Pipeline Dreams? Gender Differences in Occupational Plans and STEM Major Completion among a Recent Cohort of US College Entrants

Kim A. Weeden (Cornell University)**, Dafna Gelbgiser** (Cornell University),

and **Stephen L. Morgan** (Johns Hopkins University)

NOTE: The supplementary tables to which we refer in the text are available at http://www.kimweeden.com/manuscripts/.

Abstract

We use data from the Educational Longitudinal Survey (2002-2012) to document and examine the sources of the gender gaps in the probability of earning a BA in either a STEM or biomedical field or a health field. Even among this recent cohort, gender differences in the completion of these two types of science-related degrees are substantial, and they are strongly predicted by gender differences in the content and stability of occupational plans, but not by gender differences in prior academic achievement and math and science preparation, family-work orientation, or self-assessed math ability. Attrition from STEM/Biomed majors is also gendered: women who had declared STEM/Biomed majors as sophomores were more likely to leave without a degree than they were to graduate with a degree in STEM, the reverse pattern as men. Among the subset of college sophomores who declared STEM/Biomed degrees, gender differences in academic achievement attenuate the gender gap in STEM persistence (because of STEM women's higher GPAs), self-assessed math ability and family-work orientation have no effect on the gender gap, and gender differences in occupational plans account for between 10 and 20% of the gender gap in STEM persistence.

Women's underrepresentation among graduates in science, technology, engineering, and mathematics (STEM) and related fields remains a crucial form of inequality in higher education, contributes to women's underrepresentation in STEM-related occupations and to differences in labor market rewards that are tied to occupations (e.g., earnings, autonomy, security), and, from the perspective of workforce development, exacerbates the shortage of scientifically trained workers needed to feed the American economy. Despite a large and multidisciplinary research literature on "women in STEM," much debate and uncertainty remains about why women are less likely to enter STEM educational pathways, persist in them, graduate from college with a STEM or related degree, continue on to post-baccalaureate training in STEM, or enter STEM-related occupations after completing their education.

This paper examines the patterns and sources of gender differences in the persistence, dropout, and late college entry into STEM and other science-intensive majors, including doctoral-level biomedical degrees (e.g., pre-med, pre-veterinary degrees) and degrees that lead to MA-level health occupations. It assesses the impact of three common explanations for gender differences STEM major completion in the "supply side" theoretical and empirical literature on gender in STEM: (1) socioeconomic status and family background; (2) academic performance, standardized test performance, and math and science coursetaking in high school; (3) family-work orientation; and (4) math self-assessment. It also examines a less commonly studied, if often implied, predictor of gender differences in STEM major completion: the specific content and stability of student's occupational plans (but see Morgan, Gelbgiser, and Weeden 2013).

We make three main contributions to the extant literature on STEM in higher education. First, we extend and elaborate Morgan et al's 2013 analysis of the predictors of the gender gap in early college STEM major selection by (1) using newly released data to assess the gender gap in STEM major completion, and (2) broadening the scope of predictors to include self-assessed math ability, which has received considerable attention in the social psychological literature on college major and occupational choice (see, e.g., Correll 2001, 2004). Second, we describe gender differences in mobility into and out of STEM majors between sophomore year in college and college completion (or, in some cases, non-completion), which allows us to better pinpoint the timing of "leaks" from the STEM education pipeline. And, third, we assess how well the aforementioned predictors of STEM major completion can account for gender differences in STEM major persistence, conditional on declaring a STEM major as a college sophomore.

We also depart from much of the extant literature by differentiating, where data allow, between two types of scientific majors: STEM/Biomed, which includes "pre-med" and related majors that are educational pathways to doctoral-level medical occupations; and health majors (e.g., nursing, physical therapy), which are educational pathways to medical occupations that typically require a MA-degree or less. Although the latter group is often excluded from analyses of the patterns of sources of women's underrepresentation in scientific fields, we think that this decision fails to appreciate the scientific content and training entailed in these majors, and, from a policy perspective, their importance for workforce development. Where data allow, we will treat STEM/Biomed and Health majors as separate categories, in acknowledgement that these two types of scientific majors have quite different gender profiles, that they may draw quite different types of students with different occupational aspirations, and, more pragmatically, that some readers may prefer a narrower definition of STEM degrees.

The paper relies on the recently released 2012 wave of the Educational Longitudinal Survey (ELS), a nationally representative survey of a cohort of young men and women who were first surveyed in 2002, when they were in 10th grade, and again in 2004, 2006, and

2012. The ELS cohort is one of the first cohorts in which members began their schooling after women started to exceed men in college enrolment and during an historical moment when substantial public and policy attention focused on increasing women's participation in STEM (e.g., NSF's ADVANCE program).

To foreshadow our results, we find substantial gender differences in science-related major completion among the ELS cohort, the strongest predictor of which is the content and stability in occupational plans. We also find significant gender differences in STEM/Biomed persistence: women who declare STEM/Biomed degrees as sophomores are more likely to leave their first institution without obtaining any degree in any field than they to complete their STEM/Biomed degree, precisely the reverse pattern as men. Although more women than men enter STEM/Biomed majors after sophomore year, primarily through "upgrading" from health majors, this late-career entry into STEM/Biomed does not offset women's greater attrition from STEM/Biomed in absolute terms. Finally, we show that among the subset of college sophomores who major in STEM/Biomed in their sophomore year, occupational plans are again the strongest predictor of gender differences in the probability of completing a STEM/Biomed degree. The effect of occupational plans persists in models that condition on family-work orientation, self-assessed math ability, pre-college academic achievement and test scores, and high school math and science coursestaking; this finding implies that occupational plans are not reducible to these factors.

The rest of our paper adopts a slightly different structure than usual. First, we discuss the data set, our measure of college outcomes, and the distribution of men and women across these outcomes. We next discuss the logic and prior research motivating each of our predictors of gender gaps in STEM major completion, how the relevant variables are constructed, and men and women's distribution on them. We then fit a series of multinomial logit models that allow us to estimate the minimum and maximum change in the gender gap in the probability of completing a STEM/Biomed degree, a health degree, a major in some other field, an Associate degree or certificate, or failing to complete a degree. In the second half of the paper, we focus on STEM persistence and mobility, first by presenting inflow and outflow tables, and then, to assess the predictors of STEM persistence, by re-estimating our logistic regression models on the subset of respondents who had declared a STEM major by their sophomore years. We conclude with a brief discussion of the implications of our findings for strong-form preference theories of gender differences in educational decisions.

Data and measures

The ELS is a nationally representative, two-stage sample of schools and students. ELS instruments include questionnaires administered to students, parents, teachers, and school administrators and, in some waves, a supplementary analysis of student transcripts. Further details of the ELS data collection procedures and sampling frame are available on the data distributors' website. Although most of our variables come from the publicly released ELS panel data from 2002 and the 2004, 2006, and 2012 follow-up waves, our measure of occupational plans (see below) is extracted from a restricted access version of the ELS.

We focus on the subset of ELS respondents who entered a 4-year baccalaureate institution within a year of graduating from high school (hereafter "traditional college student"), who participated in all four waves of the survey, and who have non-missing information on college major and type of degree earned. We examine these students' college outcomes as of the 3rd follow-up wave in 2012, by which point most had either completed baccalaureate degrees and/or entered the labor force.[1]

We weight the data by the base-year and third follow-up panel weight developed by the

[1] Our sample differs slightly from the "traditional college student" sample in Morgan et al (2013), in that (1) we lose additional students to attrition between the 2nd and 3rd followups, and (2) we include, and Morgan and colleagues exclude, students who had not declared a major by 2006, the 2nd followup.

data distributors, multiplied sequentially by two estimated inverse probabilities that account for (1) non-participation in all four waves of the survey and (2) non-response on the dependent variable of the relevant model (major and type of degree earned). The estimated probabilities for (1) and (2) were drawn from separate logit models that predict inclusion in the relevant restricted sample with demographic characteristics, family background, and base-year indicators of academic engagement. The weighted analytic sample contains 5,159 observations, 2,353 men and 2,806 women.

College Major in 2006 and 2012

The outcome variable in our models is based on the respondents' self-reported major in which he or she received a baccalaureate degree. In the tables we present in the main paper, we impose the further condition that the degree was awarded from the first institution that the respondent attended and was completed by 2012, which for the traditional college students is ELS is approximately eight years after first enrolling in a 4-year institution. In a supplementary online appendix, we present an analogous set of tables that examine college outcomes from any institution, regardless of whether it was the institution the student entered in 2004 or a different institution. Our decision to focus on the first institution reflects our assumption that students who transfer institutions are different, in both observed and unobservable ways, from students who do not transfer institutions.

Based on the detailed CIP codes provided by the data distributors, we coded the field of the BA degree into one of three categories: STEM and biomedical degrees that lead to doctoral level occupations (e.g., pre-med, pre-vet), health and related degrees that lead to MA-level occupations (e.g., nursing, physical therapy), and all other majors. It was not possible to differentiate biomedical degrees from other STEM degrees, because many institutions offer "pre-med" programs within disciplinary degrees such as Biology. We also include categories for students who obtained an AA degree or certificate (but not a BA) from the 4-year institution in which they enrolled after high school, and for students who dropped out of their first institution without receiving any kind of degree.

We also coded initial college major from the 2006 ELS follow-up wave, when most of the respondents in our analytic sample were sophomores in college. The first three categories (STEM/biomed, health, and other) are coded the same as the field of the BA degree. Valid categories for initial college major also include “Undeclared” and “Missing.”

Table 1 presents the distributions, by gender, of 2012 college outcomes (Panel A) and 2006 college major (Panel B) for the weighted ELS sample. It shows, firstly, that gender differences in college major choice are substantial in this very recent cohort of college entrants. The percentage of young men who declared a STEM/Biomed major in the 2006 wave is more than double the percentage of young women (24.3% vs. 11.0%, respectively; see Table 1, Panel B). Conversely, 11.7% of young women initially selected a health or related major, three times the analogous percentage of young men (3.5%).

Similar ratios also characterize the distribution of young men and women across college majors as of 2012, although the absolute percentage of students majoring in STEM/Biomed declined for both sexes (see Panel A). As of 2012, 15.5% of men completed a STEM/Biomed major from their first institution, more than twice the percentage of women (6.8%). Considering outcomes from all institutions, these percentages increase to 18% and 8%, respectively – still more than a 2:1 ratio (see supplementary on-line tables). Turning to health majors, 4.4% of women completed a BA degree in health or a related field from their first institution, compared to 1.8% of men. These percentages increase to 6.4% and 2.7%, respectively, when we include degrees from second or later institutions, but the gender gap itself is quite similar.

Pre-College Academic Preparation and Math and Science Coursetaking

A common explanation for gender differences in STEM major completion traces it to gender differences in academic achievement and preparation, including standardized test scores, grades, and math/science coursetaking in high school (e.g. Turner and Bowen 1999; Ayalon 2003; Hyde et al. 2008). The argument has two parts. First, although women's *mean* grades and test scores exceed (in the case of GPA) or comparable to (in the case of standardized test scores and coursetaking) men's, women are underrepresented among the students who leave high school with the greatest academic preparation in science and math. Second, these gender differences account for women's lower propensity to choose and stay in majors that require substantial math skills or prior academic preparation in math or science; or, conversely, their greater propensity to choose and stay in majors that require greater reading or verbal skills.

We measure academic preparation using data collected in the 2002 and 2004 waves of the ELS. (The college transcript information was released the day before we submitted this paper.) We are able to include math scores in 10^{th} and 12^{th} grade, reading scores in 10^{th} grade, 12^{th} grade GPA, and 6-category measures of both math coursetaking and science coursetaking (following Burkham et al 2003). Reading test scores were not collected in 12^{th} grade.

The gender distributions on these predictors are presented in Table 2. Relative to young men, young women who enter college immediately after high school have higher mean GPA and less variance in GPA, higher average reading test scores, lower average math test scores, are less likely to have taken calculus, and are less likely to have taken both chemistry and physics in high school. However, young women are more likely to have chemistry *or* physics, and also more likely to have taken pre-calculus, than young men.

Family-Work Orientation

A second set of arguments focuses on gender-differentiated preferences for high involvement in child-rearing and other time-consuming family activities coupled with the perceived or actual incompatibility between science and family life. The preferences themselves may reflect ongoing gender-differentiated socialization and normative pressures on women to conform to sex-typical work and family roles (e.g., Jacobs 1989; Hakim 2002) or deep-rooted or even innate and biologically based preferences of women for caring, nurturing, and interacting (see, e.g, Ceci and Williams 2012; Williams and Ceci 2012, 2015). Both versions anticipate substantial differences between young men and young women in their expressed preferences for family over work. They also anticipate that these differences in family-work orientation will predict men's overrepresentation in STEM/Biomed majors, whether because these fields are more "object-oriented" or because they lead to careers that are not perceived to be family friendly.

Our measure of family-work orientation replicates, to the extent possible, the "family vs. work attitude scale" proposed and used by Xie and Shauman (2003) in their analysis of the NELS data, the precursor to ELS. As high school seniors, all ELS respondents were asked to rate 18 separate items in response to the question, "How important is each of the following to you in your life?." Following Xie and Shauman (2003), we use four of these items (see Table 3) to create a scale of family-work orientation that sums the two items pertaining to family, subtracts the two items pertaining to work, and standardizes to mean 0. For respondents who were missing valid information on one or more of the component items, we imputed the overall scale scores based on the available items and demographic characteristics. In supplementary analyses (not shown), we fit each individual measure of family-work orientation separately, rather than a scale, and obtained nearly identical results.

The distributions of responses on the four component items and scale are presented in Table 3, by gender. Among traditional college students in the ELS, we find modest gender differences in preferences for work or family. Nearly all young men and young women think that it's very important to be successful in one's line of work (92.8% and 94.2%, respectively). Similarly, a supermajority of young men and young women think that it's very important to marry right and have a happy family life (83% and 85.2%, respectively), although we cannot discern what "marrying right" and "happy family life" means to the respondents.[2] A higher percentage of men (37.6%) than women (23.2%) think that it is very important to have lots of money, and a higher percentage of women (55.8%) than men (47.1%) think it is very important to have children. The gendered responses to the latter questions are responsible for the observed gender differences in the family-work scale (see Panel B, Table 3).

Math Self-Assessment

A key finding in the social-psychological research on women in STEM is that women tend to assess their math ability lower than men, even conditional on standardized math test scores and other observed measures of academic ability(see, e.g., Correll 2001, 2004). These self-assessments of competence affect students' decisions about whether or not to enter STEM and, once in, stay. The implication is that some of the gender gap in STEM/Biomed completion is driven by gender differences in self-assessed math ability.

A related argument, which is more prevalent in the psychological literature, is that young men and women use myriad information, including academic achievement and cues about their abilities from significant others, to develop an identity, or "self-concept" as someone who is good in math or science or someone who is not (e.g., Eccles 2011). Although

[2] As we will discuss below, we ran supplementary models in which we allowed the effect of each predictor on our measure of college major outcomes to vary by gender. If the meaning of "marrying right" or "having a happy family life" differs systematically by the gender of the respondent, relaxing the equality constraint on the family-work orientation covariate should increase the estimated contribution of family-work orientation on the gender gap in STEM/Biomed major completion. It doesn't.

there are subtle differences this “expectancy value” model and math self-assessment, in practice they tend to be measured with similar questions, especially in wide-ranging surveys such as the ELS.

We develop a factor variable that measures self-assessed math ability from a set of five component items in the 2004 (12^{th} grade) wave of the ELS, which is the most recent survey in which these items were asked. The items, which are listed in Table 4 along with their responses by gender, ask students questions such as “I can understand difficult math texts” and “I can do an excellent job on math assignments.” From these five questions, we computed a factor score, standardized with mean 0 and standard deviation of 1. We imputed scores for respondents who had missing values on one or more of the five questions from demographic characteristics and the questions for which we have valid information.

The responses on the individual items and factor scores are presented in Table 4. They show the predicted patterns anticipated by social psychological theory and by empirical research using older data (e.g., Correll 2001). A substantially greater percentage of young men than young women respond that they can almost always understand difficult math texts (22.5% vs. 12.4%) and math classes (26.6% vs. 15.6%), master math (35.3% vs. 28.2%), and do an excellent job on math tests (27.0% vs. 19.7%). Gender differences in the remaining item, “can do an excellent job on math assignments,” are weaker: 34.3% of young men, compared to 31.2% of young women, respond with “almost always.” Young men have an average value of 0.41 (standard deviation =1.14) on the factor score, compared to 0.05 for young women (standard deviation=1.08).

Occupational Plans

The final predictor of STEM major completion we consider is occupational plans. In much of the theoretical literature on gender differences in STEM major choice, gender differences in occupational plans are implicit. For example, in Eccles’ expectancy value model, gendered socialization and normative pressures for young women to prioritize family over work will generate gender differences in plans to enter STEM or related scientific occupations. Similarly, strong-form preference models such as Hakim (20022) and Ceci and Williams (2012; Williams and Ceci 2012, 2015) claim that women self-select out of STEM because of their innate, or at least immutable preferences for family over work, under the assumption that family and scientific careers are incompatible.

These arguments predict that the association between occupational plans and college major completion will be eliminated or greatly reduced in models that fit measures of prior academic preparation and achievement, family-work preferences, and self-assessed math ability. However, in research on initial college major selection, Morgan and his colleagues (2013) found strong residual association between occupational plans and STEM/Biomed major choice in models that adjust for prior academic achievement and family-work orientation (although not self-assessed math ability), suggesting that occupational plans are not simply endogenous to family-work orientation or academic achievement.

Morgan et al (2013) do not specify an alternative mechanism by which occupational plans have an independent effect on gender gaps in STEM/Biomed or health majors, but are anticipated by Morgan's (2001, pp 101-102) model of educational decision-making. In this model, students make "prefigurative commitments" to specific future courses of behavior (e.g., entering a science occupation). These commitments guide intermediate decisions (e.g., college major choice) that students believe, possibly erroneously or with uncertainty, will help them attain that future outcome. The prefigurative commitments themselves are a function of students' beliefs about the costs and benefits of a specific future outcome, imitative processes (e.g., what my peers do), and normative processes (e.g., what people like me do). All of these processes could be patterned by gender – for example, if women have a more difficult time envisioning themselves in a STEM occupation because of extant patterns of gender segregation, or if they have few same-sex friends who are also in STEM, on the importance of being assigned a STEM roommate in freshman year), and so on. Family-work orientation may be one factor that shapes prefigurative commitments, but it is not the only factor. The empirical implication is that gender differences in occupational plans will predict gender differences in STEM major completion, even in models that condition on family-work orientation, academic achievement and preparation, and math self-assessment.

Our measures of occupational plans are based on questions in the 2004 and 2006 ELS self-administered student questionnaires that instructed respondents to “Write in the job or occupation that you expect or plan to have at age 30.” Students could write in a response, select “you don’t know,” or skip the question. The coding of this variable in the public use data does not allow researchers to identify students with science-related occupational plans, so we coded the verbatim responses, which are available in the metadata to approved users. Our coders coded the verbatim responses from 2004 and 2006 (separately) into 1,220 distinct occupational categories elaborated from the 2000 Standard Occupational Classification. ELS respondents could list multiple occupations, and our coders coded all responses. This allows us to identify students with uncertain plans that combine STEM and non-STEM jobs.

We aggregated the coding of the verbatim responses into occupational plans variables for 2004 and 2006, each of which captures qualitatively different types of planned occupations. The categories are: STEM and doctoral level biomedical occupation (e.g., physicist, engineer, doctor, veterinarian); masters level health (e.g., nurse, therapist); other occupations; a mixture of a science-related occupation and something else; “don’t know”; and “missing.” We were forced to combine STEM and doctoral-level biomedical occupations because of sparse cell counts, a decision that mutes gender differences in occupational plans given women are more likely to plan to be a doctor or veterinarian than a physicist or engineer. In some models, the mixture category was also too sparsely populated, so we instead assigned students to the main category of their science-related occupation choice. "Don't know" is a substantively interesting category in itself, reflecting uncertainty in occupational plans, and the "missing" category is a heterogeneous mix of students who were not asked the question, students who answered or whose answers were transcribed with an illegible response, "joke" answers (e.g., "drug dealer," which we refrained from coding as science, health, or retail sales), and related types of responses.

We also created a binary measure of whether the respondents' plans changed between 2004 and 2006. Our reasoning is that students whose commitment to pursuing STEM is more deeply rooted will be more likely to remain in a STEM (or biomedical/MD, or health) major than students whose plans are more transitory. We estimate both the main effect of this "plan persistence" variable and its interaction with 2006 plans. (We also tested models with 2004 plans and a 2006 persistence variable, with nearly identical results. We lack sufficient cases to interact the 2004 and 2006 plans variables.)

The distribution of responses by gender on three occupational plans questions – 2004 plans, 2006 plans, and persistence in plans – is presented in Table 5. In 2004, 26.5% of young men and 13.3% of young women planned to enter STEM or biomedical-MD occupations. By 2006, these percentages had declined to 20.0% and 10.0%, respectively, but the same two to one ratio persists. The percentage of young men and young women who plan to enter MA-level health occupations, by contrast, was essentially unchanged across the two waves, at 4% for young men and 15.2% (2004) and 16.1% (2006) for young women. A small but interesting share of young men and women who listed a STEM plan and a non-STEM occupational plan ("mixture") in each wave, and another fifth to quarter of students listed "don't know." The data show modest gender differences in occupational plan persistence: 46.7% of young men changed their occupational plans between 2004 and 2006, compared to 42.8% of young women.

Adjustment Variables

Our models also include a comprehensive set of variables that capture demographic characteristics and socioeconomic background of the ELS students. Our measures of these attributes come from the 2002, 2004, and 2006 waves of the ELS, and are listed in Table 6. We use item-specific best subset linear regression to impute information for missing adjustment variables.

Gender differences in demographic and family background characteristics are modest, and are likely to have correspondingly modest effects on the gender gap in STEM/Biomed and Health major completion. Still, they are standard in models of college attainment, and for completeness we include them here. **Analytic Strategy**

In the first set of analyses, we estimate a series of nested multinomial logit models in which the outcome variable is Major of the BA degree and the reference category is "Other [non-STEM] degree." The first model estimates the raw gender gap in the probability of obtaining one of the Major categories (including no degree and AA/certificate), and replicates the gender gap shown in Table 1. The second model fits the demographic and socioeconomic background adjustment variables, and serves as our baseline for the subsequent models.

Models 3a-3d add sets of related predictors (e.g., all measures of prior academic achievement) to the baseline model. These models are used to estimate the *maximum* portion of the gender gap in the probability of falling into the outcome category that are attributable to family-work orientation, performance and prior coursetaking, math self-assessment, and occupational plans. Model 4 fits all predictors simultaneously. Models 5a through 5d take away each set of predictors from Model 4, and are used to estimate the *minimum* proportion of the gender gap in the outcome attributable to that set of predictors. This decomposition strategy reflects the absence of an agreed-upon causal model to guide the decision of which of the sets of predictors has temporal primacy (see also Morgan et al 2013).[3]

Panel A of Table 7 presents predicted probabilities that men and women will earn BA degrees in particular majors (or, alternatively, to receive AA/certificates or drop out) under various model assumptions, as well as the gender gap in these probabilities. Panel B presents our minimum and maximum estimate of the percentage change in each gender gap in marginal probabilities, as described above. The coefficients on which these predicted probabilities are calculated are relegated to Appendix Table A.

In the second set of analyses, we assess gender differences in STEM persistence, attrition, and late entry between 2006, when most of the ELS respondents were sophomores in college, and 2012, by which point most had finished their degrees. We begin this analysis with simple, gender-specific "mobility" tables that cross-classify 2006 college major, including "undeclared," by 2012 college outcome.

We then focus on STEM persistence with a series of multinomial logit models that have a similar structure as those in Analysis 1, but that are fit to data that have been stratified by initial college major (measured in 2006). These models allow us to assess the minimum and maximum contribution of each of our set of predictors to gender gaps in STEM/Biomed persistence. We lack sufficient cases to conduct analogous analyses of persistence in health majors or of late-career entry into STEM/Biomed from other 2006 college major categories.

Sources of Gender Differences in STEM major completion

The unadjusted gaps in Model 1 (see line 1, Panel A, Table 7) replicate the raw gaps in college major completion presented in Table 1, and need not be discussed again here. The

[3] In the results presented in the main tables, we constrain the effects of each set of predictors to be the same for young men and young women. In supplementary analyses, we relax these constraints and find, with one exception noted below, no substantively important differences in the minimum and maximum effects of any of the predictors on the gender gap in STEM/Biomed or Health major completion.

SES-adjusted gaps in STEM/Biomed and Health majors do not change appreciably: the gender gap in STEM/Biomed, for example, declines from 0.0878 to 0.0854 after adjusting for family socioeconomic status. (The gender gaps in the probability of majors in other fields, AA degrees, or no degrees change more in percentage terms, but off a base of essentially no gender gap.)

We find that very little of the gender gap in STEM/Biomed major completion can be accounted for by either prior academic performance or by family-work orientation. Prior academic preparation accounts for between 1% and 15% of the gender gap in STEM/Biomed major degree completion. It accounts for much more of the gender gap in Health in percentage terms (28-43%), but we note that this is off a much smaller base: the SES-adjusted gender gap (favouring women) in men and women's probability of earning a BA in health is 0.0257 points, compared to a gender gap of 0.0854 (favouring men) in STEM/Biomed.

Family-work orientation contributes has no impact on either the gender gap in STEM/Biomed degree attainment or Health major attainment. The weak contribution of gender differences in prior academic preparation and family-family orientation to gender gaps in STEM attainment is consistent with prior sociological research (e.g., Xie and Shauman 2003; Mann and DiPrete 2013; Morgan et al 2013; Riegle-Crumb et al 2012, Perez-Felkner et al 2012). This finding is worth reiterating, given these "zombie theories" refuse to die in the popular press and in some corners of academia (e.g., Ceci and Williams 2012).

Prior research does not anticipate the very modest contribution of gender differences in self-assessed math ability to the gender gap in STEM/Biomed or MA/Health attainment (see Table 7). For example, gender differences in math self-assessment account for between 2% and 19% of the gender gap in STEM/Biomed completion. The maximum contribution, 19%, assumes that math test scores are endogenous to math self-assessment, and gives theoretical primacy to the latter. This is an extremely generous test of the gendered self-assessment claim, which explicitly notes that gendered math self-assessments occur even conditioning on test scores. In supplementary models (not shown), we fit a variant of Model 3c that conditions on prior achievement as well as math self-assessment, and compare it to the gender gap in Model 3b (i.e., conditioning on SES and prior achievement); this yields a minimum and maximum math self-assessment effect of 1% and 6%, respectively. In the ELS cohort, gender differences in math self-assessment have a modest impact on gender differences in STEM/Biomed major completion.

Table 7 also shows that gender differences in occupational plans account for between 21% and 33% of the gender gaps in STEM/Biomed and MA-level health major completion. Put differently, the minimum estimate of the occupational plans "effect" matches or exceeds the maximum estimated effect for any other factor. Gender differences in occupational plans also account for nearly three quarters of the gender gap in MA-level health major completion, although recall that these percentages are calculated from a small gender gap (and cell counts are quite small).

The contribution of occupational plans is reduced only modestly in models that adjust for family-work orientation, academic preparation and math/science coursetaking, and math-self assessment. For example, in our estimate of the gender gap in STEM/Biomed major completion, the inclusion of these predictors reduces the contribution of occupational plans by 12 percentage points, or slightly more than 1/3. Although our goal in this paper is to focus on the consequences of gender gaps in occupational plans, not their sources, our results suggest that no more than 1/3 of the effect of gender differences in occupational plans on STEM major completion is due to the association – causal or otherwise – between family-work preferences, self-assessed math ability, or academic achievement and preparation in high school.

Patterns of STEM & Health Major Persistence, Attrition, and Late-Career Entry

The preceding findings offer important insight into the sources of the gender differences in the attainment of a college major in STEM or a related health field, but not into gender-specific patterns of persistence in, attrition from, or late college entry into STEM and related majors throughout the college career. Our second set of results turn to these differences.

We begin with simple “mobility tables” of initial major selected in 2006 (“origins”, or rows) by college outcome by 2012 (“destinations,” or columns), for men (Panel A) and women (Panel B). Table 8a provides these data as outflow tables, where the percentages in the cells sum to 100 in each row, and Table 8b shows the same data as an inflow table, where the percentages in the cells sum to 100 in each column. (Raw data are given in Appendix B.)

Table 8a shows substantial attrition from the three substantive major categories in 2006 (STEM/Biomed, Health, and Other) for both men and women, reflecting major switching as well as attrition from the first baccalaureate institution attended. This attrition differs by gender. For example, among men who majored in STEM/Biomed in 2006, more than half (51.9%) completed a STEM/Biomed degree.The next most prevalent outcome was “No Degree,” capturing 36.6% of the 2006 STEM/Biomed origin category. Among young women who declared STEM/Biomed majors in 2006, however, nearly half (48.5%) received “No Degree,” and only 38% completed their STEM/Biomed degrees. Put differently, it was more likely for a female STEM sophomore to leave college than to complete her STEM degree, whereas the reverse is true for men.[4]

Attrition from 2006 health majors was also substantial, but the gender differences comparatively modest. Only 26.7% of young men and 25.6% of young women who declared

[4] Table 8a in the online supplementary tables (see www.kimweeden.com) shows that persistence in STEM is greater when we include degrees attained from any institution, not just the first institution attended. However, the gender differences are still substantial: 58.2% of young men in STEM/Biomed in 2006 finished a STEM degree at any institution, compared to 42.5% of women.

a health major in 2006 completed a BA degree in a health major. Another 7.3% of women completed Associate degrees or certificates, some of which were likely in health-related fields (e.g., an Associate-level nursing degree), but only 2.5% of men.. Attrition from the first institution without obtaining a degree of any kind was highest for young men and women who had missing values on 2006 major, and for young men who had not declared a major by 2006.

Table 8b shows gender differences in the sophomore year majors of the students who, by 2012, had earned STEM/Biomed, health, or other majors. It shows that late-college career entry into STEM/Biomed was relatively uncommon for the male students. Among the men who completed STEM/Biomed degrees, 81% had declared STEM/Biomed in 2006, and another 9.5% had not yet declared a major. By contrast, only 62.4% of young women who completed a STEM/Biomed major in 2012 had declared STEM/Biomed by 2006. The next largest origin category for women is health majors, suggesting some “upgrading” of science majors across the college career. In absolute terms, however, there are few women in health, or any other non-STEM category in 2006, who complete STEM/Biomed degrees, than men (see Appendix B).

Sources of Gender Differences in STEM/Biomed Persistence

Our final set of results addresses the question, “why are young women who initially declare STEM/Bomed less likely to persist in these majors than young men?” Table 9 presents descriptive statistics for the 891 ELS respondents who declared STEM/Biomed majors in 2006. Compared to the full analytic sample (see Table 6), these students’ parents have slightly greater education and occupational SEI scores, but for the most part their demographic and family background traits are similar to those in the full sample.

The college sophomores in STEM do differ significantly from the full sample of traditional college students on the other core predictors. The 2006 STEM/Biomed majors scored approximately 5 points higher (just under ½ a standard deviation) on the 12th grade math test, a greater share of them entered college with extensive math and science coursetaking experience (compare Table 9 and Table 2). The young women who declared STEM in 2006 have lower scores on the family-work orientation measure than the full sample (0.04 compared to 0.20), and show little difference on this measure from the young men (-0.02). The 2006 STEM/Biomed majors have higher math self-assessment than in the full sample (compare Table 9 with Table 4), but the gender gap in math self-assessment remains substantial at 0.31 points, or just under a third of a standard deviation, compared to 0.36 points in the full sample. The shares of men (57%) and women (54%) who plan to enter STEM or doctoral-level biomedical occupations are comparable (and much higher than in the full sample; see Table 5), but a substantially higher percentage of young women (18%) than young men (3%) plan to enter MA-level health occupations.

Table 10 presents the results of multinomial logit models that are analogous to those fit to the full sample(see Table 7), but estimated on the subsample of students who had declared a STEM/Biomed major in 2006. For this analysis, we were forced to combine health majors with other fields, and AA degrees/certificates with "no degree" because of small cell counts.

The results show, as in the full sample, no effect of family-work orientation on the gender gap in STEM major completion. It also has little association with the gender gap (favoring women) in the probability of completing a degree in a non-STEM field or the gender gap ("favoring" women) of not attaining a BA degree from the first institution. If there is support for the claim that the supply of women in STEM is due to adolescent and college student's choices to privilege family over work (see, e.g., Ceci and Williams 2012; Williams and Ceci 2012, 2015), it isn't found in the ELS data.

Gender differences in academic preparation and prior coursetaking have more pronounced effects on estimated gender differences in the probability of persisting in STEM than they did in the full sample, but in the reverse pattern than is predicted by a naïve, "women are worse at math so women drop out of STEM" hypothesis. Specifically, Table 10 shows that if the young men and women in STEM majors in 2006 had the same academic preparation and coursetaking experience in high school, the gender gap in STEM major persistence would be larger than under the observed distributions on these predictors, as indicated by the positive percentage change values in Panel B. An examination of the coefficients (see Appendix C) suggests that this is because high school GPA is the strongest predictor of STEM persistence among this select subset of students, and the young women who majored in STEM in 2006 have higher GPAs, on average, than the young men (see Table 9).

Gender differences in math self-assessment contribute to as much as 12% of the gender gap in STEM persistence, although when we estimate its effect conditional on the other predictors, its contribution disappears. In supplementary models (not shown), we also estimated the contribution of gender differences in math self-assessment conditioning on academic preparation and math and science coursetaking. This more conservative test of the math self-assessment hypothesis yields no net effect of this factor on gender differences in STEM/Biomed persistence. Math self-assessment does, however, contributes between 3% and 17% (under the generous test, i.e., not conditioning on test scores) of the gender gap in "other" major completion, and -1% to 11% of the gender gap in attrition from college.

Finally, Table 10 shows that, as in the full sample, occupational plans are the strongest predictor of gender gaps in STEM persistence and, at least under the maximum estimate, of the gender gap in attrition ("favouring" women) from the first institution. Specifically, our minimum estimate of the contribution of gender differences in occupational plans is 12% of the gap, and the maximum estimate is 20% of the gap. We note that in models (not shown) that allow the effect of occupational plans on the probability of persisting in STEM/Biomed to differ by gender, the maximum estimate declines to 15% of the gender gap in STEM. Even so, the minimum estimate of the effect of gender differences in occupational plans on the difference between men and women's probability of persisting in STEM, conditional on declaring STEM/Biomed as sophomores, meets or exceeds the maximum estimate of the other predictors.

Discussion and Conclusion

Despite a now two-decade old reversal of the gender gap in college enrolment, the weakening of stereotypes about women's appropriate roles at home and in the workplace, and various educational and policy efforts to encourage girls to study science, gender differences in STEM major completion remain substantial, even for the most recent cohort of college entrants. Men are more than twice as likely to receive STEM and pre-doctoral biomedical baccalaureate degrees than women, and women are nearly four times as likely to receive degrees in health majors leading to MA-level or lower occupations.

Many of the most prominent explanations of the gender gap in STEM major completion receive little empirical support in the ELS data. Gender differences in academic preparation or performance, math/science coursetaking in high school, and family-work orientation account for between 0% and 15% of the gender gap in STEM major completion. Math self-assessment fares slightly better, but only under the "maximum" estimate in which all gender variation in math achievement is assumed to be endogenous to math self-assessment. Under the usual specification of the theory, where gender differences in math self-assessment are assumed to affect supply-side decisions conditional on test scores, its contribution shrinks to 6%.

We have also shown that the gender gap in STEM completion is exacerbated by gender-differentiated attrition from STEM majors. Less than half of young women who declare a STEM/biomed major by their sophomore years will complete those degrees, and indeed these young women are more likely to leave their first institutions without obtaining a baccalaureate degree in any field than they are to graduate with a STEM major. Although some of the impact of women's greater attrition from STEM is offset by their higher propensity to move into STEM majors (especially from health majors) after their sophomore years, this late-career entry into STEM is dwarfed, in absolute terms, by gender-differentiated outflow from STEM/Biomed majors.

Finally, we show that the gender gap in the probability of completing a STEM degree conditional on declaring one as a college sophomore has much the same relationship to the theoretically informed predictors as the gender gap for the full sample: family-work orientation, prior academic preparation, and math and science coursetaking explain very little of the gender gap in STEM persistence. Math self-assessment fares slightly better, at least in the "maximum" estimate, but the strongest predictors are, as in the analysis of STEM completion, the strongest predictors of STEM persistence.

Although our results fail to support the standard explanations for women's greater attrition out of the "STEM pipeline," they leave much room for the operation of other sources of leaks. Our most highly parameterized model leaves a gender gap of five percentage points in STEM/Biomed completion, and our full model of STEM/Biomed persistence (among sophomore STEM majors) leaves a gender gap of 12 percentage points. These models leave less of the gender gap in health major completion (favouring women) unaccounted for, but both men and women's probabilities of completing these majors relative to other outcomes are low enough that the gaps were modest to begin with.

Some of these residual gaps may have been smaller if the ELS data contained measures of academic achievement and coursetaking, family-work orientation, and math self-assessment collected early in the college career, rather than senior year in high school. Still, we think it unlikely that measures that are more temporally proximate to the outcome would alter the results appreciably (see, e.g., Morgan et al 2013). It's equally plausible that factors that we could not measure here – for example, gender-differentiated college experiences such as social pressure exerted through peers and roommates, disparate support or discouragement from teachers and other mentors (e.g., Seymour and Hewitt 1997), or ongoing sex-typed socialization – contribute to women's greater attrition from STEM/Biomed during the college career and their lower odds of completing STEM degrees.

Our analysis offers positive evidence of the importance of gender differences in occupational plans in patterning STEM major completion. These gender differences in plans account for between 21% and 33% of the gender gap in the probability of graduating from the first baccalaureate institution with a major in STEM/ Biomed. It bears emphasizing, though, that our "lower bound" estimates of the effects of occupational plans are sizeable. The implication is that the association between occupational plans and gender differences in college major attainment persists cannot be reduced to gender differences in academic performance, math/science coursetaking, family-work orientation, or math self-assessment. Our future research will focus more directly on understanding the patterns and sources of gender differences in the content, stability, and certainty of occupational plans as they develop throughout students' educational careers.

References

Ayalon, Hanna. (2003). "Women and Men go to University: Mathematical Background and Gender Differences in Choice of Field in Higher Education." Sex Roles 48(5): 277-290.

Buchmann, Claudia, and Thomas A. DiPrete. (2006). "The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement." *American Sociological Review* 71(4): 515-541.

Burkam, D. T., Lee, V. E., & Smerdon, B. A. (2003). Mathematics, foreign language, and science coursetaking and the NELS: 88 transcript data. *US Department of Education, Institute of Education Statistics, National Center for Education Statistics. Retrieved October, 10,* 2006.

Cech, Erin, Brian Rubineau, Susan Silbey, and Caroll Seron. "Professional role confidence and gendered persistence in engineering." American Sociological Review 76, no. 5 (2011): 641-666.

Ceci, Stephen J., and Wendy M. Williams. 2012. Understanding current causes of women's underrepresentation in science. PNAS 108(8):3157-3162.

Correll, Shelley J. (2001). Gender and the Career Choice Process: The Role of Biased Self-Assessments1. *American Journal of Sociology*, *106*(6), 1691-1730.

Correll, Shelley J. (2004). Constraints into preferences: Gender, status, and emerging career aspirations. *American Sociological Review*, *69*(1), 93-113.

Hakim, C. (2002). Lifestyle preferences as determinants of women's differentiated labor market careers. *Work and Occupations*, 29,428 – 459.

Hyde, Janet S., Lindberg, Sara M., Linn, Marcia C., Ellis, Amy B. and Caroline C. Williams. (2008). "Gender Similarities Characterize Math Performance." *Science 321(5888)*: 494-495

Jacobs, J. A. (1989). *Revolving doors: Sex Segregation and Women's Careers*. Stanford University Press.

Legewie, Joscha., and Thomas A DiPrete. (2014). Pathways to Science and Engineering Bachelor's Degrees for Men and Women. *Sociological Science*, *1*, 41-48.

Mann, Allison, and Thomas A. DiPrete. (2013). "Trends in gender segregation in the choice of science and engineering majors." *Social science research*, *42*(6), 1519-1541.

Morgan, Stephen L. (2001). *On the Edge of Commitment.* Stanford University Press.

Morgan, Stephen L., Dafna Gelbgiser, and Kim A. Weeden. (2013). Feeding the Pipeline: Gender, Occupational Plans, and College Major Selection. *Social Science Research*, *42*(4), 989-1005.

Perez-Felkner, L., McDonald, S. K., Schneider, B., & Grogan, E. (2012)."Female and male adolescents' subjective orientations to mathematics and the influence of those orientations on postsecondary majors." *Developmental psychology*, *48*(6), 1658.

Riegle-Crumb, C., King, B., Grodsky, E., & 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.

Seymour, Elaine, and Nancy M. Hewitt. (1997) Talking about Leaving: Why Undergraduates Leave the Sciences, 1997. Boulder, CO: Westview Press.

Turner, Sarah E., and William G. Bowen. (1999). "Choice of Major: The Changing (Unchanging) Gender Gap." *Industrial & Labor Relations Review* 52(2): 289-313.

Williams, W. M., & Ceci, S. J. (2012). "When Scientists Choose Motherhood A single factor goes a long way in explaining the dearth of women in math-intensive fields. How can we address it?" *American scientist*, *100*(2), 138-145.

Williams, Wendy M., and Steven J Ceci. (2015). National Hiring Experiments Reveal 2:1 Faculty Preference for Women on STEM Tenure Track. *PNAS* early edition.

Xie, Yu, and Kimberly A. Shauman. (2003). *Women in Science: Career Processes and Outcomes* (Vol. 26, No. 73.4). Cambridge, MA: Harvard University Press.

Table 1: Early College Major and Major of BA Earned from First Institution by 2012

	Men		Women	
	N	%	N	%
Panel A: Major of BA (by 2012)				
STEM & Biomed-MD	366	15.5	190	6.8
Health-MA/lower	43	1.8	123	4.4
Other fields	686	29.1	960	34.2
AA/Certificate only	71	3.0	91	3.2
No degree	1188	50.5	1442	51.4
Total	2,353	100.0	2,806	100.0
Panel B: Early College Major (2006)				
STEM & Biomed-MD	572	24.3	308	11.0
Health- MA/lower	81	3.5	328	11.7
Other fields	1,000	42.5	1,430	51.0
Did not declare major	373	15.9	419	14.9
Missing major	27	13.9	321	11.4
Total	2,353	100.0	2,806	100.0

Source: Educational Longitudinal Survey, 2002-2012.
N=5,159. Data are weighted.

Table 2: Pre-College Academic Preparation and Math/Science Coursetaking Among Traditional College Students

	Men		Women	
Panel A: Test scores	Mean	sd	Mean	sd
Math score, 12 grade	59.5	12.7	55.7	12.2
Math score, 10th grade	52.9	12.4	49.8	11.9
Reading score, 10th grade	35.2	8.4	35.4	7.8
12th grade GPA	3.1	0.7	3.3	0.6
Panel B: Math Pipeline	N	%	N	%
None/low/middle academic	149	6.4	129	4.6
Middle academic II	397	16.9	531	18.9
Advanced I	450	19.1	576	20.5
Advanced II/Pre-calculus	574	24.4	803	28.6
Advanced III/Calculus	644	27.4	621	22.1
Missing transcripts	139	5.9	145	5.2
Panel C: Science Pipeline				
Low level science	294	12.5	316	11.3
Chemistry 1 or physics 1	677	28.8	985	35.1
Chemistry 1 and physics 1	625	26.6	656	23.4
Chemistry 2 or physics 2 (and/or other advanced)	221	9.4	369	13.2
Chemistry and Physics 2 (and/or other advanced)	397	16.9	334	11.9
Missing transcripts	139	5.9	145	5.2

Source: Educational Longitudinal Survey, 2002-2012

Note: N=5,159. Data are weighted. The coursework pipeline measures are based on Burkham and Lee (2003). Traditional college students are students who entered a 4-year degree program within 6 months of graduating from high school.

Table 3: Pre-College Family and Work Orientation Among Traditional College Students

	Men		Women	
	N	%	N	%
Panel A: family-work values items				
Importance of being successful in line of work				
Not important	11	0.5	3	0.1
Somewhat important	159	6.7	159	5.7
Very important	2,184	92.8	2,644	94.2
Importance of having lots of money				
Not important	220	9.3	330	11.8
Somewhat important	1,250	53.1	1,824	65.0
Very important	884	37.6	651	23.2
Importance of marrying right/having happy family				
Not important	57	2.4	69	2.5
Somewhat important	342	14.5	346	12.3
Very important	1,954	83.0	2,391	85.2
Importance of having children				
Not important	254	10.8	349	12.5
Somewhat important	991	42.1	891	31.8
Very important	1,109	47.1	1,565	55.8
Panel B: Scale score		Mean		Mean
Family-work orientation scale		-0.04		0.20
		(1.15)		(1.18)

Source: Educational Longitudinal Survey, 2002-2012.
N=5,159. Data are weighted.

Table 4: Math Self-Assessment in 2004 (12th Grade) Among Traditional College Students

	Men	Women
	%	%
Panel A: Component Self-Assessment Items		
Can understand difficult math texts		
Almost never	9.3	15.2
Sometimes	35.9	42.3
Often	32.3	30.1
Almost always	22.5	12.4
Can understand difficult math class		
Almost never	9.3	14.0
Sometimes	32.6	37.9
Often	31.5	32.5
Almost always	26.6	15.6
Can master math		
Almost never	5.0	6.8
Sometimes	25.0	28.1
Often	34.7	36.8
Almost always	35.3	28.2
Can do excellent job on math tests		
Almost never	5.5	8.7
Sometimes	35.7	40.0
Often	31.8	31.6
Almost always	27.0	19.7
Can do excellent job on math assignments		
Almost never	3.4	4.8
Sometimes	22.5	24.9
Often	39.9	39.0
Almost always	34.3	31.2
Panel B: Factor scores	Mean (sd)	Mean (sd)
Math self-assessment, 12th grade (standarized)	0.41	0.05
	(1.14)	(1.08)

Source: Educational Longitudinal Survey, 2002-2012

N=5,159. Data are weighted. Traditional college students are students who entered a 4-year degree program within 6 months of graduating from high school.

Table 5: Occupational Plans in 2004 and 2006 Among Traditional College Students

	Men		Women	
	N	%	N	%
Panel A: Plans in 2004 (12th grade)				
STEM only/Biomed-MD only	623	26.5	374	13.3
Health-MA/lower only	93	4.0	425	15.2
Non-STEM only	917	39.0	1243	44.3
Mixture	36	1.5	63	2.3
Don't know	621	26.4	666	23.7
Missing	63	2.7	35	1.3
Total				
Panel B: Plans in 2006 (early college)				
STEM only/Biomed-MD only	471	20.0	279	10.0
Health-MA/lower only	93	3.9	452	16.1
Non-STEM only	1171	49.8	1405	50.1
Mixture	36	1.5	25	0.9
Don't know	549	23.3	596	21.3
Missing	34	1.4	48	1.7
Total	2,353	100.0	2,806	100.0
Panel C: Persistance of plans between 2004 & 2006				
Changed	1,098	46.7	1,201	42.8
Stay the same	1,255	53.4	1,604	57.2
Total	2,353	100.0	2,806	100.0

Source: Educational Longitudinal Survey, 2002-2012

Note: N=5,159. Data are weighted. See text for a detailed description of occupational plans.

Table 6: Social and Demographic Indicators Traditional College Students

	Men		Women	
	Mean	sd	Mean	sd
Race:				
White	0.72		0.70	
Underrepresented minority	0.19		0.21	
Asian	0.06		0.05	
Multirace	0.04		0.04	
School region in 10th grade				
Midwest	0.27		0.26	
North East	0.23		0.22	
South	0.34		0.34	
West	0.16		0.18	
Locality type 10th grade				
Suburban	0.51		0.51	
Urban	0.30		0.29	
Rural	0.19		0.19	
Family structure				
Both parents	0.82		0.81	
Mother only	0.15		0.16	
Father only	0.03		0.03	
Other family structure	0.00		0.01	
SES				
Mother education (in years)	14.6	2.3	14.4	2.3
Father education (in years)	14.8	2.6	14.7	2.7
Family income in 10th grade (logged)	11.1	0.9	11.0	0.8
Mother SEI score	49.8	13.0	48.1	13.3
Father SEI score	48.7	12.0	47.5	12.2

Source: Educational Longitudinal Survey, 2002-2012
N=5,159. Data are weighted.

Table 7: Average Marginal Effects from Multinomial Logit Models Predicting Major of BA from First Institution by 2012

	STEM			Health-MA/lower		
	Men	Women	Gap	Men	Women	Gap
Panel A: predicted probabilities that men and women will earn a BA degree in particular majors						
M1: gender only	0.1554	0.0676	0.0878	0.0183	0.0437	-0.0254
M2: gender+SES	0.1536	0.0682	0.0854	0.0182	0.0439	-0.0257
M3a: M2+family-work orientation	0.1540	0.0681	0.0859	0.0187	0.0431	-0.0244
M3b: M2+academic preparation, math/science pipeline	0.1454	0.0728	0.0726	0.0216	0.0391	-0.0175
M3c: M2+ math self-assessment	0.1430	0.0734	0.0696	0.0186	0.0435	-0.0249
M3d: M2+occupational plans	0.1364	0.0788	0.0576	0.0283	0.0339	-0.0056
M4: full model	0.1316	0.0821	0.0495	0.0341	0.0315	0.0026
M5a: M4-family-work orientation	0.1316	0.0821	0.0495	0.0332	0.0318	0.0014
M5b: M4-academic preparation, math/science pipeline	0.1311	0.0826	0.0485	0.0261	0.0353	-0.0092
M5c: M4-math self-assessment	0.1324	0.0815	0.0509	0.0337	0.0317	0.0020
M5d: M4-occupational plans	0.1421	0.0746	0.0675	0.0224	0.0382	-0.0158
Panel B: minimun and maximum estimate of the percentage change in each gender gap in marginal probabilities						
	Min	Max		Min	Max	
Family-work orientation	0%	1%		0%	0%	
Academic preparation and math/science coursetaking	1%	-15%		-43%	-28%	
Math self-assessment	-2%	-19%		2%	2%	
Occupational plans	-21%	-33%		-70%	-77%	

Source: Educational Longitudinal Survey, 2002-2012

N=5,159. Data are weighted.

Table 7: continued

	Other fields			AA/Certificate			No degree		
	Men	Women	Gap	Men	Women	Gap	Men	Women	Gap
Panel A: predicted probabilities that men and women will earn a BA degree in particular majors									
M1: gender only	0.2913	0.3423	-0.0510	0.0301	0.0324	-0.0023	0.5048	0.5141	-0.0093
M2: gender+SES	0.2883	0.3453	-0.0570	0.0310	0.0317	-0.0007	0.5089	0.5108	-0.0019
M3a: M2+family-work orientation	0.2907	0.3431	-0.0524	0.0309	0.0318	-0.0009	0.5057	0.5138	-0.0081
M3b: M2+academic preparation, math/science pipeline	0.3064	0.3312	-0.0248	0.0308	0.0319	-0.0011	0.4958	0.5250	-0.0292
M3c: M2+ math self-assessment	0.2932	0.3423	-0.0491	0.0315	0.0314	0.0001	0.5138	0.5093	0.0045
M3d: M2+occupational plans	0.2869	0.3480	-0.0611	0.0332	0.0301	0.0031	0.5151	0.5092	0.0059
M4: full model	0.3101	0.3274	-0.0173	0.0320	0.0309	0.0011	0.4921	0.5282	-0.0361
M5a: M4-family-work orientation	0.3091	0.3284	-0.0193	0.0319	0.0310	0.0009	0.4942	0.5268	-0.0326
M5b: M4-academic preparation, math/science pipeline	0.2947	0.3412	-0.0465	0.0331	0.0304	0.0027	0.5151	0.5105	0.0046
M5c: M4-math self-assessment	0.3065	0.3305	-0.0240	0.0325	0.0306	0.0019	0.4950	0.5258	-0.0308
M5d: M4-occupational plans	0.3134	0.3257	-0.0123	0.0305	0.0322	-0.0017	0.4915	0.5293	-0.0378
Panel B: minimun and maximum estimate of the percentage change in each gender gap in marginal probabilities									
	Min	Max		Min	Max		Min	Max	
Family-work orientation	0%	0%		No gender gap			No gender gap		
Academic preparation and math/science coursetaking	-52%	-53%		-	-		-	-	
Math self-assessment	-9%	-6%		-	-		-	-	
Occupational plans	13%	17%		-	-		-	-	

Table 8a: Initial Major Selected in 2006 by College Outcome in 2012 Among Traditional College Students (Outflow %)

	Type and field of degree attained from 1st institution					
Major in 2006	STEM/ Biomed-MD	Health-MA or lower	Other fields	AA or certificate	No degree	Total
Panel A: Men						
STEM & Biomed-MD	51.9	0.5	7.4	3.7	36.6	100.0
Health-MA/lower	4.7	26.7	16.2	2.5	49.8	100.0
Other Majors	1.7	1.3	50.3	3.1	43.7	100.0
Did not declare major	9.3	1.7	26.7	1.8	60.5	100.0
Missing	4.1	0.0	8.5	3.1	84.4	100.0
Total	15.5	1.8	29.1	3.0	50.5	100.0
Panel B: Women						
STEM & Biomed-MD	38.4	3.9	7.4	1.8	48.5	100.0
Health-MA/lower	7.7	25.6	6.2	7.3	53.1	100.0
Other Majors	0.9	0.9	50.7	2.8	44.8	100.0
Did not declare major	4.5	2.9	39.2	2.4	50.9	100.0
Missing	4.5	0.6	8.7	3.7	82.6	100.0
Total	6.8	4.4	34.2	3.2	51.4	100.0

Table 8b: Initial Major Selected in 2006 by College Outcome in 2012 Among Traditional College Students (Inflow %)

	Type and field of degree attained from 1st institution					
Major in 2006	STEM/ Biomed-MD	Health-MA or lower	Other fields	AA or certificate	No degree	Total
Panel A: Men						
STEM & Biomed-MD	81.0	5.9	6.2	29.8	17.6	24.3
Health- MA/lower	1.1	50.3	1.9	2.9	3.4	3.5
Other Majors	4.7	29.6	73.3	43.8	36.8	42.5
Did not declare major	9.5	14.3	14.6	9.4	19.0	15.9
Missing	3.7	0.0	4.0	14.1	23.2	13.9
Total	100.0	100.0	100.0	100.0	100.0	100.0
Panel B: Women						
STEM & Biomed-MD	62.4	9.7	2.4	6.2	10.3	11.0
Health- MA/lower	13.4	68.6	2.1	26.3	12.1	11.7
Other Majors	6.8	10.2	75.5	43.4	44.4	51.0
Did not declare major	10.0	10.0	17.1	11.2	14.8	14.9
Missing	7.5	1.5	2.9	13.0	18.4	11.4
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Educational Longitudinal Survey, 2002-2012

N=5,159. Data are weighted.

Table 9: Descriptive statistics for all predictors for students who enroll in STEM major in 2006

	Male	Female
Social and economic background factors		
Race		
Underreresented minority	19%	24%
Asian	7%	8%
Multirace	5%	2%
White	69%	66%
Geographic region		
North east	25%	17%
South	33%	41%
West	15%	17%
Midwest	26%	25%
Urbanicity		
Urban	26%	34%
Rural	21%	16%
Suburban	53%	51%
Family structure		
Both parents	85%	84%
Mother only	14%	13%
Father only	2%	3%
Other family	0%	0%
Mother education (in years)	14.8	14.6
Father education (in years)	15.1	15.1
Income (logged)	11.1	11.0
Mother SEI score	50.5	48.3
Father SEI score	49.9	48.9
Family work values		
Family-work orientation (factor score)	-0.02	0.04
Academic preparation and math/science pipeline		
Math score 12th grade	64.6	60.5
Math score 10th grade	57.0	53.3
Reading score 10th grade	36.2	36.8
GPA in 12th grade	3.4	3.6
Math pipeline		
None/Low/Middle Acade	3%	1%
Middle academic ii	9%	7%
Advanced i	13%	21%
Advanced ii/Pre-calcu	21%	22%
Advanced iii/Calculus	49%	45%
Missing Transcripts	5%	3%

Table 9: Descriptive statistics for all predictors for students who enroll in STEM major in 2006

	Male	Female
Science pipeline		
Low Level Science	7%	7%
Chemistry 1 or physi	17%	20%
Chemistry 1 and physi	29%	23%
Chemistry 2 or physic	10%	20%
Chemistry and physics	32%	28%
Missing transcripts	5%	3%
Math self concept		
Math self-assessment in 12th grade (factor score)	0.76	0.45
Occupational plans		
Content of plans in 2006		
STEM only/Biomed-MD only	57%	54%
Health-MA/lower only	3%	18%
Non-STEM only	16%	10%
Mixture	3%	2%
Don't know	18%	15%
Missing	2%	2%
Consistency of plans between 2004 and 2006		
Changed	45%	40%
Unchanged	55%	60%

Source: ELS 2002-2012

Notes: Data are weighted. N=891

Appendix A: coefficients from multinomial logit models predicting field of BA degree from 1st institution by 2012 among traditional college students

	M1				M2				M3a				M3b	
	(1)	(2)	(4)	(5)	(6)	(7)	(9)	(10)	(11)	(12)	(14)	(15)	(16)	(17)
Outcome (No BA is the reference outcome)	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health
Gender: women	-0.99**	0.71**	-0.09	-0.14+	-1.00**	0.70**	-0.16	-0.17*	-0.99**	0.67**	-0.13	-0.15+	-0.84**	0.54*
	(0.121)	(0.225)	(0.220)	(0.078)	(0.123)	(0.228)	(0.218)	(0.079)	(0.125)	(0.229)	(0.220)	(0.080)	(0.134)	(0.240)
Demography/SES														
Race (white=0)														
Underrepresented minority					0.10	0.31	0.68*	0.37**	0.08	0.36	0.63*	0.32**	0.52*	0.13
					(0.191)	(0.268)	(0.279)	(0.118)	(0.191)	(0.271)	(0.279)	(0.119)	(0.203)	(0.287)
Asian					0.48*	0.41	0.72*	0.01	0.47*	0.46	0.68*	-0.04	-0.01	0.30
					(0.188)	(0.332)	(0.344)	(0.128)	(0.188)	(0.338)	(0.341)	(0.130)	(0.216)	(0.344)
Multirace					-0.17	0.57	0.32	0.13	-0.18	0.60	0.30	0.10	-0.10	0.54
					(0.315)	(0.488)	(0.527)	(0.200)	(0.315)	(0.491)	(0.523)	(0.201)	(0.331)	(0.501)
School region in 10th grade (midwest=0)														
North east					-0.18	-0.44	0.06	-0.22+	-0.18	-0.44	0.06	-0.22+	-0.14	-0.37
					(0.173)	(0.272)	(0.354)	(0.115)	(0.173)	(0.273)	(0.354)	(0.115)	(0.183)	(0.278)
South					-0.11	-0.03	0.18	-0.00	-0.11	-0.04	0.18	-0.00	-0.06	-0.03
					(0.147)	(0.213)	(0.312)	(0.108)	(0.147)	(0.213)	(0.312)	(0.108)	(0.158)	(0.218)
West					-0.08	-0.49	0.35	-0.16	-0.09	-0.49	0.34	-0.17	-0.09	-0.50
					(0.211)	(0.380)	(0.442)	(0.136)	(0.212)	(0.380)	(0.444)	(0.136)	(0.224)	(0.387)
Urbanicity (suburban=0)														
Urban					0.01	0.00	-0.43	0.00	0.00	0.01	-0.43	-0.00	0.05	0.02
					(0.148)	(0.223)	(0.271)	(0.092)	(0.148)	(0.225)	(0.272)	(0.092)	(0.154)	(0.226)
Rural					0.27	0.30	0.28	0.06	0.26	0.30	0.27	0.06	0.31+	0.32
					(0.162)	(0.242)	(0.310)	(0.110)	(0.162)	(0.242)	(0.310)	(0.110)	(0.170)	(0.243)
Family structure (two parents=0)														
Mother only					-0.27	-0.07	-0.53	-0.14	-0.27	-0.05	-0.55	-0.17	-0.17	-0.07
					(0.192)	(0.313)	(0.332)	(0.118)	(0.194)	(0.314)	(0.333)	(0.119)	(0.191)	(0.322)
Father only					-0.13	-0.99	0.33	-0.11	-0.13	-0.98	0.32	-0.12	-0.34	-0.91
					(0.379)	(0.811)	(0.605)	(0.246)	(0.378)	(0.813)	(0.599)	(0.245)	(0.416)	(0.816)
Other					-1.59	0.05	1.25	-0.01	-1.59	0.10	1.23	-0.02	-1.24	-0.03
					(1.138)	(1.131)	(0.861)	(0.666)	(1.141)	(1.092)	(0.890)	(0.681)	(1.215)	(1.116)
Mother education (in years)					0.04	-0.06	-0.11+	-0.02	0.04	-0.06	-0.11+	-0.02	-0.01	-0.04
					(0.034)	(0.058)	(0.056)	(0.023)	(0.034)	(0.058)	(0.056)	(0.023)	(0.036)	(0.057)
Father education (in years					0.03	-0.01	-0.03	-0.05*	0.03	-0.02	-0.02	-0.05*	0.01	-0.00
					(0.031)	(0.043)	(0.051)	(0.020)	(0.031)	(0.043)	(0.051)	(0.020)	(0.032)	(0.043)
Family income in 10th grade (logged)					-0.16+	-0.02	-0.12	-0.15*	-0.16+	-0.02	-0.11	-0.15*	-0.20*	0.03
					(0.083)	(0.146)	(0.110)	(0.059)	(0.083)	(0.145)	(0.111)	(0.059)	(0.080)	(0.146)
Mother SEI score					-0.00	0.01	0.01	-0.00	-0.00	0.01	0.01	-0.00	-0.00	0.01
					(0.006)	(0.008)	(0.010)	(0.004)	(0.006)	(0.008)	(0.010)	(0.004)	(0.006)	(0.008)
Father SEI score					0.01	-0.01	-0.01	-0.00	0.01	-0.01	-0.01	-0.00	0.00	-0.01
					(0.007)	(0.010)	(0.010)	(0.004)	(0.007)	(0.010)	(0.010)	(0.004)	(0.007)	(0.010)

Appendix A: coefficients from multinomial logit models predicting field of BA degree from 1st institution by 2012 among traditional college students

	M1				M2				M3a				M3b	
	(1)	(2)	(4)	(5)	(6)	(7)	(9)	(10)	(11)	(12)	(14)	(15)	(16)	(17)
Outcome (No BA is the reference outcome)	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health
Family-work orientation (12th grade)														
Family-work scale (centered)									-0.04	0.11	-0.09	-0.11**		
									(0.056)	(0.080)	(0.082)	(0.039)		
Math and science achievements and pipeline														
Math score 12th grade													0.05**	-0.03
													(0.012)	(0.019)
Math score 10th grade													0.01	0.01
													(0.011)	(0.018)
Reading score 10th grade													-0.04**	-0.03*
													(0.011)	(0.016)
GPA in 12th grade													0.32*	0.33
													(0.163)	(0.215)
Math pipeline (Middle academic ii=0)														
None/Low/Middle Academic													0.40	-0.28
													(0.609)	(0.934)
Advanced i													0.50	0.04
													(0.314)	(0.371)
Advanced ii/Pre-calculus													0.16	0.21
													(0.300)	(0.340)
Advanced iii/Calculus													0.55+	0.12
													(0.327)	(0.411)
Missing Transcripts													0.39	2.01**
													(0.463)	(0.695)
Science pipeline (low level chemistry=0)														
Chemistry 1 or physics 1													-0.21	1.81**
													(0.324)	(0.574)
Chemistry 1 and physics 1													0.32	1.81**
													(0.323)	(0.610)
Chemistry 2 or physics 2 or adv bio													0.41	2.26**
													(0.337)	(0.660)
Chemistry and physics and level 7													0.86*	1.39*
													(0.337)	(0.671)
Math self-assessment factor score														

Appendix A: coefficients from multinomial logit models predicting field of BA degree from 1st institution by 2012 among traditional college students

	M1				M2				M3a				M3b	
	(1)	(2)	(4)	(5)	(6)	(7)	(9)	(10)	(11)	(12)	(14)	(15)	(16)	(17)
Outcome (No BA is the reference outcome)	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health
Occupational Plans 2006														
Content of plans (STEM only/biomed-MD only =0)														
Health-MA/lower only														
Non-STEM only														
Mixture														
Don't know														
Missing														
Consistency of plans														
Plans in 2006 are same as 2004 (no=0)														
Content of plans consistency interactions*														
Health-MA/lower only*consistency														
Non-STEM only*consistency														
Mixture*consistency														
Don't know*consistency														
Missing * consistency														
Constant	-0.63**	-2.77**	-2.27**	0.55**	-0.09	-1.70	0.70	3.34**	-0.10	-1.65	0.66	3.31**	-2.25*	-2.92+
	(0.080)	(0.194)	(0.174)	(0.060)	(0.876)	(1.448)	(1.085)	(0.629)	(0.877)	(1.430)	(1.088)	(0.629)	(1.035)	(1.568)
Observations	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159
Model chi-square	95.60	95.60	95.60	95.60	312.6	312.6	312.6	312.6	326.9	326.9	326.9	326.9	857.4	857.4
df	4	4	4	4	68	68	68	68	72	72	72	72	120	120
Log Likelihood	-1.389e+06	-1.389e+06	-1.389e+06	-1.389e+06	-1.364e+06	-1.364e+06	-1.364e+06	-1.364e+06	-1.362e+06	-1.362e+06	-1.362e+06	-1.362e+06	-1.282e+06	-1.282e+06
Pseudo R-squared	0.0108	0.0108	0.0108	0.0108	0.0284	0.0284	0.0284	0.0284	0.0301	0.0301	0.0301	0.0301	0.0871	0.0871

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Appendix A: coefficients from multinomi

	M3b			M3c			M3d				M4			
	(19)	(20)	(21)	(22)	(24)	(25)	(26)	(27)	(29)	(30)	(31)	(32)	(34)	(35)
Outcome (No BA is the reference outcome)	AA	None	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health	AA	None
Gender: women	-0.01	-0.00	-0.85**	0.70**	-0.15	-0.16*	-0.93**	-0.09	-0.33	-0.24**	-0.73**	-0.20	-0.10	-0.00
	(0.219)	(0.088)	(0.125)	(0.230)	(0.219)	(0.079)	(0.137)	(0.244)	(0.223)	(0.083)	(0.157)	(0.264)	(0.226)	(0.091)
Demography/SES														
Race (white=0)														
Underrepresented minority	0.13	0.07	0.15	0.31	0.68*	0.37**	-0.08	0.30	0.66*	0.34**	0.35	0.23	0.06	-0.00
	(0.273)	(0.125)	(0.193)	(0.268)	(0.281)	(0.118)	(0.212)	(0.291)	(0.272)	(0.118)	(0.226)	(0.328)	(0.263)	(0.127)
Asian	0.75*	-0.01	0.52**	0.41	0.72*	0.01	-0.04	0.11	0.60+	-0.13	-0.29	0.11	0.69+	-0.11
	(0.366)	(0.139)	(0.186)	(0.332)	(0.344)	(0.128)	(0.222)	(0.353)	(0.355)	(0.136)	(0.247)	(0.386)	(0.378)	(0.154)
Multirace	0.06	-0.04	-0.08	0.57	0.32	0.13	-0.33	0.70	0.38	0.11	-0.33	0.62	0.12	-0.07
	(0.544)	(0.203)	(0.315)	(0.490)	(0.527)	(0.200)	(0.360)	(0.483)	(0.529)	(0.213)	(0.375)	(0.468)	(0.536)	(0.216)
School region in 10th grade (midwest=0)														
North east	-0.01	-0.30*	-0.20	-0.44	0.06	-0.22+	-0.05	-0.29	0.13	-0.17	-0.04	-0.25	0.02	-0.27*
	(0.369)	(0.122)	(0.175)	(0.272)	(0.353)	(0.115)	(0.180)	(0.314)	(0.356)	(0.115)	(0.195)	(0.330)	(0.373)	(0.122)
South	0.12	-0.04	-0.13	-0.03	0.18	-0.00	-0.10	0.03	0.20	0.01	-0.06	-0.03	0.12	-0.04
	(0.315)	(0.109)	(0.149)	(0.213)	(0.312)	(0.108)	(0.168)	(0.231)	(0.317)	(0.109)	(0.182)	(0.233)	(0.319)	(0.109)
West	0.51	-0.03	-0.09	-0.49	0.34	-0.16	-0.02	-0.36	0.37	-0.15	0.01	-0.44	0.54	-0.02
	(0.422)	(0.137)	(0.209)	(0.380)	(0.441)	(0.136)	(0.251)	(0.414)	(0.446)	(0.142)	(0.261)	(0.424)	(0.434)	(0.143)
Urbanicity (suburban=0)														
Urban	-0.56*	-0.09	0.04	0.00	-0.42	0.00	0.09	0.03	-0.41	0.01	0.12	-0.02	-0.55*	-0.09
	(0.265)	(0.093)	(0.146)	(0.223)	(0.271)	(0.092)	(0.171)	(0.240)	(0.271)	(0.095)	(0.179)	(0.259)	(0.266)	(0.096)
Rural	0.33	0.11	0.27+	0.30	0.28	0.06	0.17	0.11	0.24	0.03	0.20	0.15	0.29	0.08
	(0.320)	(0.114)	(0.165)	(0.241)	(0.310)	(0.110)	(0.188)	(0.284)	(0.310)	(0.112)	(0.196)	(0.291)	(0.322)	(0.116)
Family structure (two parents=0)														
Mother only	-0.72*	-0.26*	-0.27	-0.07	-0.53	-0.14	-0.07	0.05	-0.47	-0.10	-0.07	0.05	-0.68*	-0.23+
	(0.348)	(0.121)	(0.198)	(0.313)	(0.331)	(0.118)	(0.216)	(0.339)	(0.332)	(0.122)	(0.226)	(0.347)	(0.347)	(0.123)
Father only	0.35	-0.13	-0.08	-0.99	0.33	-0.11	0.23	-0.88	0.41	-0.03	0.01	-0.94	0.45	-0.03
	(0.587)	(0.251)	(0.377)	(0.811)	(0.604)	(0.245)	(0.352)	(0.880)	(0.607)	(0.246)	(0.357)	(0.927)	(0.592)	(0.250)
Other	0.90	-0.16	-1.55	0.06	1.25	-0.01	-0.76	0.29	1.25	0.10	-0.39	0.22	0.86	-0.09
	(0.956)	(0.660)	(1.102)	(1.133)	(0.862)	(0.663)	(1.081)	(0.946)	(0.843)	(0.645)	(1.140)	(0.883)	(0.953)	(0.647)
Mother education (in years)	-0.09	-0.02	0.03	-0.06	-0.11+	-0.02	0.05	-0.04	-0.10+	-0.02	0.01	-0.04	-0.08	-0.01
	(0.057)	(0.023)	(0.034)	(0.058)	(0.056)	(0.023)	(0.041)	(0.060)	(0.057)	(0.023)	(0.042)	(0.061)	(0.057)	(0.024)
Father education (in years	0.02	-0.03	0.02	-0.01	-0.03	-0.05*	0.02	-0.03	-0.03	-0.05*	0.01	-0.02	0.01	-0.04+
	(0.051)	(0.020)	(0.031)	(0.043)	(0.051)	(0.020)	(0.036)	(0.047)	(0.052)	(0.021)	(0.037)	(0.048)	(0.051)	(0.021)
Family income in 10th grade (logged)	-0.07	-0.15*	-0.17*	-0.02	-0.11	-0.15*	-0.17+	0.00	-0.11	-0.14*	-0.19*	0.04	-0.05	-0.13*
	(0.120)	(0.062)	(0.086)	(0.146)	(0.109)	(0.059)	(0.096)	(0.154)	(0.113)	(0.060)	(0.092)	(0.147)	(0.123)	(0.060)
Mother SEI score	0.01	-0.00	-0.00	0.01	0.01	-0.00	0.00	0.02+	0.01	-0.00	0.00	0.02+	0.01	-0.00
	(0.010)	(0.004)	(0.006)	(0.008)	(0.010)	(0.004)	(0.006)	(0.009)	(0.010)	(0.004)	(0.006)	(0.009)	(0.010)	(0.004)
Father SEI score	-0.00	0.00	0.01	-0.01	-0.01	-0.00	0.01	0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00
	(0.010)	(0.004)	(0.007)	(0.010)	(0.010)	(0.004)	(0.007)	(0.012)	(0.010)	(0.004)	(0.008)	(0.012)	(0.010)	(0.004)

Appendix A: coefficients from multinomi

	M3b		M3c				M3d				M4			
	(19)	(20)	(21)	(22)	(24)	(25)	(26)	(27)	(29)	(30)	(31)	(32)	(34)	(35)
Outcome (No BA is the reference outcome)	AA	None	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health	AA	None
Family-work orientation (12th grade)														
Family-work scale (centered)											-0.00	0.15	-0.02	-0.06
											(0.065)	(0.096)	(0.090)	(0.041)
Math and science achievements and pipe														
Math score 12th grade	-0.02	-0.01									0.04**	-0.03	-0.03	-0.02*
	(0.020)	(0.007)									(0.014)	(0.023)	(0.021)	(0.008)
Math score 10th grade	-0.00	-0.00									0.00	0.01	-0.00	-0.00
	(0.019)	(0.007)									(0.013)	(0.022)	(0.019)	(0.007)
Reading score 10th grade	-0.03+	-0.01									-0.02+	-0.02	-0.02	-0.00
	(0.017)	(0.007)									(0.013)	(0.019)	(0.017)	(0.008)
GPA in 12th grade	-0.63**	-0.61**									0.08	0.21	-0.70**	-0.66**
	(0.240)	(0.095)									(0.175)	(0.222)	(0.244)	(0.096)
Math pipeline (Middle academic ii=0)														
None/Low/Middle Academic	0.52	0.63**									0.22	-0.14	0.57	0.66**
	(0.437)	(0.230)									(0.610)	(0.932)	(0.439)	(0.230)
Advanced i	-0.16	0.17									0.41	-0.12	-0.21	0.14
	(0.328)	(0.139)									(0.320)	(0.404)	(0.332)	(0.144)
Advanced ii/Pre-calculus	-0.21	0.08									-0.06	-0.07	-0.28	0.01
	(0.365)	(0.132)									(0.322)	(0.359)	(0.375)	(0.139)
Advanced iii/Calculus	0.23	0.29+									0.07	-0.31	0.09	0.15
	(0.463)	(0.156)									(0.367)	(0.441)	(0.469)	(0.163)
Missing Transcripts	0.36	0.28									0.11	2.01**	0.38	0.27
	(0.549)	(0.230)									(0.506)	(0.676)	(0.540)	(0.235)
Science pipeline (low level chemistry=0)														
Chemistry 1 or physics 1	0.10	0.17									-0.21	1.99**	0.12	0.19
	(0.361)	(0.155)									(0.315)	(0.525)	(0.359)	(0.154)
Chemistry 1 and physics 1	-0.03	0.14									0.35	2.07**	-0.01	0.18
	(0.435)	(0.162)									(0.322)	(0.564)	(0.437)	(0.160)
Chemistry 2 or physics 2 or adv bio	0.28	0.24									0.27	2.15**	0.20	0.18
	(0.429)	(0.193)									(0.334)	(0.625)	(0.437)	(0.191)
Chemistry and physics and level 7	-0.09	0.53**									0.73*	1.75**	-0.10	0.51**
	(0.580)	(0.186)									(0.348)	(0.631)	(0.577)	(0.183)
Math self-assessment factor score			0.47**	-0.01	0.00	0.03					0.15*	-0.02	0.20*	0.13**
			(0.060)	(0.095)	(0.082)	(0.036)					(0.071)	(0.112)	(0.091)	(0.041)

Appendix A: coefficients from multinomi

	M3b		M3c				M3d				M4			
	(19)	(20)	(21)	(22)	(24)	(25)	(26)	(27)	(29)	(30)	(31)	(32)	(34)	(35)
Outcome (No BA is the reference outcome)	AA	None	STEM	Health	AA	None	STEM	Health	AA	None	STEM	Health	AA	None
Occupational Plans 2006														
Content of plans (STEM only/biomed-MD														
Health-MA/lower only							-0.36	3.01**	0.86	0.61+	-0.18	2.95**	0.65	0.55
							(0.435)	(0.889)	(0.726)	(0.365)	(0.447)	(0.886)	(0.721)	(0.367)
Non-STEM only							-3.15**	-0.68	-0.71	-0.98**	-3.05**	-0.74	-0.81	-1.03**
							(0.323)	(0.875)	(0.598)	(0.264)	(0.320)	(0.873)	(0.606)	(0.273)
Mixture							-0.73	1.12	-0.10	-0.50	-0.82	1.13	-0.17	-0.53
							(0.630)	(1.110)	(1.287)	(0.625)	(0.627)	(1.130)	(1.357)	(0.653)
Don't know							-2.35**	-0.86	-0.52	-0.66*	-2.19**	-0.91	-0.58	-0.68*
							(0.324)	(0.909)	(0.655)	(0.265)	(0.328)	(0.910)	(0.668)	(0.272)
Missing							-2.30**	0.86	0.96	-0.27	-2.41**	0.90	1.06	-0.19
							(0.644)	(1.329)	(0.808)	(0.422)	(0.617)	(1.314)	(0.849)	(0.448)
Consistency of plans														
Plans in 2006 are same as 2004 (no=0)							0.78*	1.29	0.41	0.33	0.54	1.33	0.60	0.44
							(0.325)	(0.916)	(0.705)	(0.324)	(0.337)	(0.913)	(0.715)	(0.332)
Content of plans consistency interaction*														
Health-MA/lower only*consistency							-0.80	-0.32	0.48	-0.16	-0.55	-0.44	0.46	-0.22
							(0.588)	(1.031)	(0.959)	(0.541)	(0.590)	(1.027)	(0.960)	(0.549)
Non-STEM only*consistency							-1.39**	-1.68	-0.60	-0.51	-1.08*	-1.75+	-0.84	-0.61+
							(0.414)	(1.026)	(0.756)	(0.346)	(0.422)	(1.021)	(0.762)	(0.357)
Mixture*consistency							-19.83**	-20.62**	-18.97**	-0.75	-21.63**	-22.99**	-21.68**	-1.41
							(1.180)	(1.568)	(1.670)	(1.569)	(1.289)	(1.695)	(1.758)	(1.589)
Don't know*consistency							-0.81+	-0.19	-0.48	-0.50	-0.76+	-0.19	-0.54	-0.55
							(0.429)	(1.042)	(0.843)	(0.343)	(0.435)	(1.035)	(0.874)	(0.352)
Missing * consistency							-17.67**	-18.56**	-18.80**	0.40	-18.86**	-21.19**	-21.58**	-0.01
							(1.069)	(1.640)	(1.169)	(1.079)	(1.042)	(1.639)	(1.224)	(1.062)
Constant	3.53*	5.46**	0.11	-1.69	0.69	3.34**	1.74+	-2.60	0.97	3.98**	0.19	-3.75*	4.14**	6.18**
	(1.383)	(0.703)	(0.893)	(1.443)	(1.076)	(0.625)	(1.058)	(1.780)	(1.234)	(0.693)	(1.207)	(1.821)	(1.534)	(0.768)
Observations	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159	5,159
Model chi-square	857.4	857.4	382.7	382.7	382.7	382.7	.	.	.	.	.	.	.	.
df	120	120	72	72	72	72	110	110	110	110	170	170	170	170
Log Likelihood	-1.282e+06	-1.282e+06	-1.350e+06	-1.350e+06	-1.350e+06	-1.350e+06	-1.220e+06	-1.220e+06	-1.220e+06	-1.220e+06	-1.156e+06	-1.156e+06	-1.156e+06	-1.156e+06
Pseudo R-squared	0.0871	0.0871	0.0381	0.0381	0.0381	0.0381	0.131	0.131	0.131	0.131	0.177	0.177	0.177	0.177

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1

Appendix B: Distribution of Major in 2006 and College Outcome by 2012

	Type and field of degree attained from 1st institution					
Major in 2006	STEM/ Biomed-MD	Health- MA or lower	Other fields	No degree	AA or certificate	Total
Panel A: Men						
STEM & Biomed-MD	296	3	42	21	209	572
Health- MA/lower	4	22	13	2	40	81
Other Majors	17	13	503	31	437	1000
Did not declare major	35	6	100	7	226	373
Missing	13	0	28	10	276	327
Total	366	43	685	71	1188	2353
Panel B: Women						
STEM & Biomed-MD	118	12	23	6	149	308
Health- MA/lower	25	84	20	24	175	329
Other Majors	13	12	725	39	641	1430
Did not declare major	19	12	164	10	213	419
Missing	14	2	28	12	265	321
Total	190	123	960	91	1443	2806

Source: Educational Longitudinal Survey, 2002-2012

N=5,159. Data are weighted.

Appendix C: Coefficients from multinomial logit models predicting field of BA degree from 1st instiution by 2012 among students in STEM major in 2006

	M1: gender		M2: Gender+SES		M3a: M2+family work values		M3b: M2+math and science pipeline		M3c: M2+ Math self concept		M3d: M2+ occupational plans		M4: Full model	
Outcome (No BA= reference)	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other
Gender: women	-0.52**	0.14	-0.53**	0.17	-0.55**	0.16	-0.64**	0.08	-0.48**	0.14	-0.43*	0.22	-0.58**	0.10
	(0.180)	(0.300)	(0.182)	(0.301)	(0.184)	(0.300)	(0.213)	(0.323)	(0.186)	(0.297)	(0.182)	(0.313)	(0.217)	(0.340)
Demography/SES														
Race (white=0)														
underrepresented minority			-0.74**	-0.28	-0.68*	-0.24	0.15	-0.10	-0.68*	-0.37	-0.77**	-0.21	0.18	0.07
			(0.280)	(0.378)	(0.284)	(0.388)	(0.309)	(0.458)	(0.281)	(0.402)	(0.294)	(0.384)	(0.315)	(0.445)
Asian			-0.05	0.00	0.01	0.04	-0.12	0.13	-0.03	-0.03	-0.16	0.07	-0.11	0.19
			(0.273)	(0.414)	(0.278)	(0.418)	(0.315)	(0.441)	(0.279)	(0.417)	(0.279)	(0.432)	(0.321)	(0.474)
Multirace			-0.20	0.45	-0.17	0.47	0.02	0.46	-0.17	0.41	-0.26	0.59	-0.01	0.55
			(0.455)	(0.689)	(0.459)	(0.692)	(0.441)	(0.759)	(0.462)	(0.733)	(0.464)	(0.713)	(0.444)	(0.843)
School region in 10th grade (midwest=0)														
North east			-0.26	-0.57	-0.23	-0.54	-0.17	-0.73	-0.25	-0.59	-0.21	-0.63	-0.11	-0.73
			(0.270)	(0.495)	(0.275)	(0.497)	(0.281)	(0.547)	(0.272)	(0.498)	(0.266)	(0.486)	(0.282)	(0.536)
South			-0.25	0.02	-0.21	0.05	-0.22	-0.07	-0.24	-0.02	-0.25	-0.08	-0.23	-0.13
			(0.246)	(0.405)	(0.248)	(0.407)	(0.255)	(0.410)	(0.247)	(0.408)	(0.259)	(0.411)	(0.264)	(0.406)
West			0.01	-0.01	0.11	0.08	-0.17	-0.09	-0.01	-0.01	0.07	-0.08	-0.10	-0.15
			(0.338)	(0.548)	(0.332)	(0.552)	(0.340)	(0.535)	(0.342)	(0.549)	(0.340)	(0.550)	(0.335)	(0.547)
Urbanicity (suburban=0)														
Urban			0.38+	-0.17	0.38+	-0.17	0.48*	-0.19	0.42+	-0.22	0.44+	-0.19	0.57*	-0.21
			(0.223)	(0.335)	(0.224)	(0.335)	(0.239)	(0.339)	(0.223)	(0.341)	(0.227)	(0.340)	(0.237)	(0.347)
Rural			0.24	-0.45	0.27	-0.42	0.13	-0.37	0.24	-0.46	0.28	-0.31	0.21	-0.22
			(0.241)	(0.477)	(0.249)	(0.482)	(0.259)	(0.489)	(0.245)	(0.480)	(0.244)	(0.464)	(0.263)	(0.475)
Family structure (two parents=0)														
Mother only			0.16	0.56	0.17	0.58	0.20	0.73	0.13	0.60	0.14	0.49	0.17	0.71
			(0.278)	(0.461)	(0.272)	(0.459)	(0.276)	(0.452)	(0.283)	(0.462)	(0.297)	(0.464)	(0.290)	(0.456)
Father only			0.75	-0.07	0.80	-0.02	0.45	-0.07	0.74	-0.14	0.63	-0.23	0.38	-0.32
			(0.713)	(1.133)	(0.743)	(1.159)	(0.652)	(1.096)	(0.685)	(1.137)	(0.690)	(1.087)	(0.662)	(1.048)
Other			-15.79**	2.82*	-17.20**	2.62*	-30.86**	3.23*	-17.27**	2.77*	-16.42**	2.79+	-16.77**	3.19+
			(0.939)	(1.219)	(0.891)	(1.157)	(0.918)	(1.430)	(0.963)	(1.183)	(1.181)	(1.472)	(1.122)	(1.696)
Mother education (in years)			0.12*	0.17+	0.12*	0.17+	0.06	0.18+	0.12*	0.18+	0.12*	0.18+	0.06	0.18+
			(0.054)	(0.096)	(0.055)	(0.097)	(0.056)	(0.094)	(0.054)	(0.096)	(0.054)	(0.098)	(0.055)	(0.094)

Appendix C: Coefficients from multinomial logit models predicting field of BA degree from 1st instiution by 2012 among students in STEM major in 2006

	M1: gender		M2: Gender+SES		M3a: M2+family work values		M3b: M2+math and science pipeline		M3c: M2+ Math self concept		M3d: M2+ occupational plans		M4: Full model	
Outcome (No BA= reference)	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other
Father education (in years			0.06	-0.11	0.06	-0.11	0.04	-0.14	0.06	-0.10	0.07	-0.11	0.04	-0.14
			(0.045)	(0.086)	(0.045)	(0.085)	(0.048)	(0.092)	(0.045)	(0.086)	(0.047)	(0.089)	(0.048)	(0.095)
Family income in 10th grade (logged)			-0.04	0.48+	-0.03	0.49+	0.07	0.60*	-0.04	0.46+	-0.04	0.52+	0.07	0.61*
			(0.124)	(0.290)	(0.123)	(0.289)	(0.137)	(0.304)	(0.124)	(0.278)	(0.125)	(0.283)	(0.132)	(0.286)
Mother SEI score			-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
			(0.008)	(0.015)	(0.008)	(0.015)	(0.008)	(0.016)	(0.008)	(0.015)	(0.008)	(0.015)	(0.009)	(0.016)
Father SEI score			-0.01	0.01	-0.01	0.01	-0.02+	0.00	-0.01	0.01	-0.00	0.01	-0.02	0.01
			(0.010)	(0.016)	(0.010)	(0.016)	(0.011)	(0.017)	(0.010)	(0.017)	(0.011)	(0.016)	(0.011)	(0.017)
Family-work orientation (12th grade)														
Family-work scale score					0.15+	0.12							0.13	0.04
					(0.086)	(0.110)							(0.092)	(0.117)
Math and science pipeline														
Math score 12th grade							0.05*	-0.02					0.05*	-0.01
							(0.021)	(0.032)					(0.022)	(0.032)
Math score 10th grade							0.00	0.03					-0.00	0.03
							(0.018)	(0.028)					(0.019)	(0.030)
Reading score 10th grade							-0.01	0.00					-0.00	-0.00
							(0.018)	(0.030)					(0.019)	(0.028)
GPA in 12th grade							0.91**	0.35					0.82**	0.44
							(0.242)	(0.399)					(0.242)	(0.383)
Math pipeline (Middle academic ii=0)														
None/Low/Middle Academic							0.05	-0.08					-0.20	-0.09
							(0.857)	(1.172)					(0.848)	(1.150)
Advanced i							0.47	-0.62					0.39	-0.60
							(0.435)	(0.663)					(0.435)	(0.640)
Advanced ii/Pre-calculus							0.26	-0.21					0.21	-0.25
							(0.419)	(0.684)					(0.421)	(0.684)
Advanced iii/Calculus							0.00	-0.94					-0.10	-0.92
							(0.458)	(0.689)					(0.459)	(0.676)
Missing Transcripts							0.21	1.15					-0.10	1.07
							(0.729)	(1.320)					(0.726)	(1.248)
Science pipeline (low level chemistry=0)														
Chemistry 1 or physics 1							-0.59	0.43					-0.67	0.36
							(0.428)	(0.880)					(0.441)	(0.815)
Chemistry 1 and physics 1							-0.03	0.65					-0.16	0.55
							(0.419)	(0.912)					(0.428)	(0.867)
Chemistry 2 or physics 2 or adv bio							-0.04	1.01					-0.09	0.95
							(0.449)	(0.935)					(0.472)	(0.881)

Appendix C: Coefficients from multinomial logit models predicting field of BA degree from 1st instiution by 2012 among students in STEM major in 2006

	M1: gender		M2: Gender+SES		M3a: M2+family work values		M3b: M2+math and science pipeline		M3c: M2+ Math self concept		M3d: M2+ occupational plans		M4: Full model	
Outcome (No BA= reference)	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other	STEM	Other
Chemistry and physics and level 7							-0.29	0.22					-0.41	0.19
							(0.439)	(0.945)					(0.461)	(0.898)
Math self-assessment factor score									0.23**	-0.21			-0.01	-0.17
									(0.085)	(0.144)			(0.099)	(0.143)
Occupational Plans 2006														
Content of plans (STEM only/biomed-MD only =0)														
Health-MA/lower only											-0.72	0.21	-0.55	-0.03
											(0.488)	(0.744)	(0.503)	(0.717)
Non-STEM only											-0.13	0.84	-0.19	0.70
											(0.398)	(0.662)	(0.426)	(0.672)
Mixture											0.24	-0.59	0.13	-0.69
											(0.592)	(0.873)	(0.563)	(0.882)
Don't know											-0.65+	-0.12	-0.70+	-0.25
											(0.369)	(0.655)	(0.374)	(0.684)
Missing											-1.21+	-0.51	-1.46*	-0.58
											(0.679)	(1.201)	(0.674)	(1.327)
Consistency of plans														
Plans in 2006 are same as 2004 (no=0)											0.05	-0.34	-0.20	-0.41
											(0.268)	(0.467)	(0.287)	(0.466)
Content of plans consistency interactions (a)*														
Health-MA/lower only*consistency											-0.76	-0.04	-0.57	0.11
											(0.817)	(1.031)	(0.830)	(0.994)
Non-STEM only*consistency											-0.70	-0.42	-0.19	-0.05
											(0.532)	(0.846)	(0.595)	(0.864)
Don't know*consistency											0.11	0.99	0.31	1.10
											(0.545)	(0.769)	(0.560)	(0.770)
Constant	0.25*	-1.64**	-1.20	-7.69**	-1.31	-7.76**	-6.71**	-10.21**	-1.16	-7.54**	-0.96	-8.03**	-6.17**	-10.70**
	(0.111)	(0.206)	(1.226)	(2.974)	(1.217)	(2.976)	(1.694)	(3.371)	(1.242)	(2.906)	(1.261)	(3.080)	(1.684)	(3.243)
Observations	891	891	891	891	891	891	891	891	891	891	891	891	891	891
Model chi-square	10.67	10.67	663.1	663.1	740.5	740.5	2341	2341	810.7	810.7	845.3	845.3	978.4	978.4
df	2	2	34	34	36	36	60	60	36	36	52	52	82	82
Log Likelihood	-189697	-189697	-181749	-181749	-181134	-181134	-167048	-167048	-179342	-179342	-176578	-176578	-163312	-163312
Pseudo R-squared	0.0091	0.0091	0.0506	0.0506	0.0538	0.0538	0.1270	0.1270	0.0632	0.0632	0.0776	0.0776	0.1470	0.1470

Robust standard errors in parentheses

** p<0.01, * p<0.05, + p<0.1

(a) the mixture*consistency and missing*consistency interactions are ommitted because there are no relevant observations in the dataset.

Appendix D: Field of BA degree earned from first instiution by occupational plans in 2006

	STEM/ Biomed-MD	Health-MA or lower	Other field	AA/ certificate	No degree	Total
All students						
STEM only/Biomed-MD only	315	12	78	16	329	750
Health-MA/lower only	53	102	46	28	315	545
Non-STEM only	79	28	1,103	74	1,292	2,577
Mixture	15	2	14	2	27	61
Don't know	88	19	387	34	618	1,145
Missing	4	2	19	7	49	81
Total	555	166	1,646	162	2,630	5,159
Men						
STEM only/Biomed-MD only	209	3	39	8	211	471
Health-MA/lower only	15	12	12	0	53	92
Non-STEM only	61	13	475	40	582	1,171
Mixture	14	1	5	2	14	36
Don't know	64	12	151	16	306	549
Missing	3	2	3	4	21	34
Total	366	43	685	71	1,188	2,353
Women						
STEM only/Biomed-MD only	106	9	38	8	118	279
Health-MA/lower only	38	90	34	28	262	452
Non-STEM only	19	15	628	34	710	1,405
Mixture	2	1	8	-	13	25
Don't know	24	7	236	18	311	596
Missing	1	-	16	3	28	48
Total	190	123	960	91	1,443	2,806

Source: Educational Longitudinal Survey, 2002-2012

Notes: N=5,159. Data are weighted.