Field Interest and the Choice of College Major

Shannon Breske†
Cory Koedel†
Eric Parsons†

April 2019

We perform exploratory research to understand the importance of interest in the field of study as a determinant of college major choice and examine how the importance of interest varies by student demographic and socioeconomic characteristics. We show that women, white students, and students from advantaged socioeconomic backgrounds rate interest as a more important factor in driving the choice of college major relative to other students. The gender gap in the importance of interest is largest, and it occurs both across and within majors. Gaps by race/ethnicity and socioeconomic status are less pronounced, but they are still large and occur entirely within majors.

Acknowledgement
† University of Missouri-Columbia
The authors thank Maggie Brooks, Cheng Qian, and Andrew Raddant for research assistance. All errors are our own.

1. Introduction

There is significant research and policy interest in understanding how students choose college majors, reflecting both micro- and macro-level concerns. At the micro level, a common concern is that some students make seemingly suboptimal choices. For example, gaps in labor market outcomes between workers in different fields have garnered attention on the grounds that some individuals—women and racial/ethnic minorities in particular—are overrepresented in traditionally lower-paying fields in college and the workforce (Carnevale et al., 2015; Riegle-Crumb and King, 2010). At the macro level, there is concern about the broader distribution of human capital. In the U.S. context, several recent high-profile publications argue that the share of degrees awarded in science and technology fields is too low and that this is a threat to long-term economic prosperity (Committee on Prospering in the Global Economy of the 21st Century, 2007; National Science Board, 2015).

Research and policy efforts aimed at altering student choices have devoted considerable attention to the potential for improved information about labor market outcomes across fields to influence behavior. A prime example is the U.S. Department of Education's College Scorecard, on which post-graduation average earnings is among the core statistics reported for colleges and degree programs. The focus on labor-market information is motivated by the theory that students have inaccurate views about earnings differences across fields, and that if the inaccuracies are remedied, students will choose different (and presumably more lucrative) majors. Numerous studies have developed and tested a variety of interventions in which students receive new information about earnings differences across fields. These studies confirm that students have inaccurate information, and moreover, that they are responsive to interventions that provide better information. However, despite fairly large inaccuracies in student knowledge about labor market outcomes, the effects of these interventions on student decisions have been modest (Baker et al., 2018; Beffy, Fougere and Maurel, 2012; Wiswall and Zafar, 2015; Zafar, 2013).

Researchers have also devoted significant attention to the hypothesis that academic skills prior to college entry drive the distribution of intended majors. However, Riegle-Crumb et al. (2012), who focus on the widely-studied gender gap in STEM enrollment, show that pre-college achievement differences account for little of the gap. Moreover, despite large achievement differences between students who differ by race and social class prior to postsecondary education, they show that students from disadvantaged backgrounds are not underrepresented in STEM fields at entry (also see Arcidiacono and Koedel, 2014).[1] Similarly, Porter and Umbach (2006) find that academic preparation is not a key driver of initial major choice.

Over the course of establishing the modest influence of labor-market expectations and pre-entry academic qualifications over major choice, recent studies have affirmed a clear empirical regularity: students consistently put the most weight in the choice of major on interest in the field of study, and relatedly, expected enjoyment in the career. Zafar (2013) and Wiswall and Zafar (2015) examine undergraduates at selective private universities (Northwestern University and New York University) and find that students place the most value, by far, on enjoyment of the material when selecting a major. Using a survey of students who attend a large, public research university, Weinberger (2004) aims to understand why women are underrepresented in technical fields and concludes that the "overwhelming majority" (p. 31) of women rule out these majors because the courses are not interesting to them. Finally, Baker et al. (2018) report that course enjoyment and grades are the main determinants of major choice in their sample of community college students.[2]

[1] Relatedly, Litzler, Semuelson, and Lorah (2014) find limited evidence that African American and Hispanic students have less STEM confidence than white students unconditionally, and that these students are much more confident than white students after controlling for personal, environmental, and behavioral factors.

[2] The list of studies that come to this general conclusion is substantially longer but we cannot cover all such studies in the text. In related studies focused on college persistence, Allen and Robbins (2009) show that students who are more interested in their majors are more likely to graduate on time and Maltese and Cooper (2016) find that student interest is the strongest predictor of persistence in STEM fields using a retrospective survey.

Motivated by the demonstrated importance of interest as a determinant of major choice, we perform exploratory research to deepen our understanding of the role of interest in student decisions (note that the term "enjoyment" is sometimes used to convey the same concept in the literature). Using data from a large-scale survey of undergraduate students at a state flagship university—University of Missouri-Columbia (MU)—we first replicate the empirical regularity that field interest is by far the most important determinant of major choice. We then examine heterogeneity in the importance of interest for students who differ by demographic and socioeconomic characteristics.

Our dataset has several desirable features that improve upon and complement similar, previous survey-based efforts. First, our student sample is large. We distributed our survey in the fall of 2017 in ten large-lecture, freshman-level courses across a variety of disciplines at MU. Our analytic sample includes over 2,200 students, which facilitates well-powered heterogeneity analyses. Second, we undertook design and implementation of the survey with an explicit focus on attaining a high response rate in order to reduce concerns about sample selection. The end result is that our survey data broadly reflect students in the classes we surveyed at MU: among those in attendance during a lecture we surveyed, our response rate is 94 percent (moreover, given that the survey was administered early in the semester, attendance rates in the courses were high). Third, while MU is a selective institution, it serves a more academically and socioeconomically diverse population than the exceptionally selective universities at which many similar studies have been conducted to date.[3]

We find substantial heterogeneity in the importance of field interest as a determinant of major choice across demographic and socioeconomic groups. Women, white students, and students of high socioeconomic status (SES) are more likely than other students to rate interest as the most important factor driving the choice of major. The gender gap in the importance of interest is largest and occurs

[3] MU is a mid-ranked flagship campus; specifically, as of 2017, it was ranked 29th out of the 50 flagship public campuses by *U.S. News and World Report* (where each state's highest-profile public campus is defined as the "flagship").

both within and across majors. The racial/ethnic and SES gaps are still large, and they occur entirely within majors. The implication of this latter result is that non-white and low-SES students enroll in the same majors as their white and higher-SES peers, but are much more likely to report doing so to improve expected earnings and employment outcomes, rather than because of their interest in the field. The same is true to for men relative to women, although women and men also sort differently to majors in ways that align with the importance they place on field interest. In the concluding section, and in the spirit of Loeb et al. (2017), we discuss how our descriptive findings give context to observed gaps in college success and can be used to inform the design of interventions aimed at modifying students' choices of majors.

2. Survey Instrument & Data

We collected survey data in 10 large-lecture, freshman-level courses taught at MU in fall 2017. We administered the survey in courses in the following fields, where each number in parenthesis indicates the number of separate lectures surveyed: Business Administration (1), Classical Humanities (1), Economics (2), Engineering (1), Mathematics (2), Political Science (1), and Psychology (2).[4] The survey was conducted during the second week of the semester, except in mathematics where it was conducted in week-4 at the request of the instructor (to accommodate the flow of the course).

The focal survey question asks students to rank five factors by their importance for the choice of major. The factors are: (a) expected salary after graduation, (b) stability of the expected career after graduation, (c) fulfillment from expected work after graduation, (d) inherent interest in the field of study, (e) perceived likelihood of success in coursework. It has been shown in other contexts that the order in which the options are given for a rank-order question can affect respondents' answers (e.g., see Krosnick and Alwin, 1987). With this in mind, we created four versions of the survey that use

[4] Note that for mathematics, we surveyed two lectures of a course that is primarily intended for STEM majors, rather than a course in a parallel track primarily intended for non-STEM majors. The reason for this choice was to increase the representation of students who intend to major in STEM fields in our sample.

different, randomly-selected orderings of the options.[5] In addition to the ranking exercise, students were asked to list their intended major, race/ethnicity, gender, level of education for their most-educated parent, residential status, and the first year they attended classes at MU (we use this variable to separate out students new to MU in a robustness test). An example survey is provided in Appendix A.[6]

Above we briefly noted several advantages of our dataset, which we expand on here. First, our large sample size and high survey response rate are the product of two design decisions we made with these objectives in mind. First, we kept the survey short, as shown in Appendix A. The key benefit of keeping the survey short is that it lowered the time cost, which made it easier instructors at MU to agree to participate. The entire process of introducing the survey, passing it out to students, and collecting the completed surveys, took less than 10 minutes. Our ability to administer the survey during class, in turn, made it easier for students to participate (similarly to Baker et al., 2018). The survey was voluntary in all 10 lectures and presented as unrelated to the course content; but our sense is that the low cost to students of participating, given that we were in the classroom already, contributed to the high response rate.

The second design decision aimed at maximizing the response rate is that the survey is anonymous—students did not provide any identifiable information. This permitted us to distribute the survey with IRB approval, but without a separate opt-in process and without raising concerns about privacy. When the survey was introduced in each class, it was made clear that no identifiable

[5] The survey versions are otherwise identical and were distributed randomly to students. See Appendix Table B.3 for tests consistent with the different versions of the survey being randomly distributed. Our preferred analytic models pool estimates across all survey versions. We also confirm that our findings are upheld qualitatively on a version-by-version basis in the appendix.

[6] In addition to the items mentioned in this paragraph, we also asked students about their second-choice majors (i.e., the major they would choose if they were prevented from choosing the first major) and about the individuals who influenced their choices (options: parent(s), friend(s), teacher(s), high school counselor(s)). We do not use data from these questions in the present analysis.

information was being collected. This is in line with our broad efforts to minimize the cost of student participation.

We estimate that just over 94 percent of students in attendance completed the survey.[7,8] If we measure the response rate relative to total posted enrollment in the classes we surveyed, rather than students in attendance when the survey was administered, the response rate declines to 81 percent (see Appendix Table B.1). This is still a high rate and certainly a lower bound because some students, especially early in the semester, can remain officially enrolled despite having dropped the course.[9]

The benefit of the high response rate is that our sample is representative of students who enroll in and attend the large-lecture classes we surveyed at the University of Missouri. This leads to the third benefit of our dataset that makes it a useful complement to the existing literature: MU enrolls a broad and diverse group of students, at least relative to the sites of other survey-based studies, which have focused disproportionately on students who attend selective private institutions (e.g., Barrea College (Stinebrickner and Stinebrickner, 2014); Duke University (Arcidiacono et al., 2012); Northwestern University (Zafar, 2013); New York University (Wiswall and Zafar, 2015)).[10] A related strength that impacts representativeness is that unlike many previous studies in the same vein, we did not recruit participants from a voluntary pool. While our instrument was voluntary and not all students filled it out, most did. This is a far different circumstance than one in which a much smaller number of students are recruited to participate in a more involved study, in which case concerns about the

[7] In cases where a student was present in more than one class that was surveyed, after the first class he or she was instructed to mark the survey with an "OC" on the front (for "other class") and submit it so we could avoid duplicate responses and track this reason for non-response. OC students are omitted from both the numerator and denominator of our response rate calculations. See Appendix Table B.1 (and the table notes).

[8] Moreover, excluding the "Economics 1" lecture, which is the first lecture we surveyed, the response rate increases to 96 percent. We made the mistake in "Economics Lecture 1" of starting the survey a few minutes before class started. This reduced our take-up rate substantially as students who did not arrive early did not get a survey, or if they did, they did not have time to fill it out. Surveys of subsequent classes were not administered until the official start time.

[9] Attendance in these large lectures is not explicitly tracked. We obtained the denominator for the response rates by manually counting students in attendance during the administration of each survey. Although there is surely some measurement error in the attendance counts, there is no reason to expect systematic bias. While not exact, the response rates we calculate should be close to the true values.

[10] Baker et al. (2018) is a recent exception; they examine students who attend two community colleges in California.

representativeness of the sample along both observed and unobserved dimensions arise.[11] This is not to say that studies relying on voluntary participant pools are uninformative—indeed, once participation has been secured, these pools can be used to examine deeper questions that we cannot explore with our data—but a contribution of our work is to bring new, more-broadly representative data to bear on the question of major choice.

Tables 1 and 2 provide summary statistics for our dataset. When available, we also report the MU averages for the 2017 entering cohort based on administrative microdata from the Missouri Department of Higher Education. Table 1 shows that men are overrepresented in our sample—at 54 percent—relative to their share of the MU student body. This is the product of our conducting the survey disproportionately in STEM and business and economics classes (see Table 2 below).[12] Our sample is also predominantly white, but this matches the racial composition of students who attend MU. Note that fewer than 5 percent of the students in our sample are Hispanic or "other race" (where the latter term indicates the student does not belong to any of the other four racial/ethnic groups or belongs to more than two groups).[13] For this reason, in the racial/ethnic comparisons below we focus on the contrasts between Asian, black and white students (although results for all comparisons are reported in the appendix). By virtue of our targeting freshman-level classes, our sample is also disproportionately early on in their college careers. Table 1 shows that two-thirds of students were in

[11] For example, Arcidiacono et al. (2012) perform their analysis using data from 173 students at Duke University, Zafar (2013) uses data from 161 students at Northwestern University, and Wiswall and Zafar (2015) use data from 488 students at New York University.

[12] The disproportionate representation of STEM and business/economics classes is partly due to convenience (business and economics both run multiple large-lecture classes and the faculty we contacted were all agreeable to the survey) and partly the result of a concerted effort on our part to increase the focus on STEM fields, most notably by the inclusion of the mathematics and engineering courses.

[13] As can be seen on the survey instrument in Appendix A, students could select as many racial/ethnic categories as they wanted. In our primary analysis we assign multi-race students who list more than two groups as "other race." For students who list exactly two groups, we assign them a single group using the following hierarchy that balances the objectives of maximizing non-white subgroup sample sizes and prioritizing traditionally underrepresented groups: black, Hispanic, Asian, other race (single designation; e.g., Native American), and white. That said, our findings are robust to other alternatives and we confirm our results are qualitatively unaffected by this issue in the appendix by estimating models where we drop all students from the sample who indicate more than one racial/ethnic designation.

their first semester at MU at the time of the survey, and only 11 percent were beyond their second year. This was by design, as it has been shown that early career major choices are consequential, particularly with respect to degree attainment in STEM fields (Darolia et al., 2018).

Table 1 also shows differences in average factor rankings and the likelihood that each factor is rated as most important by the students in our sample. Note that a factor with a lower average ranking is rated by students as being more important—i.e., the number-1 factor is the one the student indicates to be most important. The table reinforces the empirical regularity that field interest is the key driver of student decisions. It has the lowest average factor ranking by a wide margin and is identified as the most important factor almost half the time. No other factor comes close to this level of importance. Career stability is the next highest-rated factor, followed by career salary and career fulfillment. Success in coursework is the least important factor on average.[14]

Next, Table 2 documents the share of our sample that comes from classes in each field that we surveyed and the distribution of intended majors. As shown in Appendix A, we collected data on the intended major using an open-response survey question, which was numerically coded by two independent research assistants.[15] Intended majors are grouped into nine categories in the table to facilitate presentation; a detailed list of majors under each category is provided in Appendix Table B.2.

Again, we use the administrative microdata to document the MU distribution of intended majors for comparison. The high representation of business and economics courses among the 10 large lectures (in the top panel of Table 2) is reflected in the distribution of intended majors in our sample (in the bottom panel), although these majors are also the most popular at MU. Similarly, our emphasis on STEM in recruiting classes for the survey is reflected by the relatively high share of

[14] This is somewhat surprising in light of evidence from Stinebrickner and Stinebricker (2014), who show that student performance in coursework is a key determinant of persistence in STEM majors after college entry. A potential explanation based on Stinebrickner and Stinebricker (2014) is that students begin college overly optimistic about their ability to perform well academically. Given that our sample consists disproportionately of new entrants, it may be that concerns about academic performance have yet to manifest.

[15] The research team resolved the small number of discrepancies in coding between the assistants.

students with intended STEM majors—specifically in engineering and computer science. That said, there is considerable major diversity in our data, aided by the fact that many of the large-lecture classes we surveyed cover general graduation requirements and thus are attended by a broad mix of students. In total, 36 percent of the sample consists of students intending to major in fields outside of traditional STEM fields and the fields of accounting, business, and economics.

Figure 1 compares intended majors for the students in our sample by gender, race/ethnicity, and SES, where for the latter we proxy throughout our study by the level of parental education. The major groups in Figure 1 are ordered by the total enrollment shares in our sample (from Table 2). The key takeaways from the figure are (a) there are large gender differences in intended majors, (b) modest differences across racial/ethnic categories (particularly between black and white students, who we are best-situated to compare given the demographics of MU), and (c) small differences by SES category.[16]

3. Analysis Plan

The focal survey question asks each student to rank the five factors by how much they influenced the intended major. Our goal is to provide descriptive comparisons of student rankings by gender, race/ethnicity, and SES. We use basic linear regression as a tool to make the comparisons. Specifically, we estimate regressions of the following form:

$$Y_{ik}^{j} = \delta_0 + G_i \delta_1^j + \mathbf{R_i} \boldsymbol{\delta}_\mathbf{2}^\mathbf{j} + \mathbf{P_i} \boldsymbol{\delta}_\mathbf{3}^\mathbf{j} + \gamma_k^j + \varepsilon_{ik}^j \quad (1)$$

In equation (1), Y_{ik}^{j} is the rank of factor *j* (*j*=1,…,5) on a 1-5 scale, for student *i* who took version k of the survey. Per above, a lower-numbered value of Y_{ik}^{j} indicates the factor is more important in the decision. We also estimate versions of the model where Y_{ik}^{j} is an indicator for whether factor *j* is

[16] While most studies are consistent with these results (e.g., Arcidiacono and Koedel, 2014; Bowen, Chingos, and McPherson, 2009; Riegle-Crumb and King, 2010), there are exceptions. For example, Svoboda et al. (2016) find that students from families with lower parental education are less likely to take STEM courses in college. A possible explanation is that they examine course-taking whereas other studies, including ours, examine major enrollment. It is also possible that survey non-response causes sample selection bias in their study, as they use a non-random sample of individuals who responded to a survey and there was significant sample attrition.

ranked first, given the potential for the highest-ranked factor to be particularly important in the decision, in which case equation (1) is estimated as a linear probability model. G_i is an indicator variable equal to one if student *i* is female and zero otherwise; men are the omitted group. $\mathbf{R_i}$ and $\mathbf{P_i}$ are vectors of indicator variables that capture student race/ethnicity and parental education, respectively. As shown in Table 1, we group students into five racial/ethnic categories—Asian, black, Hispanic, white, and other—and three parental-education categories—less than a BA, BA, or graduate degree. White students and students with a parent who has a graduate degree are the omitted groups. γ_k^j is a survey-version fixed effect to control for the survey version. Because the survey versions were randomly given to students, including these fixed effects has no bearing on the findings, but we include them for completeness.

In principle we could estimate equation (1) as a system, which would account for the fact that ε_{ik}^j is mechanically correlated across factors within individuals *i* (i.e., the factor rankings sum to the same value within students). However, we do not use system modeling for our main results for presentational convenience, as it complicates the interpretation of the estimates. The advantage of using the simple linear regressions is that the coefficients on the student-group indicators, as estimated, are directly interpretable as average differences in rankings for each factor. For example, δ_1^j in equation (1) gives the average (conditional) difference in the ranking of factor *j* between women and men. That said, for completeness we show results from a joint modeling exercise in the appendix, from which the findings are qualitatively similar to the results we present in the main text. Specifically, we estimate a multinomial logit where the multinomial outcome captures whether each factor is ranked first. The multinomial logit is comparable to the linear probability models described above based on equation (1).

One might wonder why we bother using a regression framework at all, when we could simply report mean differences across student groups. The reason is that the regressions allow us to make comparisons between specific groups while simultaneously conditioning on student representation in all other groups. For example, the gender difference estimated by equation (1) is conditional on students' racial/ethnic and parental-education designations. This will matter if the composition of students is such that a particular group is over- or under-represented within a gender; e.g., the ratio of black women to men typically exceeds the ratio of white women to men in college (and this is the case in our data as well). We report the conditional average differences generated by equation (1) for this reason.

The regression framework also facilitates an intuitive but somewhat non-standard extension of the model to include intended-major fixed effects, as follows:

$$Y_{iqk}^{j} = \beta_0 + G_i\beta_1^j + \mathbf{R_i}\boldsymbol{\beta}_\mathbf{2}^\mathbf{j} + \mathbf{P_i}\boldsymbol{\beta}_\mathbf{3}^\mathbf{j} + \pi_k^j + \psi_q^j + u_{iqk}^j \qquad (2)$$

Equation (2) is the same as equation (1), except that we introduce the subscript *q* to index the intended major for student *i* and include a fixed effect for the intended major, ψ_q^j.

This is a non-standard extension of the model because the major a student chooses is clearly endogenous to the regression; put another way, one could easily imagine an alternative framework where the intended major is the dependent variable and the factor rankings are the independent variables. That said, equation (2) offers value in our descriptive analysis by allowing us to assess the extent to which differences in average factor rankings across student groups occur across versus within majors. For example, whereas δ_1^j from equation (1) gives the average gender difference in the ranking of factor *j* both across and within intended majors, β_1^j isolates the gender difference only among students who choose the same majors.

4. Results

4.1 Main Findings

Figure 2 shows conditional differences in the factor rankings by gender, race/ethnicity, and SES, as estimated by equation (1). The omitted comparison groups are men, white students, and high-parental-education students (i.e., graduate degree). The figure isolates the group-by-group comparisons for presentational purposes, but all of the output is generated from the same set of regressions. The regression output underlying the graphs is provided in full in Appendix Table B.4.

The graphs in the top row show results from regressions of average factor rankings, in which case a positive value indicates that students in the focal group rank the factor as *less important*, on average, than students in the omitted group. The bottom row of graphs show results from the linear probability models predicting the likelihood of ranking each factor first. In the bottom row, a positive estimate indicates the factor is *more important*—i.e., it is more likely to be ranked first among students in the focal group relative to the omitted group.

Starting with gender, there are clear gaps in the importance of the factors. Women are much more likely to rank field interest and career fulfillment highly, and less likely to rank career salary and stability highly, relative to men (the gender gap for course success is small and not statistically significant). The gender gap in the rank of expected salary is particularly large: women on average rank expected salary 0.70 places lower than men on the 5-point scale and are 10.8 percentage points less likely to list salary as the most important factor. These results align broadly with evidence elsewhere in the literature. For example, Zafar (2013) shows that non-pecuniary factors explain about three fourths of major choice behavior for women, but just half of choice behavior for men.

Next we examine differences by race/ethnicity, focusing on the comparisons between Asian and black students relative to white students due to sample-size issues (again, the full set of estimates, inclusive of Hispanic and other-race students, is in Appendix Table B.4). The differences between black and white students illustrated in Figure 2 load onto two factors: field interest and career salary. Specifically, black students are less likely than white students to report interest as an important driver of the intended major and more likely to report emphasis on career goals, in particular expected salary. The black-white gaps in factor rankings are smaller than the gender gaps, but not trivial. For example, black students on average rank field interest 0.34 places lower on the 5-point scale, and they are almost 6 percentage points less likely than white students to rank field interest first (although this latter difference is not statistically significant). Asian students are similar to black students, with a modest difference being that they put less weight on career stability than both black and white students.

Finally we turn to the comparison by parental education, which shows that interest is most influential in the choice of major for students with highly educated parents. Comparing extremes, students whose parents do not have a college degree rank interest over 0.20 places lower on average than students whose parents have a graduate-level education on the 5-point scale, and are 7.4 percentage points less likely to rank interest as the most important factor. Students with lower parental education are more likely to make their choices based on expected career stability, career salary, and career fulfillment. While the career fulfillment result is somewhat ambiguous given that it is conceptually similar to field interest, note that most of the differential weight on career outcomes falls on the other two career factors—expected salary and career stability—which are more distinct.

4.2 Within/Between Major Gaps

Next we consider the extent to which the factor gaps in Figure 2 persist within the same intended majors. Figure 3 follows the same structure as Figure 2, but is based on output from equation (2), which includes major fixed effects. Thus, the gaps in Figure 3 isolate differences within intended majors only. Note that for presentational convenience, Figure 1 and Table 2 show major categories at a fairly high level of aggregation. However, for the models that underlie Figure 3, we use fixed effects for detailed major codes to ensure the within-major gaps are truly holding the major fixed and not driven by cross-major differences within broader major categories (recall that Appendix Table B.2 shows the detailed major groups). The regression output underlying Figure 3 is available in Appendix Table B.5.

The gender differences in Figure 3 are similar directionally, but muted in magnitude, relative to Figure 2. This indicates that differences in factor rankings by gender occur both within and between majors, yielding two insights: (a) gender differences in sorting to majors partly align with gender differences in the factor weights, and (b) within the same intended majors, women and men report their choices are driven by different factors. The gaps in factor rankings by race/ethnicity and SES in Figure 3 are the same as in Figure 2, both directionally and in magnitude, which indicates that the gaps are driven entirely by student differences within the same majors. The implication is that students who differ by race/ethnicity and SES report being motivated by different factors to pursue the same majors.

5. Robustness

We examine the robustness of our findings to a variety of adjustments to the data and models. We briefly describe the various robustness tests in this section. The results from the tests are relegated to the appendix for brevity—to summarize succinctly, the substance of our findings is highly robust.

First, to test model sensitivity we replace our independently-estimated linear regressions with a multinomial logistic regression where the dependent variable is a 5-category outcome variable indicating the factor that each student lists as most important in the choice of major. The multinomial model is conceptually similar to the linear probability models we use to produce the results shown in the graphs in the bottom rows of Figures 2 and 3. The results from the multinomial model, which we show for the analog to Figure 2 in Appendix Table B.6, are substantively similar to our main findings. While the multinomial model is a more technically-accurate modeling structure given the mechanical connectedness of the factor-ranking values within students, we relegate the results to the appendix because they are difficult to interpret and needlessly complicate the comparisons we aim to make (Wulff, 2015).[17]

We also consider the sensitivity of our findings to various adjustments to, and restrictions on, the data. For ease of presentation, we show results comparable to the top panel in Figure 2 for all of our data-sensitivity analyses, although tests using the results from the other models yield similar results. Recall that regression output corresponding to Figure 2 is available for comparative purposes in Appendix Table. B.4.

First, in Appendix Table B.7 we report results after re-weighting the survey data so that the distribution of intended majors in our sample matches the distribution of intended majors for the incoming 2017 class at MU (as shown in Table 2). The practical implication of the re-weighting is that students with intended majors in the fields of accounting, business and economics, along with engineering and computer science, are down-weighted, while students majoring in other fields are either up-weighted or maintain their weights. Conceptually, the re-weighting is most likely to affect gender gaps in factor rankings because these gaps occur between majors, whereas the racial/ethnic and SES gaps are all within majors. As a practical matter, Table B.7 shows that our findings are substantively unaffected by the re-weighting throughout.

In Appendix Tables B.8 and B.9 we impose restrictions on the sample. First, in Table B.8 we restrict the analysis to students in their first semester on campus at the time of the survey. Our intent in targeting freshman-level classes for the survey was to assess students near the point of entry

[17] The appendix includes a brief description of the findings explaining the similarity of results. The similarity of output between linear and non-linear models with limited dependent variables in multiple contexts has been well-documented, including by Angrist and Pischke (2009). Similarly to here, Angrist and Pischke (2009) argue that non-linear models can needlessly complicate analyses (pp. 105-07).

into college; the data restriction in Table B.8 in simply more forceful in this intent. The results for first-year students are similar to the results for the full sample.

In Table B.9 we restrict our analysis to only students who indicate a single racial/ethnic category. Recall from above that we imposed decision rules for students who indicated multiple racial/ethnic categories to assign each student to a single racial/ethnic group. Our assignment process was such that we leaned toward assigning students to traditionally underrepresented racial/ethnic groups in cases where multiple categories were selected. The analysis in Table B.9 assesses whether our somewhat arbitrary strategy for handling these students affects our findings. Again, the results in Table B.9 are very similar to our main findings, indicating that the way we treat multi-category students by race/ethnicity does not drive our results.

Finally, in Table B.10, panels (a)-(d), we show results separately using each version of the survey instrument. These tables show that the order in which the factors are listed on the survey does impact students' rankings, which is consistent with previous research, but the key patterns in our comparative analysis are generally upheld using the different survey versions. A caveat is that the version-by-version models are noisy because the sample sizes are significantly reduced (to about one-fourth the size of the full sample for each survey version). However, there is no indication from Table B.10 that our findings are substantively sensitive to factor ordering on the survey.

6. Discussion & Conclusion

We report on exploratory research designed to improve our understanding how students choose college majors. Our findings are based on results from a high-response-rate survey completed by over 2,200 students in 10 large-lecture classes at a state flagship university. Consistent with previous research, we show that field interest is by far the most important factor driving students' choices of majors. We also document gaps in the importance of interest by gender, race/ethnicity, and SES. Specifically, students who are female, white, and whose parents are highly educated are more likely to report that interest is a more important factor in the choice of major. The gender gap in the importance of interest is larger than the gaps by race/ethnicity and SES, and occurs both within and across majors; the racial/ethnic and SES gaps occur entirely within majors.

The most obvious question raised by our analysis is what drives differences in student emphasis on field interest. Noting that career outcomes are the counterweight to interest in our study, the most straightforward explanation is that differences in the perceived need for education to pay off as an investment are important. For example, the gender gap may be driven in part by expectations of traditional gender roles among some women, which will reduce the importance of high earnings for women on average (Davis and Greenstein, 2009) and allow for an increased emphasis on field interest. In contrast, students from disadvantaged racial/ethnic or SES backgrounds—who will have weaker safety nets on average should their college investments not pay off—would be expected to put more emphasis on earnings and career outcomes.

The fact that large gaps in the importance of interest exist across demographic and socioeconomic groups, even within the same majors, likely has implications for student success in college. For example, research indicates that individuals who are intrinsically motivated are more successful than those who are extrinsically motivated (Allen and Robbins, 2010), all else equal. While there is nothing wrong with career-oriented choice—and indeed, from the perspective of labor-market efficiency there are reasons to prefer it—the gaps we document in the importance of interest across student groups raise the possibility that differences in intrinsic and extrinsic motivation differentially affect their college outcomes.

More broadly, the significance of field interest as a determinant of student decisions, and the variability of its significance across student groups, highlights the importance of understanding what influences student interest in different fields prior to college entry. This line of inquiry has been pursued most vigorously with respect to gender gaps in STEM. Factors such as stereotype threat, individual competitiveness, and gender differences in the distribution of math achievement, among others, have been hypothesized and tested as factors that drive gender gaps in STEM interest and outcomes. Kahn and Ginther (2017) review the evidence on these possibilities. There is less research on the factors that drive interest for students who differ by race/ethnicity and SES, perhaps because the distributions of majors by race/ethnicity and SES are similar (Figure 1) despite the gap in the importance of interest we document here.

We conclude with an implication of our findings for future interventions aimed at altering students' choices of majors: there are efficiency and equity consequences of interventions designed to nudge students into particular majors depending on the source of the nudge. For example, the higher weight that men put on expected salary implies that to the extent that salary-based interventions are effective at shifting behavior, men are more likely to be affected. Given the current levels of representation of men and women in high-paying fields, this would be expected to exacerbate field-based gender gaps. By the same logic, the higher weights placed on salary and career-stability considerations by black students and students from low-parental-education families suggests that such interventions may be more effective in altering behavior for these groups relative to their more-advantaged peers. In summary, efforts to influence students' major choices should acknowledge that different types of students may be driven by different factors to varying degrees—tailoring an intervention for the types of students who will be targeted can improve efficacy.

References

Allen, J., and Robbins, S. (2010). Effects of Interest-Major Congruence, Motivation, and Academic Performance on Timely Degree Attainment. *Journal of Counseling Psychology* 57(1), 23-35.

Angrist, J.D., and Pischke, J.S. (2009). *Mostly Harmless Econometrics*. New Jersey, NJ: Princeton University Press.

Arcidiacono, P., Hotz, J.V., and Kang, S. (2012). Modeling College Major Choice using Elicited Measures of Expectations and Counterfactuals. *Journal of Econometrics* 166(1), 3-16.

Arcidiacono, P., and Koedel, C. (2014). Race and College Success: Evidence from Missouri. *American Economic Journal: Applied Economics*, 6(3), 20-57.

Baker, R., Bettinger, E., Jacob, B., and Marinescu, I. (2018). The Effect of Labor Market Information on Community College Students' Major Choice. *Economics of Education Review* 65, 18-30.

Beffy, M., Fougere, D., and Maurel, A. (2012). Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter? *Review of Economics and Statistics* 94(1), 334-347.

Bowen, W.G., Chingos, M.M., and McPherson, M.S. (2009). *Crossing the Finish Line*. Princeton, New Jersey: Princeton University Press.

Carnevale, A.P., Fasules, M.L., Porter, A., & Landis-Santos, J. (2016). African Americans: College Majors and Earnings. Washington, DC: Georgetown University Center on Education and the Workforce.

Committee on Prospering in the Global Economy of the 21st Century. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington DC: The National Academies Press.

Darolia, R., Koedel, C., Main, J.B., Ndashimye, F., and Yan, J. (2018). High School Course Access and Postsecondary STEM Enrollment and Attainment. CALDER Working Paper No. 186.

Davis, S.N., and Greenstein, T.N. (2009). Gender Ideology: Components, Predictors, and Consequences. *Annual Review of Sociology* 35, 87-105.

Krosnick, J.A., and Alwin, D.F. (1987). An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. *Public Opinion Quarterly* 51(2), 201-219.

Litzler, E., Samuelson, C.C., and Lorah, J.A. (2014). Breaking it Down: Engineering Student STEM Confidence at the Intersection of Race/Ethnicity and Gender. *Research in Higher Education* 55(8), 810-32.

Loeb, S., Morris, P., Dynarski, S., Reardon, S., McFarland, D., and Reber, S. (2017). Descriptive Analysis in Education: A Guide for Researchers. Policy Report. Washington, DC: United States Department of Education.

Maltese, A.V., and Cooper, C.S. (2017). STEM Pathways: Do Men and Women Differ in Why They Enter and Exit? *AERA Open* 3(3), 1-16.

National Science Board. 2015. *Revisiting the STEM Workforce: A Companion to Science and Engineering Indicators 2014*. Washington DC: National Science Foundation.

Porter, S. and Umbach, P. (2006). College Major Choice: An Analysis of Person–Environment Fit. *Research in Higher Education*, 47(4), 429-449.

Riegle-Crumb, C., and King, B. (2010). Questioning a White Male Advantage in STEM: Examining Disparities in College Major by Gender and Race/Ethnicity. *Educational Researcher* 39(9), 656-664.

Riegle-Crumb, C., King, B., Grodsky, E., and Muller, C. (2012). The More Things Change, the More They Stay the Same? Prior Achievement Fails to Explain Gender Inequality in Entry Into STEM College Majors Over Time. *American Educational Research Journal* 49(6), 1048-1073.

Stinebrickner, R., and Stinebrickner, T.R. (2014). A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. *Review of Economic Studies* 81(1), 426-472.

Svoboda, R.C., Rozek, C.S., Hyde, J.S., Harackiewicz, J.M., and Destin, M. (2016). Understanding the Relationship Between Parental Education and STEM Course Taking Through Identity-Based and Expectancy-Value Theories of Motivation. *AERA Open* 2(3), 1-13.

Weinberger, C. (2004). "Just Ask! Why Surveyed Women Did Not Pursue Information Technology Courses or Careers?" *IEEE Technology and Society*, 23(2): 28-35.

Wiswall, M. and Zafar, B. (2015). Determinants of College Major Choice: Identification using an Information Experiment. *Review of Economic Studies*, 82(2), 791-824.

Wulff, J. (2015). Interpreting Results from the Multinomial Logit Model: Demonstrated by Foreign Market Entry. *Organizational Research Methods* 18(2), 300-325.

Zafar, B. (2013). College Major Choice and the Gender Gap. *Journal of Human Resources* 48(3), 545-595.

Figure 1. Intended Major Shares by Gender, Race/Ethnicity, and Parental Education.

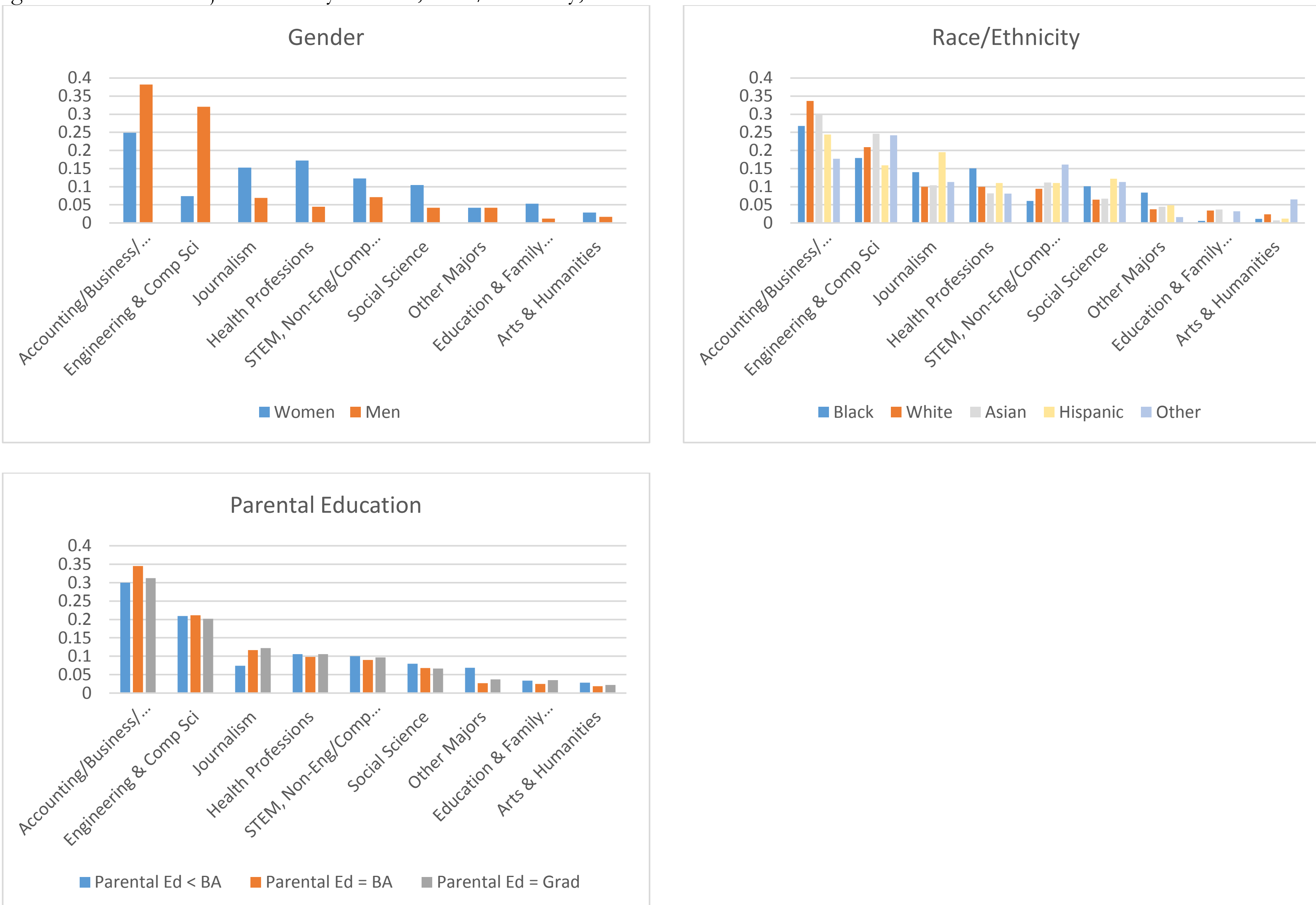
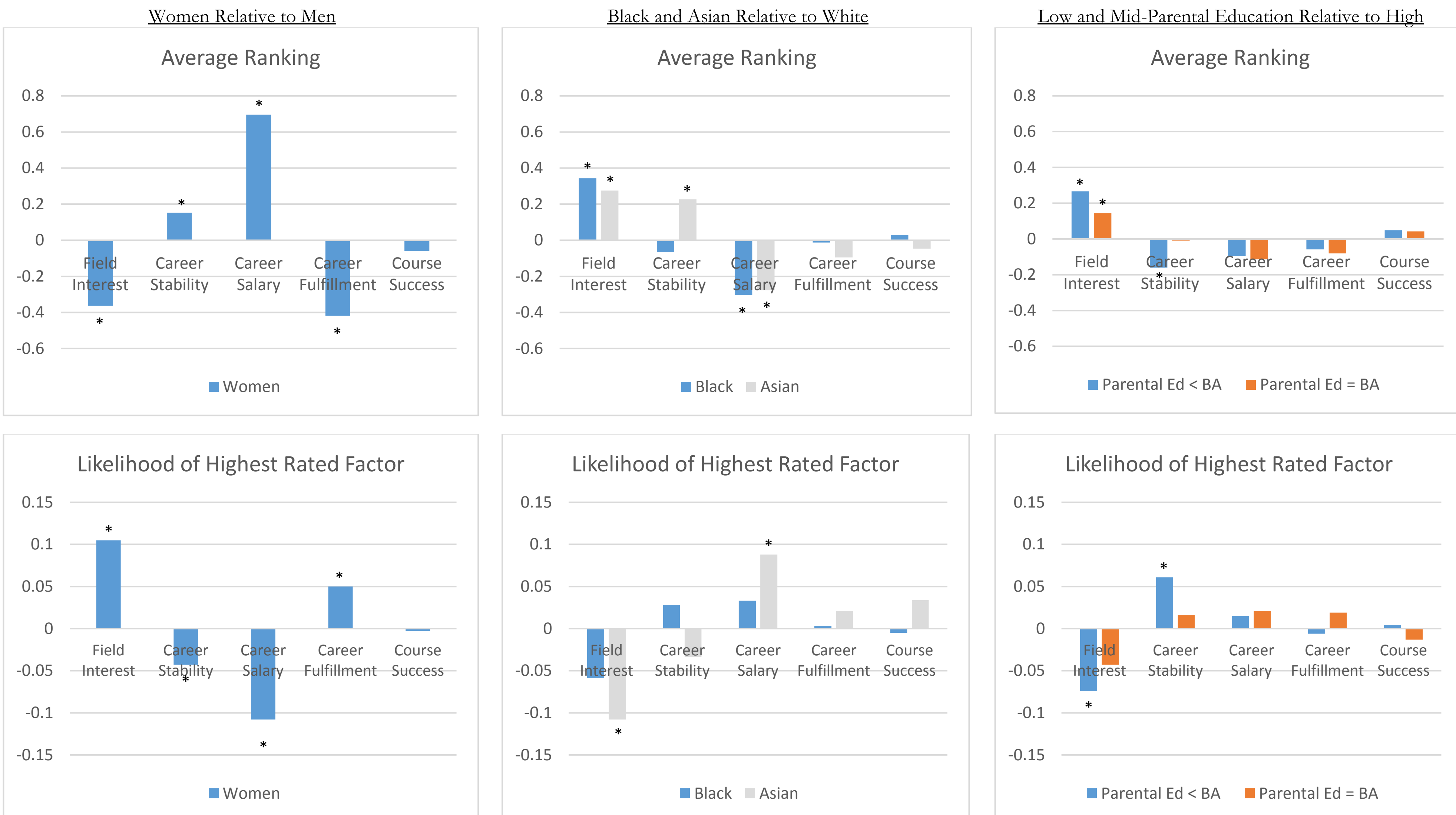
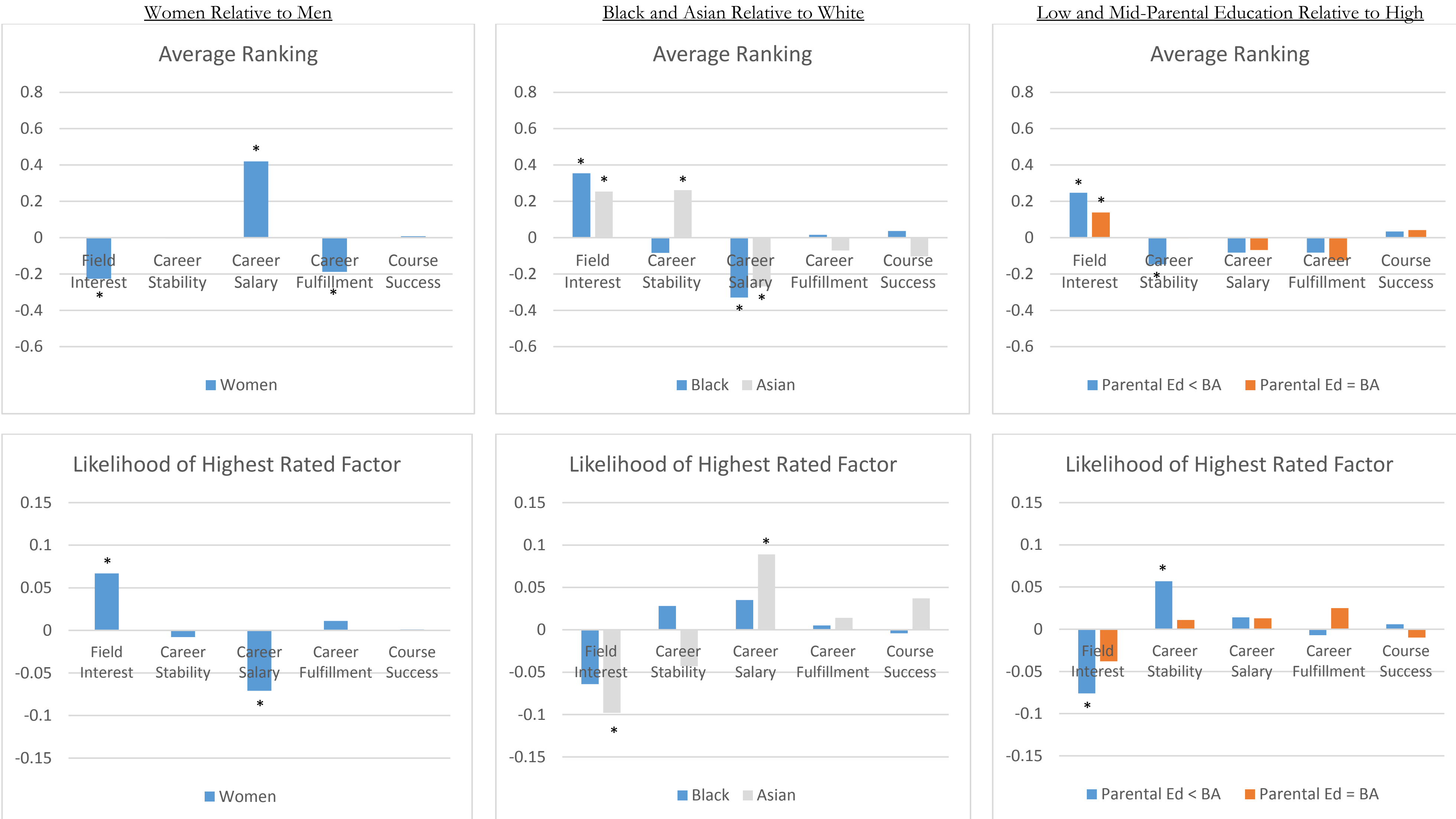


Figure 2. Factor Ranking Gaps by Gender, Race/Ethnicity, and Parental Education. Row 1: Average Ranking Gaps; Row 2: Gaps in Likelihood of Ranking Each Factor First. Corresponding Regression Output is in Appendix Table B.4.



* Indicates the difference is statistically significant from the omitted group at the 5 percent level or better in the linear model (also see Appendix Table B.4).

Figure 3. Factor Ranking Gaps by Gender, Race/Ethnicity, and Parental Education, Within Majors. Row 1: Average Ranking Gaps; Row 2: Gaps in Likelihood of Ranking Each Factor First. Corresponding Regression Output is in Appendix Table B.5.



* Indicates the difference is statistically significant from the omitted group at the 5 percent level or better in the linear model (also see Appendix Table B.5).

Table 1. Student Summary Statistics.

	Survey Sample Avg. (standard deviation)	MU Avg., Fall-2017 Entering Cohort (standard deviation)
Student Demographics and SES		
Female	0.46 (0.50)	0.53 (0.50)
Asian	0.06 (0.24)	0.02 (0.16)
Black	0.08 (0.27)	0.07 (0.26)
Hispanic	0.04 (0.19)	0.05 (0.21)
White	0.80 (0.27)	0.79 (0.41)
Other	0.03 (0.16)	0.07 (0.25)
Highest parental education < bachelor's	0.27 (0.45)	--
Highest parental education = bachelor's	0.38 (0.49)	--
Highest parental education = graduate degree	0.35 (0.48)	--
Other Student Characteristics		
First year on campus	0.67 (0.47)	1.0
Second year on campus	0.22 (0.41)	0.0
In-state student	0.66 (0.47)	0.68 (0.47)
Average Factor Rankings (Most to Least Influential)		
Field interest	2.21 (1.41)	--
Career Stability	2.84 (1.25)	--
Career Salary	3.16 (1.36)	--
Career fulfillment	3.20 (1.38)	--
Course success	3.61 (1.28)	--
Shares of Surveys Where Each Factor is Listed as Most Important		
Field interest	0.47 (0.50)	--
Career Stability	0.18 (0.38)	--
Career Salary	0.14 (0.35)	--
Career fulfillment	0.14 (0.34)	--
Course success	0.08 (0.27)	--
N	2240	5050

Notes: The gender and racial/ethnic shares reported in column (2) are taken from administrative microdata collected by the Missouri Department of Higher Education and are for all first-time MU students in fall 2017. An "--" indicates that information from the administrative data is unavailable.

Table 2. Sample Shares: Courses Surveyed and Intended Majors.

Course Type Surveyed (# of Sections Surveyed)	Sample Shares	MU Shares Fall-2017 Entering Cohort
Business Administration (1)	0.12	--
Classical Humanities (1)	0.15	--
Economics (2)	0.22	--
Engineering (1)	0.05	--
Mathematics (2)	0.09	--
Political Science (1)	0.15	--
Psychology (2)	0.21	--
Major Groups		
Accounting, Business & Economics	0.32	0.23
Engineering and Computer Science	0.21	0.12
Journalism	0.11	0.11
Health Professions	0.10	0.16
Non-engineering/computer science STEM	0.10	0.14
Social Science (excluding economics)	0.07	0.06
Other Majors	0.04	0.08
Education and Family Studies	0.03	0.05
Arts and Humanities	0.02	0.04
N	2240	5050

Notes: The intended-major shares reported in column (2) are taken from administrative microdata collected by the Missouri Department of Higher Education and are for all first-time MU students in fall 2017. An "--" indicates that information from the administrative data is unavailable (note that we do not have access to university-level transcript microdata to construct enrollment shares for individual courses).

Appendix A
Example Survey Instrument

Please answer the following questions to support our study of how Mizzou students select their fields of study. This survey is anonymous and voluntary.

<u>Please tell us about your choice of major:</u>

1. **What is your current intended major?**
(Only list *undecided* if you truly have no idea about what field you plan to study. We are interested in your intent at this point in time, recognizing that it may be uncertain and could change in the future.)

Intended Major:____

----

2. **If you had to choose a different major, what major would you choose?**

Second-Choice Major:____

----

3. **Rank the following factors in order, *from most to least important*, in terms of how much they influenced your choice of major:** ***[NOTE: Options re-ordered at random across surveys]***

A. Expected salary after graduation
B. Stability of expected career after graduation
C. Fulfillment from expected work after graduation
D. Inherent interest in the field of study
E. Perceived likelihood of success in coursework

Rank Ordering (letter of most important first; letter of least important last):

____

----

4. **Rank the following individuals in order, *from most to least important*, in terms of influencing your intended major (outside of yourself):** ***[NOTE: Options re-ordered at random across surveys]***

A. Parent(s) and/or other family
B. Friend(s)
C. Teacher(s)
D. High school guidance counselor(s)

Rank Ordering (letter of most important first; letter of least important last):

____

----

5. **Indicate how strongly you agree with this statement: Individuals other than myself were influential in my choice of major. (circle one)**

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

----

6. **Indicate how strongly you agree with this statement: I am confident that my current major will be my final major. (circle one)**

Strongly Agree   Agree   Neutral   Disagree   Strongly Disagree

----

<u>Now please tell us about yourself:</u>

7. **Race/Ethnicity (circle all that apply)**

Asian   Black   Hispanic   Native American

Pacific Islander   White   Other   Prefer not to Answer

----

8. **Gender Identity (circle one)**

Female   Male   Other (specify):___________________________.

Prefer not to Answer

----

9. **Did you attend high school in Missouri? (circle one)**

Yes   No

----

10. **How would you describe the environment you lived in for the majority of your life? (circle one)**

Urban   Suburban   Rural

----

11. **When did you take your first on-campus class at Mizzou? (circle one)**

2017   2016   2015   2014   2013 or before

----

12. **What is the highest level of education that either of your parents completed? (circle one)**

Less than HS   High School   Some College   Associate’s Degree

Bachelor’s Degree   Graduate Degree

Appendix B
Supplementary Tables

Appendix Table B.1. Data Exclusions and Final Analytic Sample.

Total Listed Enrollment in Surveyed Classes from Registrar	3202
Total Attendance in Surveyed Classes as Measured by the Research Team	2792
Surveys Collected	2655
Sample after Dropping Students Who:	
Indicated taking the survey in a previous class	(-270) 2385
Actively chose not to participate	(-21) 2364
Did not provide basic demographics and socioeconomic status	(-82) 2282
Did not answer the factor ranking question	(-5) 2277
Did not provide an intended major	(-37) 2240
Final Analytic Sample	2240

Notes: Students who did not want to participate were asked to write an "N" on the front of the survey and return it. We code these students as actively choosing to opt out (there were 21 such instances in the sample). Passive non-participation is reflected in the gap between rows 2 and 3 of the table (i.e., some students simply did not submit a survey), although in some cases this was due to, say, a student showing up a few minutes late to class such that she was counted by our research team but did not actually get a survey. Students who had taken the survey in another class were instructed to write "OC" on the front of the survey and return it (there were 270 such instances in the sample). These are students who were already surveyed, and as such they are ignored in calculating the response rate—i.e., the response rate among students in attendance is calculated as (2655-270)/(2792-270). The number of lost observations reported for each reason in the table is dependent on the order in which the reasons are listed; e.g., some students who did not provide basic demographic information also did not answer the ranking question and/or did not list the intended major.

Appendix Table B.2. Major Groups.

Accounting, Bus, Econ	**Engineering/ Computer Science**	**Journalism**	**Health Professions**	**Non-Engineering/Non-Comp-Sci STEM**	**Social Science**	**Other Majors**	**Education & Family Stud.**	**Arts & Humanities**
Accounting	Mechanical Engineering	Journalism (various)	Health Science	Biology	Political Science	Athletic Training	Early Childhood Education	Art
Business	Civil Engineering		Clinical and Diag. Sciences	Chemistry	Sociology	Agriculture	Educational Studies	Music
International Business	Chemical Engineering		Nursing	Physics	Psychology	Digital Storytelling	Agricultural Education	Anthropology
Economics	Biological Engineering		Occupational Therapy	Animal Sciences	Communications	Geography	Elementary Education	English
Finance	Computer Engineering		Physical Therapy	Biochemistry	Public Health	Hospitality Management	Middle School Education	History
Marketing	Computer Science		Respiratory Therapy	Environmental Sciences	Pre-Law	Interdisciplinary	Secondary Education	Art History
Management	Electrical Engineering			Food Science and Nutrition		International Studies	Human Development & Family Studies	Classics
Agribusiness Management	Industrial Engineering			Geological Sciences		Parks, Recreation and Sport	Special Education	Film Studies
Agricultural Economics	Information Technology			Mathematics		Personal Financial Planning		Non-English Language
	Aerospace Engineering			Natural Resources Science & Management		Social Work		Linguistics
	General Engineering			Nutritional Sciences		Textile and Apparel Management		Philosophy
				Plant Sciences		Architectural Studies		Religious Studies
				Statistics				Romance Languages
				Pre-Medical/Dentistry				Theater
								Graphic Design

Notes: This table provides details on how individual majors were aggregated into the nine groups used in some of the analysis in the main text. The majors listed here also indicate the degree of detail in the major groups we use for the major-fixed-effects models (Figure 3).

Table B.3. Balancing Tests Across the Four Versions of the Survey.

	F-statistic	P-value
Female	0.99	0.40
Asian	0.30	0.82
Black	1.22	0.30
Hispanic	0.52	0.67
Other Race	0.98	0.40
Parental Education < BA	1.31	0.27
Parental Education = BA	1.48	0.22

Notes: Each F-statistic is from a regression of the indicated student characteristic on the survey version indicators, where one indicator is omitted. The F-statistic is for a test of the joint significance of the survey version indicators. Null results are consistent with the survey versions being uncorrelated with student characteristics, which is the expectation given that the survey versions were handed out to students at random.

Appendix Table B.4. Linear Regression Output Underlying Figure 2.

	Rank Value Models (top row of Figure 2)						Highest Ranked Factor Models (bottom row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success		Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.363 (0.058)**	0.153 (0.053)**	0.696 (0.055)**	-0.419 (0.058)**	-0.060 (0.054)		0.105 (0.021)**	-0.044 (0.016)**	-0.108 (0.014)**	0.050 (0.015)**	-0.003 (0.012)
Asian	0.275 (0.128)*	0.226 (0.120)*	-0.276 (0.122)**	-0.095 (0.122)	-0.047 (0.120)		-0.108 (0.042)**	-0.034 (0.033)	0.088 (0.037)*	0.021 (0.032)	0.034 (0.028)
Black	0.343 (0.118)**	-0.067 (0.102)	-0.304 (0.100)**	-0.012 (0.105)	0.030 (0.101)		-0.059 (0.038)	0.028 (0.032)	0.033 (0.028)	0.003 (0.028)	-0.005 (0.020)
Hispanic	-0.130 (0.159)	0.134 (0.140)	0.402 (0.152)**	0.045 (0.154)	-0.385 (0.147)**		0.076 (0.055)	-0.045 (0.040)	-0.057 (0.032)	-0.044 (0.032)	0.070 (0.040)
Other Race	0.040 (0.184)	-0.035 (0.135)	0.129 (0.181)	-0.093 (0.174)	-0.043 (0.173)		0.012 (0.063)	-0.068 (0.040)	0.036 (0.049)	0.012 (0.045)	0.008 (0.036)
Parental Ed < BA	0.267 (0.077)**	-0.161 (0.068)*	-0.095 (0.070)	-0.058 (0.074)	0.049 (0.071)		-0.074 (0.027)**	0.061 (0.021)**	0.015 (0.018)	-0.006 (0.018)	0.004 (0.015)
Parental Ed = BA	0.145 (0.067)*	-0.010 (0.062)	-0.114 (0.064)	-0.081 (0.068)	0.043 (0.063)		-0.043 (0.025)	0.016 (0.018)	0.021 (0.017)	0.019 (0.017)	-0.013 (0.013)
Survey Version 1	-0.326 (0.081)**	0.127 (0.074)	0.177 (0.077)*	0.277 (0.082)**	-0.241 (0.075)**		0.144 (0.029)**	-0.061 (0.023)**	-0.036 (0.019)	-0.058 (0.021)*	0.012 (0.016)
Survey Version 2	-0.129 (0.083)	0.266 (0.075)**	-0.300 (0.078)**	0.203 (0.082)**	-0.038 (0.075)		0.059 (0.029)*	-0.067 (0.023)**	0.050 (0.022)*	-0.046 (0.021)*	0.003 (0.015)
Survey Version 3	-0.011 (0.085)	0.043 (0.075)	0.073 (0.079)	0.147 (0.081)	-0.219 (0.079)**		0.034 (0.029)	-0.016 (0.024)	-0.011 (.020)	-0.045 (0.021)*	0.037 (0.017)*
R-Squared	0.038	0.015	0.092	0.029	0.011		0.031	0.015	0.039	0.011	0.007
N	2240	2240	2240	2240	2240		2240	2240	2240	2240	2240

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, students with a parent who has a graduate degree, and students who received survey version four.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.5. Linear Regression Output Underlying Figure 3.

	Rank Value Models (top row of Figure 3)						Highest Ranked Factor Models (bottom row of Figure 3)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success		Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.226 (0.065)**	-0.004 (0.059)	0.420 (0.059)**	-0.188 (0.064)**	0.008 (0.062)		0.067 (0.023)**	-0.008 (0.018)	-0.071 (0.016)**	0.011 (0.017)	0.001 (0.013)
Asian	0.254 (0.123)*	0.262 (0.112)*	-0.268 (0.112)**	-0.071 (0.121)	-0.103 (0.117)		-0.098 (0.044)*	-0.043 (0.035)	0.089 (0.031)**	0.014 (0.031)	0.037 (0.025)
Black	0.354 (0.107)**	-0.084 (0.098)	-0.329 (0.098)**	0.016 (0.106)	0.037 (0.102)		-0.064 (0.039)	0.028 (0.030)	0.035 (0.027)	0.005 (0.027)	-0.004 (0.021)
Hispanic	-0.051 (0.154)	0.010 (0.141)	0.302 (0.140)*	0.153 (0.151)	-0.353 (0.146)*		0.047 (0.055)	-0.022 (0.043)	-0.042 (0.039)	-0.057 (0.039)	0.074 (0.031)*
Other Race	0.211 (0.176)	-0.167 (0.161)	0.007 (0.160)	0.053 (0.172)	-0.116 (0.167)		-0.047 (0.063)	-0.035 (0.050)	0.054 (0.045)	0.001 (0.045)	0.028 (0.035)
Parental Ed < BA	0.276 (0.075)**	-0.146 (0.068)*	-0.082 (0.068)	-0.083 (0.073)	0.034 (0.071)		-0.076 (0.027)**	0.057 (0.021)**	0.014 (0.019)	-0.001 (0.019)	0.006 (0.015)
Parental Ed = BA	0.139 (0.067)*	-0.004 (0.061)	-0.068 (0.061)	-0.124 (0.065)	0.042 (0.063)		-0.038 (0.024)	0.011 (0.019)	0.013 (0.017)	0.025 (0.017)	-0.010 (0.013)
Survey Version 1	-0.372 (0.082)	0.188 (0.074)**	0.219 (0.074)**	0.239 (0.080)**	-0.259 (0.077)**		0.162 (0.029)**	-0.072 (0.023)**	-0.046 (0.021)*	-0.056 (0.021)**	0.012 (0.016)
Survey Version 2	-0.110 (0.081)	0.271 (0.074)**	-0.329 (0.073)**	0.234 (0.079)**	-0.069 (0.076)		0.057 (0.029)*	-0.065 (0.023)**	0.054 (0.020)**	-0.053 (0.020)**	0.006 (0.016)
Survey Version 3	0.008 (0.081)	0.072 (0.074)	0.060 (0.074)	0.122 (0.079)	-0.229 (.076)**		0.030 (0.029)	-0.019 (0.023)	-0.010 (0.021)	-0.041 (0.021)*	0.039 (0.016)
Major Fixed Effects	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
R-squared (within only)	0.032	0.013	0.059	0.023	0.011		0.028	0.011	0.026	0.007	0.008
N	2240	2240	2240	2240	2240		2240	2240	2240	2240	2240

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, students with a parent who has a graduate degree, and students who received survey version four.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.6. Multinomial Logistic Regression Output Where the Dependent Variable Captures Five Possible Outcomes: The Ranking of Each Factor as Most Important. The Omitted Factor is Career Salary.

	Highest Ranked Factor is:			
	Field Interest	Career Stability	Career Fulfillment	Course Success
Female	1.089 (0.144)**	0.609 (0.166)**	1.223 (0.174)**	0.819 (0.200)**
Asian	-0.776 (0.249)**	-0.692 (0.299)**	-0.357 (0.307)	-0.138 (0.338)
Black	-0.377 (0.244)	-0.085 (0.273)	-0.230 (0.300)	-0.309 (0.361)
Hispanic	0.707 (0.422)	0.275 (0.489)	0.132 (0.550)	1.202 (0.496)*
Other Race	-0.215 (0.367)	-0.716 (0.495)	-0.157 (0.466)	-0.137 (0.555)
Parental Ed < BA	-0.295 (0.181)	0.208 (0.197)	-0.178 (0.216)	-0.079 (0.239)
Parental Ed = BA	-0.259 (0.156)	-0.064 (0.184)	-0.027 (0.192)	-0.340 (0.227)
Survey Version 1	0.627 (0.200)**	-0.030 (0.228)	-0.086 (0.241)	0.483 (0.287)
Survey Version 2	-0.180 (0.178)	-0.695 (0.208)**	-0.626 (0.219)**	-0.264 (0.274)
Survey Version 3	0.166 (0.190)	0.005 (0.212)	-0.215 (0.229)	0.536 (0.268)*
N	2240			

Notes: The output in this table is from a single multinomial model and is most comparable conceptually to the output from the separate regressions shown in the bottom panel of Figure 2. The omitted factor to which all coefficients are relative is career salary. The omitted groups are men, white students, students with a parent who has a graduate degree, and students who received survey version four.
**/* indicates statistical significance at the 1-percent/5-percent level.

Explanation of Table B.6

Table B.6 shows output from a multinomial logistic regression analogous to the output reported in the bottom panel of Figure 2, (in Figure 2, recall that we report estimates based on independently-estimated linear probability models where the outcome variables are indicators for whether each factor is ranked highest). As noted in the text, the multinomial model output is qualitatively similar to what we show in Figure 2, although more difficult to interpret because all estimates are relative to an omitted category (Wulff, 2015). In our model, the omitted outcome is ranking expected career salary highest.

We briefly highlight the similarity of findings with two examples in the table. First, the coefficient on the female indicator variable is consistently large and positive for the outcomes in which a non-salary factor is ranked first. This means that the relative probability of ranking any of these variables first, rather than career salary, is higher for women relative to men (holding other characteristics fixed), which is what we report in the main text. Second, 7 of 8 coefficients for the lower parental education groups (the omitted group is students with a parent who has a graduate degree) are negative, and they are approaching statistical significance for field interest (in fact, both coefficients are marginally significant at the 10 percent level). The one positive coefficient is for career stability for the lowest parental-education group, which matches the result in Figure 2 that this group rates the importance of career stability relatively higher than all other groups. The general pattern of negative estimates indicates that the relative probabilities of ranking the non-salary variables first are lower among the lower-SES students; put another way, these students are relatively more likely to rank expected salary higher. The differences are generally modest and not statistically distinguishable, which again is similar to what we show in the bottom panel of Figure 2.

Appendix Table B.7. Replication of Main Findings (Figure 2, top row; or Appendix Table B.4) Using Re-weighted Data to Match the Distribution of Majors at MU (Based on Administrative Data).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.354 (0.062)**	0.118 (0.056)*	0.651 (0.059)**	-0.357 (0.063)**	-0.058 (0.059)
Asian	0.303 (0.138)*	0.228 (0.126)	-0.362 (0.129)**	0.024 (0.141)	-0.107 (0.136)
Black	0.450 (0.123)**	-0.130 (0.106)	-0.420 (0.116)**	0.020 (0.117)	0.064 (0.112)
Hispanic	-0.069 (0.180)	0.198 (0.152)	0.331 (0.160)*	0.005 (0.173)	-0.430 (0.151)**
Other Race	-0.039 (0.183)	-0.039 (0.151)	0.159 (0.190)	-0.097 (0.194)	0.010 (0.182)
Parental Ed < BA	0.316 (0.082)**	-0.202 (0.073)**	-0.105 (0.076)	-0.064 (0.082)	0.078 (0.077)
Parental Ed = BA	0.141 (0.069)*	-0.013 (0.066)	-0.132 (0.069)	-0.064 (0.074)	0.057 (0.067)
Survey Version 1	-0.298 (0.085)**	0.127 (0.078)	0.160 (0.082)	0.261 (0.091)**	-0.235 (0.081)
Survey Version 2	-0.163 (0.087)	0.286 (0.081)**	-0.315 (0.084)**	0.211 (0.090)*	-0.018 (0.082)
Survey Version 3	-0.053 (0.091)	0.053 (0.080)	0.120 (0.084)	0.143 (0.089)	-0.216 (0.084)
N	2240	2240	2240	2240	2240

Notes: Each column shows results from an independently estimated weighted regression. The weights are at the large-group major level and equal to the ratio of the MU-sample major share divided by the survey-sample major share. For example, for accounting/business/economics majors the weight is 0.23/0.32 (per Table 2). Robust standard errors are in parenthesis. The omitted groups are men, white students, students with a parent who has a graduate degree, and students who received survey version four. R-squared values are suppressed because the meaning of this statistic changes with the weighting.

**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.8. Replication of Main Findings Using Students in their First Year at MU Only (Comparable to Figure 2, top row).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.359 (0.072)**	0.173 (0.064)**	0.738 (0.067)**	-0.466 (0.071)**	-0.067 (0.066)
Asian	0.209 (0.165)	0.325 (0.156)*	-0.167 (0.159)	-0.221 (0.154)	0.002 (0.160)
Black	0.407 (0.155)**	-0.194 (0.136)	-0.384 (0.123)**	-0.059 (0.127)	0.225 (0.125)
Hispanic	-0.075 (0.206)	0.078 (0.169)	0.313 (0.195)	0.072 (0.194)	-0.390 (0.203)
Other Race	0.189 (0.240)	0.021 (0.180)	0.190 (0.230)	-0.222 (0.216)	-0.178 (0.222)
Parental Ed < BA	0.273 (0.095)**	-0.169 (0.083)*	-0.010 (0.086)	-0.066 (0.089)	-0.026 (0.086)
Parental Ed = BA	0.144 (0.083)	0.063 (0.075)	-0.037 (0.079)	-0.118 (0.084)	-0.061 (0.078)
Survey Version 1	-0.393 (0.099)*	0.124 (0.090)	0.095 (0.094)	0.381 (0.102)**	-0.191 (0.092)*
Survey Version 2	-0.040 (0.104)	0.191 (0.090)*	-0.359 (0.095)**	0.217 (0.102)*	-0.011 (0.092)
Survey Version 3	0.000 (0.103)	0.026 (0.091)	0.061 (0.095)	0.142 (0.101)	-0.215 (0.095)*
R-Squared	0.042	0.019	0.099	0.039	0.013
N	1499	1499	1499	1499	1499

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, students with a parent who has a graduate degree, and students who received survey version four.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.9. Replication of Main Findings Using Students Who Report Belonging to Just One Racial/Ethnic Category (Comparable to Figure 2, top row).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.345 (0.060)**	0.141 (0.055)*	0.661 (0.057)**	-0.392 (0.060)**	-0.054 (0.056)
Asian	0.256 (0.139)	0.262 (0.131)*	-0.306 (0.134)**	-0.132 (0.136)	0.100 (0.135)
Black	0.387 (0.136)**	-0.122 (0.113)	-0.372 (0.111)**	0.033 (0.119)	0.063 (0.109)
Hispanic	-0.284 (0.186)	0.124 (0.168)	0.394 (0.186)*	0.119 (0.206)	-0.364 (0.178)*
Other Race	0.597 (0.312)	0.023 (0.230)	-0.523 (0.322)	0.083 (0.285)	-0.180 (0.284)
Parental Ed < BA	0.295 (0.080)**	-0.176 (0.070)*	-0.112 (0.073)	-0.051 (0.077)	0.057 (0.073)
Parental Ed = BA	0.141 (0.069)*	-0.033 (0.064)	-0.132 (0.066)*	-0.052 (0.070)	0.067 (0.065)
Survey Version 1	-0.318 (0.083)**	0.105 (0.077)	0.188 (0.080)**	0.297 (0.085)**	-0.258 (0.078)**
Survey Version 2	-0.136 (0.085)	0.280 (0.078)	-0.298 (0.081)**	0.190 (0.085)*	-0.035 (0.078)
Survey Version 3	-0.011 (0.089)	0.042 (0.077)	0.077 (0.081)	0.134 (0.084)	-0.219 (0.081)**
R-Squared	0.040	0.016	0.091	0.026	0.011
N	2106	2106	2106	2106	2106

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, students with a parent who has a graduate degree, and students who received survey version four.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.10a. Replication of Main Models, Survey Version 1 Only (Comparable to Figure 2, top row).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.476 (0.111)**	0.317 (0.104)**	0.790 (0.108)**	-0.430 (0.117)**	-0.170 (0.105)
Asian	0.530 (0.285)	0.361 (0.252)	-0.327 (0.251)	-0.360 (0.250)	-0.008 (0.230)
Black	0.372 (0.220)	0.043 (0.020)	-0.222 (0.178)	-0.390 (0.220)	0.194 (0.175)
Hispanic	-0.314 (0.259)	0.242 (0.299)	0.552 (0.236)*	0.257 (0.249)	-0.734 (0.264)**
Other Race	-0.100 (0.032)	0.139 (0.304)	-0.328 (0.337)	-0.024 (0.293)	0.315 (0.307)
Parental Ed < BA	0.310 (0.156)*	-0.241 (0.144)	-0.258 (0.143)	-0.063 (0.159)	0.221 (0.140)
Parental Ed = BA	0.108 (0.124)	-0.018 (0.119)	-0.144 (0.124)	-0.109 (0.133)	0.130 (0.124)
R-Squared	0.056	0.026	0.108	0.038	0.025
N	557	557	557	557	557

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, and students with a parent who has a graduate degree. The factors on survey version 1 were listed to students in the following order: (A) Field Interest, (B) Course Success, (C) Career Stability, (D) Career Fulfillment, and (E) Career Salary.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.10b. Replication of Main Models, Survey Version 2 Only (Comparable to Figure 2, top row).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.190 (0.119)	0.087 (0.107)	0.688 (0.112)**	-0.341 (0.117)**	-0.247 (0.107)*
Asian	0.086 (0.240)	-0.127 (0.241)	0.177 (0.257)	-0.168 (0.242)	0.029 (0.213)
Black	0.457 (0.246)	-0.274 (0.217)	-0.366 (0.199)	0.151 (0.176)	0.028 (0.186)
Hispanic	0.077 (0.300)	0.127 (0.245)	0.526 (0.297)	-0.047 (0.299)	-0.688 (0.302)*
Other Race	0.247 (0.352)	-0.083 (0.205)	0.326 (0.320)	0.028 (0.359)	-0.521 (0.329)
Parental Ed < BA	0.342 (0.149)*	-0.204 (0.137)	-0.073 (0.144)	-0.014 (0.144)	-0.044 (0.135)
Parental Ed = BA	0.377 (0.136)**	-0.145 (0.125)	-0.262 (0.128)	-0.076 (0.140)	-0.046 (0.126)
R-Squared	0.028	0.010	0.080	0.018	0.027
N	566	566	566	566	566

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, and students with a parent who has a graduate degree. The factors on survey version 2 were listed to students in the following order: (A) Career Salary, (B) Field Interest, (C) Career Fulfillment, (D) Course Success, and (E) Career Stability.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.10c. Replication of Main Models, Survey Version 3 Only (Comparable to Figure 2, top row).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.327 (0.123)**	0.078 (0.110)	0.629 (0.113)**	-0.429 (0.115)**	0.044 (0.115)
Asian	0.337 (0.250)	0.133 (0.228)	-0.422 (0.236)	0.105 (0.235)	-0.022 (0.237)
Black	0.257 (0.230)	0.092 (0.192)	-0.426 (0.201)*	0.190 (0.192)	-0.134 (0.222)
Hispanic	0.030 (0.400)	0.184 (0.289)	0.396 (0.314)	-0.102 (0.345)	-0.174 (0.334)
Other Race	0.216 (0.514)	-0.432 (0.331)	-0.325 (0.440)	0.120 (0.406)	0.397 (0.399)
Parental Ed < BA	0.200 (0.163)	-0.045 (0.136)	-0.040 (0.140)	0.044 (0.147)	-0.112 (0.153)
Parental Ed = BA	-0.013 (0.143)	0.032 (0.124)	-0.075 (0.132)	-0.038 (0.135)	0.069 (0.130)
R-Squared	0.023	0.006	0.069	0.028	0.006
N	561	561	561	561	561

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, and students with a parent who has a graduate degree. The factors on survey version 3 were listed to students in the following order: (A) Course Success, (B) Career Fulfillment, (C) Career Stability, (D) Career Salary, and (E) Field Interest.
**/* indicates statistical significance at the 1-percent/5-percent level.

Appendix Table B.10d. Replication of Main Models, Survey Version 4 Only (Comparable to Figure 2, top row).

	Rank Value Models (Comparable to the Top Row of Figure 2)				
	Field Interest	Career Stability	Career Salary	Career Fulfillment	Course Success
Female	-0.456 (0.118)**	0.129 (0.106)	0.669 (0.111)**	-0.481 (0.118)**	0.139 (0.109)
Asian	0.139 (0.247)	0.530 (0.236)*	-0.521 (0.222)*	0.023 (0.246)	-0.170 (0.268)
Black	0.285 (0.262)	-0.176 (0.194)	-0.155 (0.240)	0.048 (0.268)	-0.002 (0.227)
Hispanic	-0.309 (0.313)	-0.002 (0.289)	0.125 (0.343)	0.121 (0.332)	0.065 (0.215)
Other Race	-0.140 (0.325)	0.063 (0.262)	0.660 (0.238)*	-0.433 (0.322)	-0.152 (0.308)
Parental Ed < BA	0.210 (0.151)	-0.159 (0.133)	-0.039 (0.142)	-0.184 (0.148)	0.173 (0.139)
Parental Ed = BA	0.113 (0.136)	0.080 (0.127)	0.020 (0.132)	-0.252 (0.139)	0.039 (0.127)
R-Squared	0.034	0.019	0.079	0.038	0.008
N	556	556	556	556	556

Notes: Each column shows results from an independently estimated regression. Robust standard errors are in parenthesis. The omitted groups are men, white students, and students with a parent who has a graduate degree. The factors on survey version 3 were listed to students in the following order: (A) Career Fulfillment, (B) Career Stability, (C) Career Salary, (D) Field Interest, and (E) Course Success.
**/* indicates statistical significance at the 1-percent/5-percent level.