is chosen and P_{nB_k} is the probability of choosing an alternative in nest B_k.

"Mixed logit probabilities are the integrals of standard logit probabilities times the density of parameters" (ibid, p. 135):

$$P_{ni} = \int \left(\frac{\exp(x_{ni}\beta)}{\sum_{j=1}^{J} \exp(x_{nj}\beta)} \right) f(\beta) d\beta \tag{3}$$

where the dependent variable has J categories numbered from 1 to J; x is a vector of independent variables; β is a vector of regression coefficients; and $f(\beta)$ is a density that provides the weights. Thus, "[t]he mixed logit probability is a weighted average of the logit formula evaluated at different values of β, with the weights given by the density $f(\beta)$." (ibid, p. 135). The standard logit model is a special case of mixed logit model where $f(\beta) = 1$ if β equals fixed parameter b and 0 if $\beta \neq b$.

Given our research focus, we would ideally like to accompany our comparison of regression coefficients across gender with statistical tests for the significance of differences. In linear regression this could be done either by estimating two separate models for males and females and using Wald Chi square statistic to test the difference between coefficients or by estimating a single model that interacts gender with other independent variables in the model. Unfortunately, use of these methods in logit and probit models has been questioned in several studies. Using an example of logit regressions predicting the probability of promotion to associate professor for males and females, Allison (1999; cited in Hoetker 2004; Long 2009; Williams 2009; Mood 2010) illustrates that the difference in the two coefficients for article counts is an artifact of differences in the degree of residual variation for men and women. Allison (1999) concludes that unless we are willing to assume that the residual variation is constant across groups, both standard tests—Wald Chi square test for coefficients in two separate models or a single model that interacts group variable with other variables—tell us nothing about actual differences in the effects of our variables. With respect to single equation approach, Hoetker (2004) states: "Econometric theory and simulation results suggest that tests interacting coefficients with a dummy variable for group membership in a single equation are particularly misleading. Forcing observations from both groups to have the same residual variation yields coefficients that tell us nothing about how a covariate's impact varies across groups" (p. 17). The problem is "very similar to the well-known problems with comparing standardized ordinary least square (OLS) coefficients across groups" (Williams 2009, p. 534). But while the solution for comparing coefficients in OLS—i.e., use of unstandardized coefficients as opposed to standardized coefficients—is clear, approaches suggested for addressing residual variation in logit and probit models have limitations.

Allison (1999) proposes a test that accounts for residual variation, but adds an assumption of identical regression coefficients for certain variables across groups. This latter assumption is difficult to justify. Williams (2009) suggests that in some cases "Allison's procedure could make things worse rather than better" (p. 547). Other proposed solutions to the problem include use of predicted probabilities (Long 2009), use of ratios of coefficients (as opposed to coefficients themselves) or abandoning direct comparisons and analyzing the pattern of coefficient significance between the two models (Hoetker 2004).[1] Use of ratio of coefficients is straightforward and appealing, but it requires strong theoretical justification (ibid). Use of expected probabilities, as suggested by Long (2009), is another attractive option, but comparison of expected probabilities is complicated by nonlinear and non-additive associations between independent and dependent variables in logit models. The effects of the same magnitude would translate into different effects on the probabilities depending on the level of the initial probabilities.[2] Therefore, supplementing the coefficient estimates with changes in expected probabilities is informative, but comparisons in effects between groups with different initial probabilities should be done with caution. To summarize, at the time of writing "[t]here are no simple all-purpose solutions to the problems of interpretability and comparison of effect estimates from logistic regression" (Mood 2010, p. 80).

Hoetker (2004) indicates that one way to compare effects for two groups is to explore the direction and statistical significance of the coefficients. "If we model the two groups separately, the coefficients and standard errors are consistent within each group. The pattern of coefficient significance between the two models may provide some information." (ibid, p. 16). While this information is not sufficient to be able to compare the magnitude of differences in effects, this solution is better than spurious results. We combine this approach with exploring changes in probabilities when we manipulate one independent variable at a time. Because probabilities are unaffected by residual variation (Long 2009), we can use them to better understand effects for females and males and compare findings from multinomial logit and mixed logit models.

Models were estimated using the R mlogit package (Croissant 2015; Train and Croissant 2015).

Model Selection

Our analysis was based on three models—multinomial logit, nested logit, and mixed logit. Following estimation of multinomial logit full models, we built models that excluded investigative and enterprising categories to test for IIA assumption. The IIA hypotheses were rejected for both models (i.e., the model for female students and the model for male students). Comparison of models for female students yields χ^2 (31) = 5147.89, significant at the 0.01 alpha level. Comparison of models for male students yields χ^2 (31) = 316.26, significant at the 0.01 alpha level. These results indicate the need to explore models that relax IIA assumption. Two approaches that relax IIA assumption—nested logit and mixed logit—were explored.

[1] Williams (2009) suggests use of heterogeneous choice models for binary and ordinal dependent variables. While use of mixed logit models allows to incorporate heterogeneity of individuals in the model, we could not find a study devoted to a comparison of mixed logit coefficients across groups. Hence, we took a conservative approach by comparing expected probabilities and directions and significances of coefficients within groups.

[2] Partial derivatives—the logistic regression coefficients multiplied by a given probability and 1 minus a given probability—illustrate this point (Pampel 2000). "The effect [of b in terms of logged odds] will be at its maximum when P equals 0.5 since $0.5 \times 0.5 = 0.25$; $0.6 \times 0.4 = 0.24$, $0.7 \times 0.3 = 0.21$, and so on. The closer P comes to the ceiling or floor, the smaller the value of $P \times (1 - P)$, and the smaller the effect of a unit change in X has on the probability" (ibid, p. 25).

Our nested logit models (not presented here, but available from authors upon request) were based on two nests—a nest for investigative and realistic fields and a nest for social, artistic and enterprising fields. Based on the likelihood ratio test, the nested logit model for female students fits data better than the multinomial logit model for female students. The likelihood ratio test is $\chi^2(1) = 63.00$, significant at the 0.01 % alpha level. At the same time, the nested logit for male students does not fit data better than the multinomial logit model. The likelihood ratio test $\chi^2(1) = 0.19$, not significant at the 5 % alpha level. The concerning aspect for both nested models is that the log-sum coefficients are greater than 1—1.74 and 1.03 for females and males respectively. This means there is more substitution across nests than within nests. We concur with Train and Croissant (2015) that having a greater substitution across nests than within nests is not reasonable. With two nests—a nest for investigative and realistic fields and a nest for social, artistic and enterprising fields—we would expect that, if we excluded an investigative field from the choice set, the probability of choosing a realistic field would rise proportionately more than the probability of social, artistic, or enterprising fields. Similarly, if we excluded an artistic field from the choice set, the probability of choosing social and enterprising fields would increase proportionately more than the probability of realistic or investigative fields (Train 2009). Yet, based on the log-sum coefficients, the opposite is true. We also attempted other nest combinations, which led to improved model fit as evidenced by likelihood ratio tests. Yet none of our models led to the solution with greater substitution patterns within nests than between nests. Thus we failed to find nested models with reasonable substitution patterns for our choice of an academic discipline problem.

The mixed logit model in Table 3 extends the standard conditional logit model by allowing intercepts to be randomly distributed. Based on the likelihood ratio test, the mixed models fit data better than the multinomial logit model. For the models for male students the likelihood ratio test $\chi^2(4) = 467.31$, significant at the 0.01 % alpha level. For the models for female students the likelihood ratio test $\chi^2(4) = 156.4$, significant at the 0.01 % alpha level. One can observe that some of the random intercepts in our mixed logit model are negative, which intuitively does not seem to be right. However, as indicated by Hole (2007), the sign of estimated random effects in mixed logit models is irrelevant. Although in practice random effects in mixed logit might be negative, one can interpret them as being positive. Another observation is related to the fact that the random effects for only one category of the dependent variable—realistic—are statistically significant both for males and females.

In our calculations of expected probabilities, we followed Train and Croissant's (2015) illustration of calculation of expected probabilities in *R's mlogit* package. First we estimated the models with the actual values of all independent variables. The results of these models—referred to as fitted probabilities—are presented in Table 4 for female students and Table 5 for male students. After estimating models with actual values of independent variables, we change one variable at a time and calculate probabilities with the new values of this variable. Changes in expected probabilities between the fitted model and new models provide information on the magnitude of the effect of each independent variable. We present expected probabilities for multinomial logit models and mixed logit to gauge the consistency across these models when IIA assumption is violated. (The coefficients, standard errors, and indices of model fit for multinomial logit models are not presented here, but available from authors upon request).

Consistent across models, the *fitted* or *average expected* probabilities for females (see models in Table 4) are 0.33 for social disciplines, 0.13 for artistic disciplines, 0.10 for

Table 3 Mixed logit models of choice of academic majors with social disciplines as a reference group

	Model 1: females				Model 2: males			
	Artistic	Enterprising	Investigative	Realistic	Artistic	Enterprising	Investigative	Realistic
Intercept	−0.98 (1.16)	−24.2 (1.03)***	−15.67 (0.90)***	−48.99 (3.30)***	5.05 (1.7)**	−19.4 (1.46)***	−10.56 (1.49)***	−34.50 (3.76)***
Race, social class and highest degree expectation								
Caucasian	0.04 (0.15)	−0.16 (0.13)	−0.13 (0.09)	−2.55 (0.40)***	−0.04 (0.22)	0.27 (0.17)	0.06 (0.17)	−1.87 (0.47)***
First generation status	0.05 (0.15)	0.05 (0.13)	0.11 (0.09)	−0.05 (0.34)	0.25 (0.24)	−0.25 (0.19)	0.21 (0.18)	0.97 (0.45)*
Financial stress	−0.06 (0.05)	−0.14 (0.06)**	0.13 (0.04)***	0.40 (0.14)**	0.01 (0.09)	−0.08 (0.07)	0.11 (0.07)[a]	0.50 (0.18)**
Highest degree	−0.20 (0.09)*	−0.34 (0.13)*	1.18 (0.09)***	−1.51 (0.29)***	−0.25 (0.12)*	−0.57 (0.13)***	0.77 (0.13)***	−4.23 (0.48)***
Prior achievement and high school curriculum								
High school GPA	−0.17 (0.13)	−0.27 (0.13)*	0.25 (0.10)*	0.39 (0.36)	−0.60 (0.21)**	−0.37 (0.15)*	0.35 (0.16)*	0.82 (0.44)[a]
ACT verbal	0.04 (0.010)*	−0.03 (0.01)*	0.00 (0.01)	0.02 (0.04)	0.02 (0.02)	−0.04 (0.02)**	−0.02 (0.02)	−0.05 (0.04)
ACT quantitative	0.00 (0.02)	0.06 (0.02)***	0.04 (0.01)***	0.33 (0.06)***	−0.01 (0.03)	0.04 (0.02)*	0.02 (0.02)	0.48 (0.07)***
A passing grade in calculus	0.05 (0.11)	0.24 (0.10)*	0.32 (0.08)***	1.59 (0.34)***	0.49 (0.19)**	0.15 (0.14)	0.47 (0.14)***	3.42 (0.49)***
Years of english/literature	0.30 (0.24)	0.11 (0.21)	−0.10 (0.17)	−0.79 (0.60)	−0.30 (0.39)	0.16 (0.31)	−0.24 (0.31)	−1.68 (0.86)[a]
Years of math	−0.07 (0.12)	0.18 (0.12)	0.04 (0.09)	1.03 (0.31)***	−0.26 (0.21)	−0.04 (0.15)	−0.09 (0.15)	1.16 (0.35)**
Years of science	−0.31 (0.12)**	−0.36 (0.13)**	0.67 (0.09)***	0.59 (0.28)*	−0.02 (0.18)	−0.02 (0.14)	0.83 (0.15)***	1.36 (0.36)***
Years of history/social sciences	−0.03 (0.10)	0.03 (0.10)	−0.20 (0.07)**	0.07 (0.27)	−0.26 (0.17)	−0.32 (0.13)*	−0.55 (0.14)***	−1.13 (0.36)**
Years of foreign language	0.13 (0.05)**	0.12 (0.05)*	−0.10 (0.04)**	−0.34 (0.13)**	0.05 (0.08)	0.10 (0.06)	−0.11 (0.06)[a]	−0.89 (0.19)***

Table 3 continued

	Model 1: females				Model 2: males			
	Artistic	Enterprising	Investigative	Realistic	Artistic	Enterprising	Investigative	Realistic
Involvement in extra-curricular activities								
Performing or visual arts	0.19 (0.03)***	−0.04 (0.02)[a]	0.00 (0.02)	0.31 (0.07)***	0.15 (0.05)**	−0.13 (0.03)***	−0.06 (0.03)[a]	0.17 (0.09)*
Athletic teams	−0.06 (0.02)*	−0.02 (0.02)	0.01 (0.02)	0.06 (0.06)	−0.05 (0.04)	0.11 (0.03)**	0.05 (0.03)	−0.37 (0.09)***
Student government	−0.02 (0.03)	0.05 (0.03)[a]	−0.01 (0.02)	−0.07 (0.07)	−0.16 (0.06)*	−0.02 (0.04)	−0.09 (0.04)*	−0.35 (0.12)**
Publications	0.06 (0.03)*	0.01 (0.02)	−0.03 (0.02)	−0.06 (0.07)	0.15 (0.06)*	0.04 (0.05)	−0.05 (0.05)	−0.40 (0.15)**
Academic honor societies	0.00 (0.03)	0.06 (0.03)*	0.04 (0.02)[a]	0.15 (0.08)[a]	0.08 (0.06)	0.03 (0.04)	0.07 (0.04)[a]	0.24 (0.11)*
Academic clubs	0.04 (0.03)	0.02 (0.03)	0.04 (0.02)*	0.21 (0.07)**	−0.13 (0.06)*	−0.16 (0.04)***	−0.06 (0.04)	0.08 (0.11)
Vocational clubs	−0.05 (0.04)	0.11 (0.03)**	0.03 (0.02)	0.00 (0.09)	0.00 (0.07)	0.17 (0.05)***	0.05 (0.05)	0.16 (0.11)
Religious youth groups	−0.04 (0.02)[a]	−0.04 (0.02)[a]	−0.07 (0.02)***	−0.29 (0.07)***	0.03 (0.04)	0.06 (0.03)[a]	0.07 (0.03)*	0.03 (0.08)
Community service	−0.15 (0.04)***	−0.07 (0.03)*	−0.07 (0.03)**	−0.26 (0.10)**	0.02 (0.06)	−0.03 (0.04)	−0.06 (0.04)	−0.08 (0.11)
Self-efficacy for academic achievement								
Write clearly and effectively	0.17 (0.06)**	0.01 (0.05)	0.01 (0.04)	−0.20 (0.15)	0.15 (0.08)[a]	0.07 (0.06)	0.08 (0.06)	−0.37 (0.17)*
Speak clearly and effectively	0.00 (0.05)	0.02 (0.05)	−0.06 (0.04)	−0.26 (0.14)[a]	−0.10 (0.09)	−0.05 (0.07)	−0.26 (0.07)***	−0.82 (0.19)***
Think critically and analytically	0.12 (0.07)[a]	−0.14 (0.06)*	−0.13 (0.04)**	0.29 (0.17)[a]	−0.09 (0.11)	−0.24 (0.08)**	−0.26 (0.08)***	−0.47 (0.22)*
Analyze math	−0.17 (0.05)**	0.13 (0.04)**	0.23 (0.03)***	1.09 (0.18)***	0.03 (0.08)	0.32 (0.05)***	0.47 (0.06)***	2.73 (0.31)***
Use computing	−0.01 (0.04)	0.08 (0.04)[a]	0.07 (0.03)*	0.56 (0.13)***	0.13 (0.08)[a]	0.08 (0.05)	0.02 (0.06)	1.92 (0.23)***

Table 3 continued

	Model 1: females				Model 2: males			
	Artistic	Enterprising	Investigative	Realistic	Artistic	Enterprising	Investigative	Realistic
Work effectively with others	−0.09 (0.06)	−0.02 (0.06)	−0.05 (0.04)	−0.73 (0.17)***	−0.10 (0.10)	−0.02 (0.07)	0.01 (0.07)	−0.97 (0.20)***
Learn effectively on your own	−0.02 (0.06)	−0.08 (0.05)	−0.05 (0.04)	−0.28 (0.15)[a]	−0.01 (0.09)	−0.07 (0.07)	−0.06 (0.07)	−0.75 (0.19)***
Random effect	0.36 (1.17)	−0.61 (0.98)	−0.18 (0.85)	5.86 (0.76)***	−0.40 (1.25)	−0.35 (0.83)	0.34 (0.74)	9.88 (1.04)***
ln (salary)	65.29 (0.28)***				62.39 (0.17)***			
Log-likelihood	−10,593				−9671.5			
McFadden R^2	0.27604				0.21365			
Likelihood ratio test	8077.9				5255.4			

[a] Significant at the 5 % alpha level

* Significant at the 1 % alpha level

** Significant at the 0.1 % alpha level

*** Significant at the 0.01 % alpha level

Table 4 Effects of independent variables on choices of females: multinomial logit versus mixed logit models of choice of academic majors

	Multinomial logit					Mixed logit				
	Social	Artistic	Enterprising	Investigative	Realistic	Social	Artistic	Enterprising	Investigative	Realistic
Fitted probabilities	0.33	0.13	0.10	0.32	0.11	0.33	0.13	0.10	0.32	0.11
Race, social class and highest degree expectation										
Caucasian*	0.37	0.14	0.10	0.33	0.06	0.37	0.14	0.10	0.33	0.06
First generation status*	0.32	0.13	0.10	0.34	0.11	0.32	0.13	0.10	0.33	0.11
Financial stress**	0.33	0.13	0.09	0.34	0.12	0.33	0.12	0.09	0.34	0.12
Highest degree*	0.25	0.10	0.05	0.53	0.06	0.26	0.10	0.05	0.52	0.07
Prior achievement and high school curriculum										
High school GPA*	0.32	0.12	0.08	0.36	0.12	0.32	0.12	0.08	0.36	0.12
ACT verbal*	0.33	0.13	0.10	0.32	0.11	0.33	0.13	0.10	0.32	0.11
ACT quantitative*	0.33	0.13	0.11	0.32	0.12	0.33	0.13	0.11	0.32	0.12
A passing grade in calculus*	0.29	0.12	0.10	0.33	0.15	0.29	0.12	0.10	0.34	0.15
Years of english/literature*	0.33	0.15	0.12	0.30	0.09	0.33	0.15	0.12	0.31	0.09
Years of math*	0.32	0.12	0.11	0.31	0.14	0.32	0.12	0.11	0.31	0.14
Years of science*	0.29	0.10	0.06	0.43	0.12	0.29	0.10	0.06	0.43	0.12
Years of history/social sciences*	0.35	0.13	0.11	0.29	0.11	0.35	0.13	0.11	0.29	0.12
Years of foreign language*	0.34	0.14	0.12	0.30	0.10	0.34	0.14	0.12	0.30	0.10
Involvement in extra-curricular activities										
Performing or visual arts*	0.33	0.14	0.10	0.31	0.12	0.33	0.14	0.10	0.32	0.12
Athletic teams*	0.34	0.13	0.10	0.32	0.11	0.33	0.13	0.10	0.32	0.11
Student government*	0.33	0.13	0.11	0.32	0.11	0.33	0.13	0.11	0.32	0.11
Publications*	0.34	0.14	0.10	0.32	0.11	0.33	0.13	0.10	0.32	0.11
Academic honor societies*	0.33	0.13	0.11	0.32	0.11	0.33	0.13	0.11	0.32	0.12
Academic clubs*	0.33	0.13	0.10	0.32	0.12	0.33	0.13	0.10	0.32	0.12
Vocational clubs*	0.33	0.13	0.11	0.32	0.11	0.33	0.12	0.11	0.32	0.11

Table 4 continued

	Multinomial logit					Mixed logit				
	Social	Artistic	Enterprising	Investigative	Realistic	Social	Artistic	Enterprising	Investigative	Realistic
Religious youth groups*	0.35	0.13	0.10	0.31	0.10	0.34	0.13	0.10	0.32	0.11
Community service*	0.35	0.12	0.10	0.32	0.11	0.35	0.12	0.10	0.32	0.11
Self-efficacy for academic achievement										
Write clearly and effectively*	0.33	0.14	0.10	0.32	0.10	0.33	0.14	0.10	0.32	0.11
Speak clearly and effectively*	0.34	0.13	0.11	0.31	0.10	0.34	0.13	0.11	0.32	0.11
Think critically and analytically*	0.34	0.14	0.09	0.30	0.12	0.34	0.14	0.09	0.30	0.12
Analyze math*	0.31	0.11	0.10	0.34	0.14	0.31	0.11	0.10	0.34	0.14
Use computing*	0.32	0.13	0.11	0.32	0.12	0.32	0.13	0.11	0.32	0.13
Work effectively with others*	0.35	0.13	0.11	0.32	0.09	0.35	0.13	0.11	0.33	0.09
Learn effectively on your own*	0.33	0.13	0.10	0.32	0.11	0.34	0.13	0.10	0.32	0.11
Salary for social disciplines***	0.55	0.10	0.06	0.22	0.08	0.62	0.09	0.04	0.16	0.09
Salary for artistic disciplines***	0.29	0.24	0.09	0.28	0.10	0.26	0.30	0.07	0.26	0.10
Salary for enterprising disciplines***	0.30	0.12	0.20	0.29	0.10	0.28	0.11	0.23	0.27	0.11
Salary for investigative disciplines***	0.25	0.11	0.07	0.49	0.07	0.22	0.10	0.06	0.54	0.09
Salary for realistic disciplines***	0.31	0.12	0.09	0.29	0.18	0.32	0.13	0.10	0.31	0.14

* Average expected probabilities are calculated for a one point increase in the independent variable

** Average expected probabilities are calculated for a one standard deviation increase in the independent variable

*** ln (salary) → ln (salary + $1000)

Table 5 Effects of independent variables on choices of males: multinomial logit versus mixed logit models of choice of academic majors

	Multinomial logit					Mixed logit				
	Social	Artistic	Enterprising	Investigative	Realistic	Social	Artistic	Enterprising	Investigative	Realistic
Fitted probabilities	0.08	0.06	0.19	0.22	0.45	0.08	0.06	0.19	0.22	0.45
Race, social class and highest degree expectation										
Caucasian*	0.08	0.05	0.25	0.23	0.38	0.08	0.06	0.24	0.24	0.39
First generation status*	0.08	0.06	0.14	0.23	0.48	0.07	0.07	0.14	0.24	0.48
Financial stress**	0.08	0.06	0.17	0.22	0.46	0.07	0.06	0.17	0.23	0.47
Highest degree*	0.08	0.05	0.12	0.46	0.29	0.08	0.05	0.13	0.42	0.31
Prior achievement and high school curriculum										
High school GPA*	0.08	0.04	0.14	0.27	0.47	0.08	0.04	0.14	0.27	0.48
ACT verbal*	0.08	0.06	0.19	0.22	0.45	0.08	0.06	0.19	0.22	0.45
ACT quantitative*	0.08	0.06	0.19	0.21	0.47	0.07	0.06	0.19	0.21	0.47
A passing grade in calculus*	0.05	0.06	0.14	0.21	0.54	0.05	0.06	0.14	0.20	0.55
Years of english/literature*	0.09	0.05	0.25	0.21	0.39	0.09	0.05	0.25	0.21	0.40
Years of math*	0.08	0.05	0.18	0.19	0.49	0.08	0.05	0.18	0.20	0.49
Years of science*	0.06	0.04	0.13	0.30	0.47	0.06	0.04	0.13	0.28	0.48
Years of history/social sciences*	0.11	0.06	0.21	0.20	0.42	0.10	0.06	0.21	0.20	0.43
Years of foreign language*	0.08	0.06	0.22	0.21	0.42	0.08	0.06	0.22	0.21	0.42
Involvement in extra-curricular activities										
Performing or visual arts*	0.08	0.07	0.18	0.21	0.46	0.08	0.07	0.18	0.22	0.46
Athletic teams*	0.08	0.05	0.21	0.22	0.43	0.07	0.06	0.21	0.22	0.44
Student government*	0.09	0.05	0.20	0.22	0.44	0.08	0.06	0.20	0.22	0.44
Publications*	0.08	0.07	0.20	0.21	0.44	0.08	0.07	0.20	0.22	0.44
Academic honor societies*	0.08	0.06	0.19	0.22	0.46	0.07	0.06	0.19	0.22	0.46
Academic clubs*	0.09	0.06	0.18	0.22	0.46	0.08	0.06	0.18	0.22	0.46
Vocational clubs*	0.07	0.05	0.21	0.21	0.45	0.07	0.06	0.20	0.21	0.46

Table 5 continued

	Multinomial logit					Mixed logit				
	Social	Artistic	Enterprising	Investigative	Realistic	Social	Artistic	Enterprising	Investigative	Realistic
Religious youth groups*	0.08	0.06	0.19	0.22	0.45	0.07	0.06	0.19	0.22	0.45
Community service*	0.08	0.06	0.19	0.21	0.45	0.08	0.06	0.19	0.22	0.45
Self-efficacy for academic achievement										
Write clearly and effectively*	0.08	0.06	0.19	0.22	0.45	0.07	0.06	0.20	0.23	0.44
Speak clearly and effectively*	0.09	0.06	0.21	0.21	0.43	0.09	0.06	0.21	0.21	0.43
Think critically and analytically*	0.09	0.06	0.19	0.21	0.44	0.09	0.06	0.19	0.21	0.44
Analyze math*	0.06	0.04	0.17	0.21	0.52	0.05	0.04	0.16	0.21	0.53
Use computing*	0.07	0.05	0.17	0.19	0.52	0.07	0.06	0.17	0.19	0.51
Work effectively with others*	0.09	0.06	0.20	0.23	0.42	0.08	0.06	0.20	0.24	0.42
Learn effectively on your own*	0.09	0.06	0.20	0.23	0.43	0.08	0.06	0.20	0.23	0.43
Salary for social fields***	0.14	0.05	0.18	0.20	0.43	0.21	0.04	0.14	0.16	0.44
Salary for artistic fields***	0.08	0.09	0.18	0.21	0.44	0.06	0.16	0.15	0.18	0.44
Salary for enterprising fields***	0.07	0.05	0.26	0.20	0.41	0.05	0.04	0.32	0.15	0.44
Salary for investigative fields***	0.07	0.05	0.17	0.31	0.40	0.05	0.04	0.11	0.38	0.43
Salary for realistic fields***	0.07	0.05	0.16	0.18	0.53	0.07	0.06	0.18	0.21	0.49

* Average expected probabilities are calculated for a one point increase in the independent variable

** Average expected probabilities are calculated for a one standard deviation increase in the independent variable

*** ln (salary) → ln (salary + $1000)

enterprising disciplines, 0.32 for investigative disciplines, and 0.11 for realistic disciplines. The fitted probabilities for males (see models in Table 5) are 0.08 for social disciplines, 0.06 for artistic disciplines, 0.19 for enterprising disciplines, 0.22 for investigative disciplines, and 0.45 for realistic disciplines. Fitted probabilities also match the observed shares of students who chose each of these disciplines, see Table 1. Next, we will discuss our findings and consistency of effects based on multinomial logit and mixed logit models.

Findings

Based on the models in Table 3, salary expectation has a significant positive effect on choice of a discipline by females and males. To translate these effects into expected probabilities, we estimated the effects of $1000 increase in expected earnings on the probabilities of choosing each of the disciplines. The effects of expected salary increases for females and males are presented in Tables 4 and 5. For example, the increase in expected salary for social disciplines leads to the 0.29 and 0.13 increases in expected probabilities of choosing a social discipline by females and males, see mixed logit models in Tables 4 and 5. The effects of salary expectations are generally greater based on mixed logit compared to multinomial logit models. Overall, substantial effects of salary expectations on choice of a discipline by females challenge our initial speculation that females overlook wages in their career choice.

Models in Table 3 indicate statistically significant associations between ethnicity and choice of realistic disciplines over social disciplines for both males and females. The expected probabilities in Tables 4 and 5 indicate that Caucasian students are less likely to choose a realistic discipline. The effect of ethnicity is consistent across models and holds for both males and females. As illustrated in Fig. 1, non-Caucasian female students are 1.70 (0.17/0.10) times as likely to choose a realistic discipline as Caucasian female students. Compared to Caucasian males, non-Caucasian male students have 1.16 (0.51/0.44) times the likelihood of choosing a realistic field. Due to a small number of racial and ethnic minority students at a study institution, we did not analyze different racial and ethnic groups separately. Further studies are needed to describe these differences in choices by Caucasian and non-Caucasian students.

Two variables—financial stress and first generation status—were used to measure the family socio-economic status. Based models in Table 3, the effect of first generation status

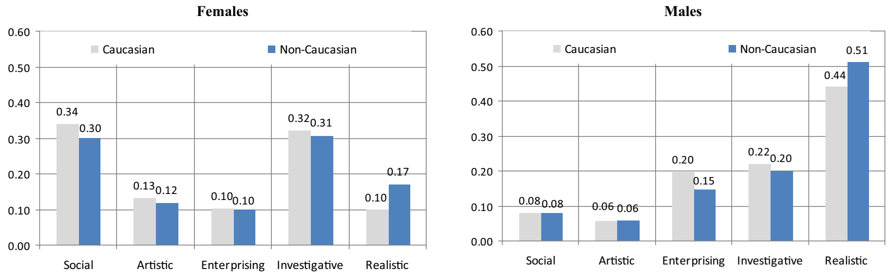

Fig. 1 Expected probabilities of choosing academic disciplines by Caucasian and non-Caucasian students. *Note* Expected probabilities in Fig. 1 are based on multinomial logit models

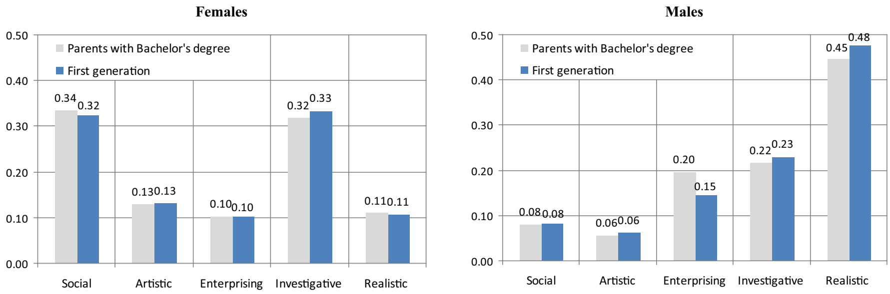

Fig. 2 Expected probabilities of choosing academic disciplines by first generation students and students whose parent completed a Bachelor's degree. *Note* Expected probabilities in Fig. 2 are based on multinomial logit models

on choice of realistic disciplines over social disciplines by males is significant at the 1 % alpha level. Consistently across multinomial logit and mixed logit models, first generation status is associated with greater likelihood of choosing realistic disciplines by males. The expected probabilities by first generation status are also provided in Fig. 2. For males the expected probability of choosing a realistic discipline is 0.45 if their parents completed a four-year degree and 0.48 if their parents did not complete a four-year degree. Yet the expected probability of choosing a realistic discipline is the same for first-generation and non-first-generation females.

Financial stress is associated with greater odds of choosing investigative and realistic disciplines over social disciplines by males and females and greater odds of choosing social disciplines over enterprising disciplines by females. In terms of predicted probabilities (see mixed logit models in Tables 4, 5), with one standard deviation increase in financial stress, the likelihood of choosing an investigative field goes up from 0.32 to 0.34 for females and from 0.22 to 0.23 for males, while the likelihood of choosing a realistic field goes up from 0.11 to 0.12 for females and from 0.45 to 0.47 for males. The effect of financial stress on the odds of choosing investigative and realistic fields is consistent with several prior studies (Goyette and Mullen 2006; Ma 2009): students who expect greater financial stress are more likely to choose technical and life sciences fields or "safe bet" majors with more favorable career outlooks and greater salary expectations. Unlike Davies and Guppy (1997), we did not use expected earnings as the dependent variable. The latter approach might be more appropriate for a study of the association between choice of a lucrative career and socio-economic status.

Consistent with Eide and Waehrer (1998), student choice of major is strongly associated with the degree they intend to complete. In our study degree expectations are measured on a scale from 1-Associate's to 4-Doctoral. And, a one-point increase in degree expectations leads to an increase in the probability of choosing an investigative field from 0.22 to 0.42 for males and from 0.32 to 0.52 for females, see results for mixed logit models in Tables 4 and 5. Changes in expected probabilities of choosing investigative fields with one-point increase in degree expectations based on multinomial logit models are somewhat greater, but consistent with changes based on mixed logit. The higher the level of degree expected, the greater the odds of picking investigative disciplines by females and males.

Models in Table 3 indicate a statistically significant association between ACT quantitative (ACTQ) and having a passing grade in calculus, on the one hand, and choice of a discipline, on the other hand. Consistently across multinomial logit and mixed logit

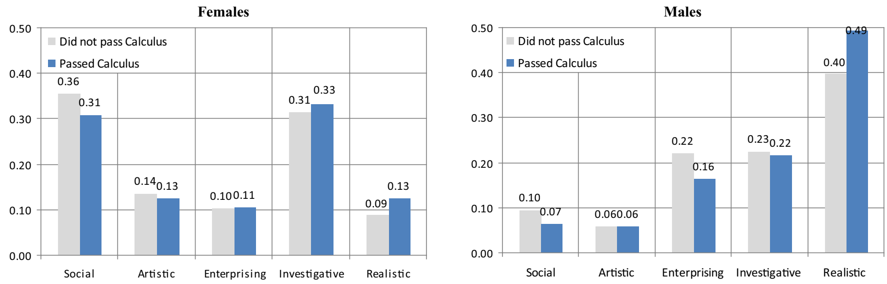

Fig. 3 Expected probabilities of choosing academic disciplines by students who did and did not earn a passing grade in Calculus in high school. *Note* Expected probabilities in Fig. 3 are based on multinomial logit models

models, passing grades in calculus are positively associated with the likelihood of choosing realistic and investigative fields by females and realistic fields by males (see expected probabilities in Tables 4, 5). Females who pass calculus have 0.04 greater likelihood of selecting a realistic discipline, see Fig. 3. Passing calculus in high school by a female also leads to a 0.02 increase in the likelihood of going into an investigative field. Males who pass calculus have 0.09 greater likelihood of choosing realistic fields, see Fig. 3.

Greater ACTQ scores are also associated with greater probabilities of choosing realistic disciplines by males and females. This effect holds for both multinomial logit and mixed logit models, see Table 4. According to Fig. 4, females with ACTQ 30 are 1.40 (0.14/0.10) times as likely to choose realistic fields as females with ACTQ 25. Males with ACTQ 30 are 1.19 (0.50/0.42) times as likely to choose realistic fields as males with ACTQ 25.

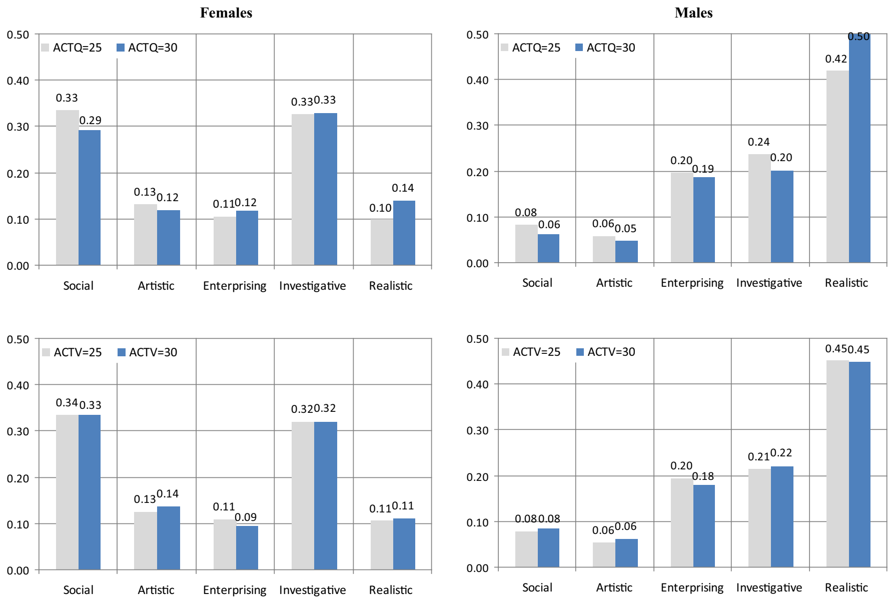

Fig. 4 Expected probabilities of choosing academic disciplines by ACT Scores. *Note* Expected probabilities in Fig. 4 are based on multinomial logit models

Interestingly, for females the probability of choosing investigative fields does not change as the ACTQ goes up. At the same time, the probability of choosing an investigative field is 0.04 lower for males with ACTQ 30 than for males with ACTQ 25, see Fig. 4.

Effects of ACT verbal (ACTV) on choice of enterprising disciplines over social disciplines are negative and statistically significant for both females and males (see Table 3). Consistently across models, these effects translate into 0.02 decrease in the probabilities of selecting enterprising fields as ACTV increases from 25 to 30 (see Fig. 4 and Tables 4, 5 to compare consistency across models).

Based on models in Table 3, high school GPA is positively associated with the choices of investigative and realistic fields over social fields and negatively associated with the choices of enterprising and artistic fields over social fields. Some of these effects are not statistically significant for females, however. A point increase in high school GPA is associated with 0.04 and 0.05 increases in expected probabilities of choosing investigative fields by females and males (see expected probabilities for mixed logit models in Tables 4, 5). Probabilities of choosing realistic fields by females and males increase by 0.01 and 0.03. Changes in expected probabilities are consistent across multinomial logit and mixed logit models.

Years of mathematics in high school are positively correlated with choosing realistic fields, see Table 3. Adding one year of mathematics in high school translates into increases in the probabilities of choosing realistic fields from 0.11 to 0.14 for females and from 0.45 to 0.49 for males. Not surprisingly, years of science are associated with greater odds of choosing investigative and realistic fields. With one additional year of science in high school, the expected probability of choosing investigative fields changes from 0.32 to 0.43 for females and from 0.22 to 0.28 for males, while the expected probability of choosing realistic field changes from 0.11 to 0.12 for females and from 0.45 to 0.48 for males. The expected probabilities of choosing social, artistic, and enterprising fields go down as the number of years of science in high school goes up. Based on models in Table 3, taking an additional year of history or social science in high school is associated with a higher likelihood of choosing social fields over enterprising, investigative, and realistic disciplines by males and a higher likelihood of choosing social fields over investigative fields for females. An additional year of history or social science translates into 0.02 increase in the probability of choosing a social field by females and males (see mixed logit models in Tables 4, 5). Students who took more years of foreign language are less likely to enter investigative and realistic fields (see Table 3). Each additional year of foreign language translates into 0.02 and 0.01 decreases in probabilities of selecting investigative disciplines and in 0.01 and 0.03 decreases in probabilities of selecting realistic disciplines by females and males. These findings suggest that high school curriculum is an important factor in choice of a discipline both by males and females. Next, we will explore the effects of involvement in different extra-curricular activities on choice of an academic major.

On a scale from 1-not involved to 6-highly involved, students were asked to evaluate their involvement in extra-curricular activities during high school years. Based on models in Table 3, involvement in performing or visual arts is associated with greater chances of selecting artistic and realistic fields by females and males. This translates into 0.01 increases in expected probabilities with a point increase of involvement in performing or visual arts (see Tables 4, 5). The model in Table 3 indicates that females who report greater involvement in athletic teams are more likely to choose a social discipline over an artistic discipline, but this effect does not translate into a visible change in expected probability of choosing artistic fields with a point change in female involvement in athletic teams (see Table 4). For males a point increase in involvement in athletic teams is associated with 0.02 greater likelihood of choosing enterprising fields (see Table 5). A point increase in involvement in student government is associated with 0.01 increases in the probabilities of selecting enterprising fields by females and males. A point increase in involvement in publications does not lead to visible changes in expected probabilities for females (see mixed logit results in Table 4), but is associated with 0.01 increase in expected probabilities of choosing artistic and enterprising fields by males (see mixed logit results in Table 5). One-point increase in involvement in both academic honor societies and academic clubs leads to a 0.01 increase in expected probabilities of selecting realistic fields by males and females. A point increase in involvement in vocational clubs leads to 0.01 increases in probabilities of choosing enterprising disciplines by females and males. Involvement in religious youth groups and community service or volunteer work is positively associated with greater odds of picking social disciplines by female students. This finding does not hold for male students. It is remarkable that females also report greater involvement in religious youth groups, community service, and volunteer work (see Table 1). While prior involvements in extra-curricular activities have statistically significant effects on choice of discipline by females and males, these effects are rather small in terms of changes in expected probabilities. Some types of involvement—i.e., religious youth groups or community service—have visible effects on choices of females but not on choices of males.

Student self-ratings of their preparation to perform different academic tasks are measured on a scale from 1-not at all prepared to 6-very prepared. Models in Table 3 indicate that student self-ratings of their preparation to write clearly and effectively have significant positive effects on expected probabilities of choosing artistic fields over social fields for females and males and a negative effect on the expected probability of selecting a realistic field by males. A point increase in self-ratings of preparation to write is associated with the 0.01 increase in expected probability of choosing artistic disciplines by females (see Table 4) and 0.01 decrease in expected probability of choosing realistic disciplines by males (see Table 5). Students who are confident in their ability to speak clearly and effectively prefer social and enterprising disciplines over investigative and realistic disciplines. A point increase in self-rating of preparation to think critically and analytically is associated with slight increases in probabilities of choosing social, artistic and realistic fields for females and increase in the probability of choosing social disciplines by males. The effect of self-rating of preparation to analyze math and quantitative problems is stronger than the effects of self-ratings of ability to write, speak or think critically. For instance, a one-point increase is associated with 0.03 and 0.08 increases in the expected probability of selecting realistic fields by females and males. Confidence in one's ability to use computing or information technology is associated with greater odds of choosing realistic over social fields (see models in Table 3). Finally, one's greater confidence in ability to work effectively with others and learn effectively on one's own is associated with lower odds of choosing a realistic discipline (see Table 3). To summarize, controlling for high school performance, test scores, high school curriculum and prior involvement in extra-curricular activities, confidence in one's abilities and preparation has a significant effect on choice of major. Self-ratings of preparation to analyze math and use computing are most consequential for the choice of major. In most instances, the direction and significance of effects of self-ratings on selection of disciplines are consistent for males and females. At the same time, females tend to rate themselves lower in areas that have a strong positive effect on the choice of enterprising, realistic and investigative fields—e.g., ability to analyze math and quantitative problems—and higher in areas that have a positive effect on the choice of artistic fields—writing—and social fields—speaking.

Limitations

Our study has several important limitations. First, the analysis is based on data from a single moderately large research institution. Major choices and determinants of these choices might vary across institutions; and findings presented here might not apply to other institutions and institution types.

Our comparisons across different models—multinomial logit, nested logit, and mixed logit—and across models for different groups—males and females—is limited to stating significant positive and negative coefficients and analysis of changes in expected probabilities. Ideally, we would compare regression coefficients across genders and accompany these comparisons with statistical tests for the significance of differences. However, while this could be done in linear regression, similar comparisons for logit and probit models have been found misleading. While we took a rather conservative approach, this approach is superior to reporting spurious results.

Next, we did not go beyond random intercepts in our mixed model. Exploring the hypotheses of random effects of each of our independent variables might lead to an improved model fit.

Finally, we included a discipline-specific salary expectation variable—natural logarithm of expected salary—in our models. Based on this approach, the magnitude of the effect of salary expectations is very substantial and warrants additional exploration of different representations of salary expectations on choice of majors.

Conclusions and Implications

Our analysis of college major choices of entering first-time freshmen suggests that, compared to males, female students are more likely to self-select into social (Education, Psychology, Social Sciences, Health Sciences, and Nursing), artistic (Arts, Communications, English, Foreign Language, Music, and Theatre) and investigative (Agriculture, Biosystems Engineering, Sciences, and Mathematics) disciplines. Males are more likely to choose realistic (Architecture, Building Science, Engineering, and Computer Science) and enterprising (Business, Industrial and Systems Engineering, and Economics) disciplines. With one exception of a higher proportion of females in investigative disciplines, our findings align with prior studies (Fouad 2002; Betz and Gwilliam 2002).

Consistent with several prior studies (e.g., Goyette and Mullen 2006; Ma 2009), lower socio-economic status is associated with a slightly greater probability of choice of more lucrative careers and careers with more favorable outlooks. With respect to gender differences, we find that first-generation status leads to a greater likelihood of choosing realistic fields for males, but not for females. Financial stress has a greater effect on selecting investigative disciplines than realistic disciplines by females. The opposite is true for males. Overall, the association between lower socio-economic status and choice of sciences and engineering is a "glimpse of hope that students and their families take advantage of college major choice to ameliorate the constraining effects of low family SES" (Ma 2009, p. 227). However, further studies are needed to review low SES student persistence in these majors and their actual job placements, since better-off families might be able to help their children locate a more prosperous career.

"[M]any students choose to major in fields never intending to terminate their education with an undergraduate degree, but rather they intend to enroll in professional or academic graduate programs" (Eide and Waehrer 1998, p. 73). Consistent with findings of Eide and Waehrer (1998), highest degree intentions are strongly associated with the choice of investigative fields over other fields. However, in contrast to findings of Eide and Waehrer (1998), the effects of the highest degree intensions are not associated with greater expected probabilities of choosing liberal arts (social or artistic) disciplines. The expected probability of choosing an artistic or social discipline by females goes down as the degree expectation goes up. For males the expected probability of choosing an artistic or social discipline remains about the same as the highest degree expectation goes up.

Not surprisingly, student high school academic performance and curriculum are strongly associated with their choice of major. Consistent with Paglin and Rufolo (1990), quantitative ability is one of the most important factors in choice of major. For example, higher ACT quantitative scores are associated with greater likelihood of selecting realistic fields by both males and females. Interestingly, female students who had passing grades in calculus are more likely to enter either investigative or realistic fields, while males who passed calculus are more likely to enter realistic fields. One should also note here that fewer females report a passing grade in calculus; while males have higher ACT quantitative scores.

Turner and Bowen (1999) suggested that differences in choice of major between men and women reflect the effects of pre-collegiate preparation. Our study aligns with their observation. Students with more years of science are more likely to choose investigative and realistic fields. Years of history and social sciences are positively associated with the choice of social fields. Students with more years of mathematics are more likely to choose realistic fields. Years of foreign language are associated with greater probability of choosing an enterprising field. While the effects of years of mathematics and foreign language are consistent across gender, females report fewer years of mathematics and more years of foreign language completed in high school.

Involvement in religious youth groups and community service has a significant positive association with the choice of a social field over all other fields for female students. No such association was found for male students. Furthermore, our descriptive analysis reveals that females report greater involvement in these activities. Females' selflessness appears to be a significant factor in their choice of social fields, which aligns with Ma's (2009) observation of females' intrinsic and altruistic motivation and males' extrinsic motivation. Overall, the effect of involvement in extra-curricular activities on choice of major is statistically significant yet not always substantial in terms of expected probabilities.

Controlling for other characteristics, student confidence in their preparation to perform different academic tasks—writing, speaking, critical thinking, analyzing math, using computing, working effectively with others and on one's own—has a significant association with choice of major. The magnitude of these effects is the greatest for the ability to analyze math and use computing thus supporting prior findings (Betz and Hackett 1983). Our study also supports the claim that the mathematics self-efficacy is stronger for males than for females (ibid). Overall, consistent with Smart et al. (2000), our study provides empirical evidence of association between student self-rated abilities and chosen academic fields. Because of weaker ratings of one's ability to analyze quantitative problems among female students, college counselors need to take into account that encouragement and reassurance might be more important to females than to males.

Given the alignment between prior involvement in extra-curricular activities, self-rated abilities, and a major choice, we concur with Smart et al. (2000) suggestion of an alternative curricular that—unlike the typical curricular pattern of general-education courses in the first year of studies—allows students to take courses in subjects that are more congruent with their interests and abilities. Since high school experiences and expectations of females are different from those of males, further studies are needed to better understand the reasons behind less interest in mathematics and greater interest in volunteering and religious groups among females. Meanwhile, it is not clear whether counselors should encourage female participation in engineering or business careers as these careers might not always fit their interests and aspirations. Further studies are needed to understand the association between interests and aspirations, on the one hand, and retention in major or satisfaction with the chosen career path, on the other hand.

While our study sheds light on some factors that affect choices of academic disciplines by males and females, it left many questions unanswered. "[T]he career development of women … is demonstrably more complex due to a socialization process that has emphasized the dichotomy of work and family" (Fitzgerald and Weitzman 1992, p. 125, cited by Patton and McMahon 1999, p. 91). Further studies are needed to explore the gender role socialization, occupational stereotypes, or gender-biased counseling (Betz 1994) on choice of an academic discipline by female students. For example, is women's choice of teaching or nursing related to their desire to marry upward and secure status as suggested by Psathas (1968)?

From the methodological point of view, our study explores the effect of violation of independence of irrelevant alternatives (IIA) assumption on the estimates of multinomial logit model and draws attention to the possible problem—a greater substitution across nests than within nests—that can be encountered while using a nested model. Despite the violation of IIA assumption, the findings from multinomial logit and mixed logit are generally consistent. Our study also demonstrates the importance of exploring changes in expected probabilities in addition to statistically significant effects. In some instances, manipulating independent variables that have statistically significant effects do not lead to visible changes in expected probabilities.

Acknowledgments The authors are grateful to Drew Clark, Kyrylo Kobzyev, Robert Toutkoushian, and two anonymous reviewers for their valuable comments and suggestions. All errors remain our own.

References

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219–245.

Adelman, C. (2003). *Postsecondary attainment, attendance, curriculum, and performance: Selected results from the NELS: 88/2000 Postsecondary Education Transcript Study (PETS), 2000*. Washington, DC: U.S. Department of Education.

Allison, P. D. (1999). Comparing logit and probit coefficients across groups. *Sociological Methods and Research, 28*(2), 186–208.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics, 121*, 343–375.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs: Prentice Hall.

Bandura, A. (1995). *Self-efficacy in changing societies*. Cambridge: Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman and Company.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Greenwich: Information Age Publishing.

Berger, M. C. (1988). Predicted future earnings and choice of college major. *Industrial and Labor Relations Review, 41*, 418–429.

Betz, N. E. (1994). Basic issues and concepts in career counseling for women. In W. N. Walsh & S. H. Osipow (Eds.), *Career counseling for women* (pp. 1–41). Hillsdale: Erlbaum.

Betz, N. E., & Gwilliam, L. R. (2002). The utility of measures of self-efficacy for the Holland themes in African American and European American college students. *Journal of Career Assessment, 10*(3), 283–300.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*(3), 329–345.

Betz, N. E., Harmon, L. W., & Borgen, F. H. (1996). The relationship of self-efficacy for the Holland themes to gender, occupational group membership, and vocational interests. *Journal of Counseling Psychology, 43*(1), 90–98.

Bubany, S. T., & Hansen, J. I. (2010). Ability self-estimates and self-efficacy: Meaningfully distinct? *Measurement and Evaluation in Counseling and Development, 43*(3), 168–187.

Bureau of Labor Statistics (2014–2015). *Occupational outlook handbook*. Retrieved March 15, 2015, from http://www.bls.gov/ooh.

Cole, J. (2012). Using BCSSE and NSSE data to investigate first-year student financial stress and engagement. NSSE Webinar. Retrieved June 12, 2015, from http://www.nsse.indiana.edu/webinars/.

Croissant, Y. (2015). Estimation of multinomial logit models in R: The mlogit Packages. Retrieved March 19, 2015, from http://www.cran.r-project.org/web/packages/mlogit/vignettes/mlogit.pdf.

Davies, S., & Guppy, N. (1997). Fields of study, college selectivity, and student inequalities in higher education. *Social Forces, 75*(4), 1417–1438.

DiPrete, T. A.,& Buchmann, C. (2013). *The rise of women: The growing gender gap in education and what it means for American schools*. Russell Sage Foundation. Kindle Edition.

Dow, J. K., & Endersby, J. W. (2004). Multinomial probit and multinomial logit: A comparison of choice models for voting research. *Electoral Studies, 23*, 107–122.

Eide, E., & Waehrer, G. (1998). The role of the option value of college attendance in college major choice. *Economics of Education Review, 17*(1), 73–82.

Fitzgerald, L. F., & Weitzman, L. (1992). Women's career development: Theory and practice from a feminist perspective. In Z. Leibowitz & D. Lea (Eds.), *Adult career development: concepts, issues, and practices* (pp. 125–157). Alexandria: National Career Development Association.

Fouad, N. A. (2002). Cross-cultural differences in vocational interests: Between-group differences on the strong interest inventory. *Journal of Counseling Psychology, 49*(3), 283–289.

Görlitz, K., & Gravert, K. (2015). *The Effects of a High School Curriculum Reform on University Enrollment and the Choice of College Major*. The Institute for the Study of Labor (IZA) Discussion Paper No.8983. Retrieved January 5, 2016, from http://www.ftp.iza.org/dp8983.pdf.

Gottfredson, L. S. (1981). Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology, 28*(6), 545–579.

Goyette, K., & Mullen, A. (2006). Who studies the Arts and Sciences? social background and the choice and consequences of undergraduate field of study. *The Journal of Higher Education, 77*(3), 497–538.

Green, K. C. (1992). *After the boom: Management majors in the 1990s*. New York: McGraw-Hill.

Hoetker, G. P. (2004). Confounded coefficients: Extending recent advances in the accurate comparison of logit and probit coefficients across groups. Available at SSRN 609104.

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal, 7*(3), 388–401.

Holland, J. (1966). *The psychology of vocational choice*. Waltham: Blaisdell.

Holland, J. (1973). *Making vocational choices: A theory of careers*. Englewood Cliffs: Prentice Hall.

Holland, J. (1987). Current status of Holland's theory of careers: Another perspective. *Career Development Quarterly, 36*, 24–30.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa: Psychological Assessment Resources Inc.

Holland, J. L., & Lutz, S.W. (1967). *Predicting a Student's Vocational Choice*. American College Testing Program. Retrieved March 17, 2015, from http://www.act.org/research/researchers/reports/pdf/ACT_RR18.pdf.

Huss, M. T., Randall, B. A., Patry, M., Davis, S. F., & Hansen, D. J. (2002). Factors influencing self-rated preparedness for graduate school: A survey of graduate students. *Teaching of Psychology, 29*(4), 275–281.

Jacobs, J. A. (1986). The sex-segregation of fields of study: Trends during the college years. *Journal of Higher Education, 57*(2), 134–154.

James, E., Nabeel, A., Conaty, J., & To, D. (1989). College quality and future earnings: where should you send your child to college? *American Economic Review, 79*, 247–252.
Jones, W. A. (2011). Variation among academic disciplines: An update on analytical frameworks and research. *Journal of the Professoriate, 1*(6), 9–27.
Krumboltz, J. D. (1996). A learning theory of career counseling. In M. L. Savickas & W. B. Walsh (Eds.), *Handbook of Career Counseling Theory and Practice* (pp. 55–80). Palo Alto, CA: Davies-Black.
Lackland, A. C. (2001). Students' choices of college majors that are gender traditional and nontraditional. *Journal of College Student Development, 42*(1), 39–47.
Long, J.S. (2009). *Group Comparisons in Logit and Probit Using Predicted Probabilities*. Retrieved from http://www.indiana.edu/~jslsoc/files_research/groupdif/groupwithprobabilities/groups-with-prob-2009-06-25.pdf.
Long, J. S., & Freeze, J. (2014). *Regression models for categorical dependent variables using stata*. College Station: Stata Press.
Ma, Y. (2009). Family socioeconomic status, parental involvement and college major choices—gender, race/ethnic, and nativity patterns. *Sociological Perspectives, 52*(2), 211–234.
Ma, Y., & Savas, G. (2014). Which is more consequential: Fields of study or institutional selectivity? *Review of Higher Education, 37*(2), 221–247.
Mood, C. (2010). Logistic Regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review, 26*(1), 67–82.
National Association of Colleges and Employers (2014). *NACE salary survey*. Retrieved from http://www.naceweb.org/uploadedFiles/Content/static-assets/downloads/executive-summary/2014-january-salary-survey-executive-summary.pdf.
National Science Foundation (2014). Science and engineering indicators, SESTAT (1993–2010). Retrieved March 15, 2015, from http://www.sestat.nsf.gov.
Nauta, M. M. (2010). The development, evolution, and status of Holland's theory of vocational personalities: Reflections and future directions for counseling psychology. *Journal of Counseling Psychology, 57*(1), 11–22.
Oakes, J. (1990). Opportunities, achievement, and choice: Women and minority students in science and mathematics. *Review of Research in Education, 16*, 153–222.
Paglin, M., & Rufolo, A. (1990). Heterogeneous human capital, occupational choice, and men-women earnings differences. *Journal of Labor Economics, 8*, 123–144.
Pampel, F. C. (2000). *Logistic regression: A primer.* Sage University Papers Series on quantitative applications in the social sciences, 07–132. Thousand Oaks, CA: Sage.
Patrick, J., Niles, S., Margetiak, C. J., & Cunning, T. J. (1993). The longitudinal relationship between extracurricular activities and congruence between initial and final major among college students. *NACADA Journal, 13*(1), 28–33.
Patton, W., & McMahon, M. (1999). *Career development and systems theory: A new relationship*. Pacific Grove: Brooks/Cole.
Porter, S. R., & Umbach, P. D. (2006). College major choice: An analysis of person–environment fit. *Research in Higher Education, 47*(4), 429–449.
Psathas, G. (1968). Toward a theory of occupational choice for women. *Sociology and Social Research, 52*, 253–268.
Randahl, G. J. (1991). A typological analysis of the relations between measured vocational interests and abilities. *Journal of Vocational Behavior, 38*(3), 333–350.
Ryan, J. M., Tracey, T. J. G., & Rounds, J. (1996). Generalizability of Holland's structure of vocational interests across ethnicity, gender and socioeconomic status. *Journal of Counseling Psychology, 43*(3), 330–337.
Schneider and Overton. (1983). Holland personality types and academic achievement. *Journal of Counseling Psychology, 30*(2), 287–289.
Smart, J. C., Feldman, K. A., & Ethington, C. A. (2000). *Academic disciplines: Holland's theory and the study of college students and faculty*. Nashville: Vanderbilt University Press.
Snyder, T. D., & Dillow, S. A. (2012). *Digest of Education Statistics, 2011*. Washington: U.S. Department of Education, National Center for Education Statistics. Retrieved August 30, 2012, from http://www.nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2012001.
Solnick, S. (1995). Changes in women's majors from entrance to graduation at women's and coeducational colleges. *Industrial and Labor Relations Review, 48*(3), 505–514.
Song, C., & Glick, J. E. (2004). College attendance and choice of college majors among asicn-american students. *Social Science Quarterly, 85*(5), 1401–1421.
Swanson, J. L. (1992). The structure of vocational interests for African-American college students. *Journal of Vocational Behavior, 40*(2), 144–157.

Swanson, J. L. (1993). Integrated assessment of vocational interests and self-rated skills and abilities. *Journal of Career Assessment, 1*(1), 50–65.

Train, K. E. (2009). *Discrete choice models with simulation*. New York: Cambridge University Press.

Train, K., & Croissant, Y. (2015). Kenneth train's exercises using the mlogit package for R. Retrieved March 19, 2015, from http://www.cran.r-project.org/web/packages/mlogit/vignettes/Exercises.pdf.

Trusty, J., Robinson, C. R., Plata, M., & Ng, K.-M. (2000). Effects of gender, socioeconomic status, and early academic performance on postsecondary educational choice. *Journal of Counseling and Development, 78*(4), 463–472.

Turner, S., & Bowen, W. G. (1999). Choice of major: The changing (via changing) gender gap. *Industrial and Labor Relations Review, 52*(2), 289–311.

Williams, R. (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. *Sociological Methods & Research, 37*(4), 531–559.