Does Beauty Matter in Undergraduate Education?

Deryugina, Tatyana and Shurchkov, Olga

5 August 2013

Does Beauty Matter in Undergraduate Education?*

Tatyana Deryugina, University of Illinois+

Olga Shurchkov, Wellesley College++

Abstract:

Physically attractive individuals achieve greater success in terms of earnings and status than those who are less attractive. However, much about the mechanism behind this "beauty premium" remains unknown. We use a rich dataset to shed light on its nature at the college level. We find that students judged to be more attractive perform significantly worse on standardized tests but, conditional on test scores, are not evaluated more favorably at the point of admission. Controlling for test scores, more attractive students receive marginally better grades in some cases. Finally, there is substantial beauty-based sorting into areas of study and occupations.

* We thank an anonymous college for providing us with student photos and academic records. Specifically, we thank the college staff and administration for assistance with acquiring and understanding the data. We also thank Jeff Brown, Kristin Butcher, Daniel Hamermesh, Robin McKnight, Nolan Miller, Julian Reif, Casey Rothschild, and Akila Weerapana for helpful comments. Sizhe Zhang and Mehrnoush Shahhosseini provided excellent research assistance. All remaining errors are our own.

+ Department of Finance, University of Illinois at Urbana-Champaign, 1206 South Sixth St., Champaign, IL, USA. e-mail: deryugin@illinois.edu. Phone: 217-333-9498.

++ Department of Economics, Wellesley College, 106 Central St., Wellesley, MA, USA. e-mail: olga.shurchkov@wellesley.edu.

I. Introduction

In most settings, discrimination based on characteristics such as gender, age, race, and national origin is illegal. Appearance-based discrimination, while not currently unlawful, has been the subject of several lawsuits in recent years.[1] In parallel, the academic literature has documented a positive correlation between earnings and perceived attractiveness for both men and women (Hamermesh and Biddle, 1994; Biddle and Hamermesh, 1998).[2] However, much about the mechanism behind this "beauty premium" remains unknown, including the extent to which beauty is a signal of innate or acquired ability, the extent of bias in favor of more attractive people, and the extent to which sorting plays a role.

In this paper, we use a unique and rich dataset to examine the nature of the beauty premium among students at an all-women's college. Our principal goals are (1) to test whether attractive students appear more academically capable when they begin college, as measured by their standardized test scores and admission ratings; (2) to test whether they appear more capable when they graduate, as measured by their GPA; and (3) to estimate the extent of beauty-based sorting into areas of study and occupations, a potentially important beauty premium mechanism.

---

[1] See for example Yanowitz v. L'Oreal USA, Inc. (2005) and Brice v. Resch and Krueger Int'l, Inc. (Corbett, 2011).

[2] The importance of beauty has been studied in contexts other than the labor market. See, for example, Ravina (2009) for the beauty premium in credit markets, Andreoni and Petrie (2008) for the beauty premium in public goods games, Wilson and Eckel (2006) for the beauty premium in trust games and Berggren et al. (2010) for the beauty premium in electoral outcomes.

To achieve these objectives, we estimate the relationships between attractiveness, standardized test scores, course grades, admissions scores, and major/career choices.

We find that, even though the admissions committee does not observe applicant appearance, more attractive individuals receive lower scores on the college's own formula for rating applicants, in this context an important measure of overall applicant quality that reflects a very wide range of student characteristics. This finding is completely explained by the fact that more attractive individuals have lower standardized test scores. Specifically, a one standard deviation (s.d) increase in attractiveness is associated with scoring 0.10 s.d. lower on the math SAT section, 0.14 s.d. lower on the verbal SAT section, and 0.45 s.d. lower on the SAT writing section. To our knowledge, our study is the first to find that attractiveness is *negatively* correlated with ability, as measured by these tests. Once we control for math and verbal SAT scores, we find no relationship between the college's admission scores and attractiveness. In addition to showing that more attractive individuals do not appear to be more capable at the beginning of college, the absence of such a relationship implies that the negative correlation between attractiveness and SAT scores among admitted students is unlikely to be driven by preferential treatment of attractive students during the admissions process.

Once we control for SAT scores and the admission rating, our results show that more attractive women have a marginally higher overall GPA. We show that this is most likely driven by sorting into types of courses: after controlling for a rich set of course characteristics, our analysis shows no significant relationship between course-level grades and attractiveness, although estimates for some course types are marginally significant.

If more attractive people are aware of the beauty premium in the labor market, they may respond by sorting into areas of study or occupations where their attractiveness generates higher returns. Indeed, we find that there is substantial beauty-based selection into study areas. Specifically, more attractive women are considerably less likely to major in the sciences and much more likely to major in economics. We find no corresponding selection into humanities, other social sciences, or another group of majors that we label "area studies." Overall, we conclude that the beauty premium at the undergraduate level is largely attributable to selection into study areas rather than ability or bias in favor of more attractive students.

Finally, we estimate the extent of beauty-based selection into various occupational categories. Consistent with our results on academic major selection, we find that more attractive women are much more likely to become consultants and managers and much less likely to become scientists and technical workers (including paralegals, technical writers, technicians, and computer programmers). Because previous work has shown that earnings vary substantially by major and occupation, this suggests that at least part of the beauty premium in the labor market is explained by major/occupational choice.[3]

We contribute to several streams of literature. The first assesses the relationship between attractiveness and ability; it has thus far produced mixed findings. Using assortative mating

[3] For more on the relationship between earnings and academic major choice, see Daymont and Andrisani (1984), Berger (1988), James et al. (1989), Grogger and Eide (1995), Loury and Garman (1995), Loury (1997), Blundell (2000), Bratti and Mancini (2003), Arcidiacono (2004), Kelly et al. (2010), Arcidiacono et al.(2011), Andrews et al. (2012), and Wiswal and Zafar (2012).

arguments and observed facts, Kanazawa and Kovar (2004) provide indirect evidence suggesting why beauty and intelligence should be positively correlated in humans. Satoshi (2011) shows empirically that there is a positive association between IQ test results and physical attractiveness in British and American children of both sexes. Several studies have also found that body symmetry is positively correlated with cognitive performance (Prokosch et al. 2005, Bates 2007).[4] However, in a sample of American men, Scholz and Sicinski (2011) find no relationship between attractiveness and IQ or high school class rank. In a laboratory experiment, Mobius and Rosenblat (2006) show that more attractive subjects do not perform better in a maze-solving task. In another experimental setting, Deryugina and Shurchkov (2013) use labor-market-relevant tasks to test for both the existence of a beauty premium and performance differentials between less and more attractive subjects. They find that there is no significant performance differential by attractiveness related to any of the tasks.

To our knowledge, our study is the first to consider the relationship between attractiveness and SAT scores. Although standardized test scores have been shown to be correlated with broad measures of intelligence and cognitive ability (Frey and Detterman, 2004; Beaujean et al., 2006; Rohde and Thompson, 2006; Koenig et al., 2008), we do not claim to show that more attractive students are less intelligent. More conservatively, our results imply that, even if there is a correlation between attractiveness and fundamental intelligence, it does not translate into higher test scores. To the extent that both intelligence and effort are necessary to score well on standardized tests, it is possible that more attractive people perform worse on

[4] Body symmetry has been shown to be strongly correlated with attractiveness (see e.g., Rhodes et al. 1999, Rhodes 2006).

aptitude tests *despite* being more intelligent. Our finding suggests that, if more attractive people are more intelligent, they may invest less time and effort in human capital formation, at least along the dimensions we study (SAT scores and GPA). Alternatively, more attractive students may expect to be more likely to get into college, conditional on their SAT scores, and may therefore rationally exert less effort preparing for standardized tests. Although we cannot test expectations directly, we show that more attractive students do not receive higher admissions scores in our data.

We also contribute to the broad body of literature on the beauty premium. Despite numerous papers on the subject (Hamermesh and Biddle 1994, Biddle and Hamermesh 1998, Fletcher 2009, Ravina 2009, Mocan and Tekin 2010, Berggren et al. 2010, Berri et al. 2011, Scholz and Sicinski 2011, von Bose 2012), much remains to be understood about the beauty premium. In particular, with the exceptions of Scholz and Sicinski (2011) and von Bose (2012), neither its origins nor persistence has been studied. Moreover, with the exception of Fletcher (2009), the extent to which the beauty premium is driven by differences in ability has not been explicitly estimated. In a sample of high-ability law students, Biddle and Hamermesh (1998) find that there are no observable skill differences (including LSAT scores) between more and less attractive individuals. However, von Bose (2012) shows that more attractive teenagers receive higher high school GPAs than less attractive ones. Our findings do not definitively resolve whether more attractive individuals are more capable of academic achievement; while more attractive students perform worse on standardized tests, they are not viewed more or less favorably by admissions officers overall and at best earn marginally higher grades than their less attractive peers. However, we find strong support for another hypothesis about the beauty premium mechanism: sorting.

The policy implications of our findings hinge on whether the observed sorting is efficient. It would be efficient, for example, if more attractive students are selecting into certain majors and occupations because of productivity expectations. If, on the other hand, the sorting is due to attractive students' anticipating a pro-beauty *bias* in some professions then it may not be optimal from a social point of view. This line of inquiry falls outside the scope of this paper. However, it is worthwhile to note that even if the observed sorting is not socially optimal, policies designed to prevent it would most likely be impracticable.

The rest of the paper is organized as follows. Section II describes our sample and data. Section III outlines the empirical strategy. Section IV presents and discusses the findings. Section V concludes.

II. Data

Our dataset consists of 794 alumnae who graduated from an anonymous women's college between the years 2002 and 2011.[5] To measure attractiveness, we use pictures taken when the alumnae were first-year students. A key advantage of our data is that the pictures are not chosen by the student: all are photographed for their student ID cards by campus officials. The pictures were subsequently rated by current male and female students from a college in another state.

[5] This study was conducted with IRB approval. Individuals had to consent to have their photographs included in the study. About 5,000 alumnae were contacted for consent. Informed consent form and description of the project sent to the alumnae are available upon request.

Each picture was rated by at least 25 male and 25 female raters. We then demean the ratings to remove rater fixed effects and average them to obtain the mean attractiveness rating of each alumna. Due to the large number of alumnae, not every picture was rated by the same set of raters. For additional details about the rating procedure, see Appendix A.

The attractiveness rating is then matched to the alumna's academic record, which includes her major, SAT scores, course-level grades, race, non-merit-based financial aid awards, international status, and scores from a quantitative reasoning (QR) test that all first-year students are required to take. Like the SAT, the QR test is scored blindly, without observing the test-taker's appearance. Moreover, we observe each student's admission rating, as assigned by three or more application reviewers. With few exceptions, application reviewers do not observe the student's appearance.[6] At the request of the college, we use a non-disclosed linear transformation to disguise the true rating scale. Finally, we have detailed characteristics for each course, including department, course level (introductory, intermediate, or advanced), total enrollment, and the gender of the instructor.

Starting in the fall semester of 2004, the college implemented an anti-grade-inflation policy that capped the average grade in introductory and intermediate courses with ten or more

---

[6] Some international applicants have TOEFL scores that are accompanied by a picture. In some cases, applicants are interviewed by a member of the admissions staff or by an alumna. However, the application reviewers have access only to the interviewer's comments, which do not contain information about the applicant's appearance.

students to a B+.[7] This policy change disproportionately affected humanities courses. If there is beauty-based selection into humanities courses, this policy change may bias our estimates. To control for the potential impacts of the anti-grade inflation policy, we identify departments that had average grades exceeding a B+ and label beginning and intermediate courses with more than ten students in those departments as "treated." We then control for the treated indicator and its interaction with a "post-fall-semester-2004" indicator in all course-level regressions.

To create our course-level controls and test for sorting, we classify the courses and majors offered at the college into six categories: humanities, sciences, social sciences, area studies, economics, and other. To do this, we use a publication provided by the college, which classifies courses and majors into "Humanities," "Social Sciences," "Science and Mathematics," and "Interdepartmental Programs." Because the "Interdepartmental Programs" category contains a significant share of the majors, we reclassify some of them into the first three categories. In addition, we classify majors such as "South Asia Studies" and "German Studies," which are originally listed as interdepartmental into a new "Area Studies" category. We place economics in its own category because the college does not have a separate business major. Thus, the students who elect to study economics may be different from students choosing other social sciences as their major. The courses and majors that do not fit into any of the above categories are classified as "Other." See Appendix B for the exact classification.

Finally, data on occupations come from alumnae surveys and are available for slightly over half of the alumnae in our sample. We categorize occupations into ten broad categories:

---

[7] The full impact of the anti-grade-inflation policy has been analyzed by Butcher, McEwan, and Weerapana (2013).

consultant/manager, administrator, art/advertising, teacher, technical, scientist, lawyer, doctor, other medical, and non-profit/government. In a few cases, the categories overlap: someone who is working in an administrative position in a non-profit would be placed in both categories, for example. There are a few alumnae reporting occupations that cannot be classified into one or more of these categories, because the occupation descriptions are either vague or very unique. Although we cannot list the specific occupations due to confidentiality concerns, we provide a general list of occupations in each category in Appendix B.

Our data are not without limitations. First, we do not observe parental income, a potentially important control. However, we do observe the amount of need-based and non-need-based loans and grants that a student receives, which we use as a proxy for parental income. We also do not observe post-college earnings. Thus, we cannot test whether more attractive students in our sample also end up earning a higher salary.

Table 1 presents summary statistics for the key variables in the data, broken down by whether individuals are below or above the median attractiveness rating of -0.03. The attractiveness rating itself ranges from -2.7 to 2.4 and has a mean of 0 by construction. The admissions ratings range from 0 to 10, with higher ratings corresponding to a higher chance of admission. The average GPA in the sample is fairly high, ranging from 3.23 in economics to 3.52 in area studies. On average, 94% of the students pass the quantitative reasoning test, which is scored out of 18 points. The average grant amount is about $50,000. Need-based and other loans are substantially smaller, averaging around $1,600 and $700, respectively.

There are some significant differences between those who are above and those who are below the median attractiveness rating. More attractive students are more likely to be

Hispanic/Latina and have about $400 more in non-need-based loans. They score significantly lower on the math, verbal, and writing sections of the SAT as well as on the QR test. In addition, more attractive students have lower admissions ratings, on average. Finally, there are no significant differences between the two groups in terms of GPA, need-based loans and grants, or other racial categories. We later perform a formal regression analysis to test whether the differences in test scores and admissions ratings hold once controls for student characteristics are included.

One potential concern is that the sample of women who consented to participate in our study may not represent all students at the college. To test for this, we compare the mean test scores, admissions ratings, and year of enrollment for the entire population of alumnae who graduated between 2002 and 2011 with those of the consenting group. The results are shown in Table 2. Overall, the consenting students have significantly higher test scores and admission ratings. They also enrolled in the college about half a year later than the general population of students, on average. Because of the necessity to obtain informed consent, we cannot do anything to correct for this or test whether there is beauty-based selection. However, as long as there is no selection on the *relationship* between attractiveness and other outcomes, such as test scores and GPA, our analysis is valid despite the baseline differences. While we view such selection as highly unlikely, we recognize that the validity of our analysis relies on the assumption that it did not occur. We also note that the college from which we obtain data is fairly selective, as evidenced by the high average SAT scores of admitted students (see Table 2). Thus, our study complements some earlier work such as Fletcher (2009), who focuses on individuals with high school diplomas only.

A final concern is that, because our sample comes from a women's college, it may not be representative of colleges as a whole. Again, this would bias our results only if the beauty premium varies by college. The college from which we obtain our data draws from a pool of students and faculty similar to those of other top-tier universities and liberal arts colleges. We do not see any obvious model of sorting that would cast doubt on the generalizability of our results. Moreover, our focus on women complements some earlier work that looks exclusively at men (e.g., Biddle and Hamermesh 1998; Scholz and Sicinski 2011). Finally, the prior literature has found that the beauty premium exists for both men and women and is similar in magnitude. However, replicating the study in a co-educational setting should be an important validation exercise.

III. Empirical framework

Conceptually, we might expect a positive correlation between attractiveness and academic success or, more generally, between attractiveness and some outcome of interest, for several reasons. First, attractiveness may be correlated with a particular characteristic, such as intelligence, trustworthiness, or confidence. If this characteristic is unobserved or poorly measured, omitted variable bias will result, creating a spurious correlation between beauty and the outcome of interest. In the context of academic success, if more attractive people are also more intelligent, then controlling for intelligence is crucial for isolating the effect of attractiveness itself.

Second, attractiveness may itself be productive in some settings. For example, more attractive solicitors may bring in more donations, justifying paying them a higher salary (Landry et al. 2006). This mechanism is less likely to apply in a college setting, however. Finally, people may be biased in favor of more attractive individuals, conditional on their characteristics and productivity.

If more attractive people are aware of these channels, they may respond by exerting greater effort into their academic work and preparing for relevant tests or sorting into occupations where their attractiveness or other characteristics is more productive. Unfortunately, effort in such a context is rarely readily observable. However, the richness of our data allows us to test for beauty-based sorting into both areas of study and occupations. In addition, we estimate the relationships between beauty, aptitude test scores, and course grades to shed light on the other channels through which attractiveness may operate.

In theory, it is possible that a portion of attractiveness can be explained by investment rather than inherent beauty. To our knowledge, there is little work addressing the potential endogeneity of beauty and virtually no work in this area is able to fully eliminate such endogeneity concerns.[8] We control for race and financial aid in all our regressions, which should eliminate some of the components of beauty that may be correlated with socioeconomic characteristics and thus a student's ability to invest in appearing more attractive. Our results are very robust to excluding these controls, and our summary statistics suggest that beauty is not strongly correlated with most of these characteristics.

We first estimate the relationship between the attractiveness rating and admission scores. Because the admissions committee does not observe applicant appearance directly, any

---

[8] One paper that explicitly considers investment in attractiveness is Hamermesh et al. (2002), who find that there is a positive relationship between attractiveness and spending on clothing and cosmetics.

correlation between attractiveness and the admission rating will be due to more attractive students differing in the quality of their recommendation letters, extracurricular activities, personal essays, and other application characteristics.

$$Admissions_i = \alpha Rating_i + X_i'\gamma + \varepsilon_i \quad (1)$$

where i represents the individual alumna and $Admissions_i$ is the average admission score assigned to her by three or more raters. The variable $Rating_i$ is the alumna's attractiveness rating, and X_i is a vector of student characteristics, including math and verbal SAT scores, a set of race indicators, the logs of grant and loan amounts, and year-of-enrollment fixed effects. We add 1 to the grant and loan amounts prior to taking their logs to avoid missing observations. Our results are generally robust to the exclusion of controls for financial aid and race, however. In a related specification, we allow the coefficient on the attractiveness rating to vary by attractiveness quintile to test for non-linear effects.

We then estimate the relationship between attractiveness and GPA, controlling for standardized test scores, the admissions rating, and student characteristics.

$$GPA_i = \beta Rating_i + X_i'\gamma + \varepsilon_i \quad (2)$$

where GPA_i is the student's grade point average on a 0–4 scale. In this case, β may be capturing the effect of bias, sorting, or skill differences that are correlated with attractiveness but are not adequately controlled for by our ability measures. In theory, β may also be capturing direct productivity differences associated with attractiveness itself (e.g., a more attractive model or actor may earn more money because her *attractiveness* is more productive). However, we think direct productivity differences are highly unlikely to be present in a college setting.

In order to remove the influence of some of these factors, we also estimate the relationship between attractiveness and course-level grades. Specifically, we include a rich set of course-level controls to eliminate any beauty premium driven by differential course choices.

$$Grade_{ijt} = \beta Rating_i + Ability_i'\delta + X_i'\rho + Z_j'\gamma + \theta_d + \mu_{at} + \pi T_d + \sigma T_d P_t + \varepsilon_{ijt} \qquad (3)$$

where i indexes individuals, j indexes courses, and t indexes semesters.[9] The variable $Grade_{ijt}$ is the course grade, measured on a 0–4 scale. The vector Z_j is a set of course-level characteristics, namely the gender of the instructor, total enrollment (in logs), and whether the course is a beginning, intermediate, or advanced course. Finally, θ_d is a set of department fixed effects (e.g., English, Mathematics, Physics), and μ_{at} denotes course-area-by-semester fixed effects (e.g., humanities in Fall 2005, sciences in Spring 2008). The variable T_d indicates whether the department had a grade average exceeding a B+ prior to the implementation of the anti-grade-inflation policy and P_t is equal to one for the fall semester of 2004 and later. Standard errors in this specification are clustered by student.

Finally, we estimate the amount of beauty-based sorting into distinct fields of study, using a probit specification.

$$I[Major = M]_i = \beta Rating_i + X_i'\gamma + \varepsilon_i \qquad (4)$$

where $I[Major = M]_i$ is an indicator equal to 1 if a student i is majoring in area M and 0 otherwise. We estimate this relationship separately for five areas of study: humanities, sciences, social sciences, economics, and area studies. The classification of majors into these five areas is detailed in Appendix B. We estimate an analogous equation for career choices.

[9] Fall semesters in two different years are treated as different semesters.

IV. Results

i. Admission ratings and test scores

Our first line of inquiry is to test whether attractiveness is correlated with the admission rating of the student. This test addresses the important question as to whether more attractive applicants differ from less attractive ones prior to college attendance, at least in our sample. Although the admissions committee does not observe everything about the applicant, the applications contain much more information than is available in our data, including extracurricular activities, recommendation letters, and personal essays. Through these, it is possible that the admissions committee receives signals about other skills that predict college success and that may be correlated with attractiveness. Because the admission rating is assigned without observing the student's appearance, any correlation between the two will be due to beauty-based differences in application characteristics, such as those listed above, rather than bias.

The results of this analysis are shown in Table 3. All specifications include controls for the year of enrollment, the student's race, and logs of financial aid amounts by category (need-based loans, other loans, and grants), with 1's added to avoid missing values. Although the admission ratings of more attractive students are worse on average (Column 1), we find that this is entirely driven by SAT scores. Once we control for math and verbal SAT scores (Column 3), there is no relationship between the admission rating and attractiveness. We do not control for the SAT writing section score because it was not offered until 2005, and including it would significantly reduce our sample size. The non-linear specification in Column 2 shows that the lower admission rating of more attractive applicants is driven mainly by those in the top quintile of attractiveness. However, this difference also disappears once we control for SAT test scores (Column 4). Using the estimates in Column 3, we can reject a very small beauty premium of 0.044 or larger in admissions ratings with 95% confidence, which is equivalent to about 0.68% of the mean admission rating.

Table 3 rules out the possibility that attractiveness is correlated (on net) with characteristics that admissions officers can observe but we cannot. It also provides evidence that more attractive applicants are not more likely to get into college, all else remaining equal. In fact, if we do not condition on SAT scores, more attractive applicants seem less likely to be admitted.

The fact that more attractive applicants have worse application packages could be driven by differential ability, effort, or both. If effort is a factor, then test scores and other components of the college application, such as personal essays and extracurricular activities, are endogenously determined. As we discuss in more detail below, if effort is costly and the outside options of more attractive individuals are superior, it is possible that they will expend less effort on college preparation, even if their ability is equal to or exceeds that of their less attractive peers.

One potential concern with the results in Table 3 is that it is conditional on being admitted to the college. While we do not observe the admission ratings of applicants who were not admitted, it is highly implausible that more attractive applicants are more likely to be admitted but do *not* have higher admission ratings conditional on getting in, for several reasons. First, the college that provided the data does not have a strict numeric cutoff for admissions. The density of admission ratings in our sample, shown in Figure 1, confirms this. The distribution looks smooth, with no apparent discontinuity. In order for more attractive applicants to be more likely to be admitted in our case, attractiveness would have to increase the chances of the applicant's being admitted *conditional* on her rating, also implying that the marginal and infra-marginal applicants would have to significantly differ from each other. Second, because the admissions committee does not observe the applicant's appearance, it is unlikely that appearance has an effect separate from the admission rating.[10] Third, we find no significant relationship between attractiveness and the admission rating when we look at students with above-median or below-median admission ratings. In addition, we explicitly test whether attractiveness moderates the importance of test scores by interacting test scores with the attractiveness rating and find that it does not.[11]

The summary statistics in Table 1 along with the results in Table 3 suggest that more attractive students perform worse on standardized tests. To test this directly, we estimate the relationship between (a) SAT and QR scores and (b) attractiveness. Because these tests are scored blindly, there is no concern that examiners are discriminating against or in favor of more attractive people. In other words, it is clear that attractiveness does not directly cause a higher test score in this case. Rather, any correlation between test scores and attractiveness will reflect some unobserved (to us) variable, such as intelligence or effort.

The results are shown in Table 4. Columns 1, 3, 5, and 7 show the results of considering a simple linear relationship between attractiveness and test scores. A one s.d. increase in

---

[10] Factors that may affect the chance of admission conditional on the admission rating include race, the applicant's place of residence (e.g., Massachusetts versus Nebraska), high school quality, parental income, whether the student is the first to go to college in her family, and legacy status, among others.

[11] Results are available upon request.

attractiveness is associated with a 0.10 s.d. decrease in the individual's math SAT score, a 0.14 s.d. decrease in the verbal score, and a 0.45 s.d. decrease in the writing score. Finally, more attractive students also score about 0.20 s.d. lower on the first-year QR score. These results are very robust to varying the set of included controls.

In Columns 2, 4, 6, and 8, we show the results of allowing the relationship between attractiveness and test scores to vary by attractiveness quintile. For the math SAT section (Column 2), there is a sharp and significant drop in scores only for the top attractiveness quintile: the most attractive students score about 0.29 s.d. lower than the least attractive students. The same pattern holds for the QR test (Column 8). For the verbal section of the SAT, the drop is more gradual, with students in the 3rd, 4th, and 5th quintiles performing significantly and progressively worse than students in the bottom quintile. The most attractive students score about 0.40 s.d. lower than the least attractive students. Finally, we cannot detect any differences by quintile on the writing section of the SAT.

A potential objection to these results is that our sample consists of people who have been admitted to the college. If more attractive students have characteristics other than SAT scores that are more likely to result in their being admitted, this would lead to a mechanical negative correlation between SAT scores and attractiveness among the sample of admitted women. However, as we demonstrate in Table 3, the admission scores of more attractive students are *not* higher, providing strong evidence against their being more likely to be admitted. Moreover, restricting the sample to students with below-median admission ratings does not substantively change our results. Thus, our finding of a negative correlation between attractiveness and SAT scores is unlikely to be driven by this type of selection.

To our knowledge, the finding that more attractive people perform worse on standardized tests is new and adds nuance to the hypothesis that more attractive individuals earn more because of some unobservable skill, such as intelligence. In contrast to our results, previous work has found that more attractive individuals attain either equal or higher test scores relative to their less attractive peers. However, with the exception of Biddle and Hamermesh (1998), who estimate the relationship between men's looks and LSAT scores, previous papers have used tests that aim to measure fundamental intelligence, such as IQ tests. SAT scores, on the other hand, are likely determined by some inherent skill *and* by effort, which may explain the difference between our findings and previous work. If more attractive individuals generally expect to receive preferential treatment in life or have better outside options for other reasons, they may rationally choose to exert less effort in advancing their credentials, all else remaining equal. Although some of the previous work on beauty has considered beauty-driven *selection*, none has considered differential investment in human capital. Separating differential investment from ability is beyond the scope of our data. However, our results highlight an important possibility: even if intelligence and attractiveness are positively correlated, the ability of individuals to modify their effort level may counteract or even reverse the relationship between outcomes and beauty in settings where effort matters.

ii. Grades

The estimates in Tables 3 and 4 demonstrate that more attractive students do not begin college with better credentials. If anything, more attractive students have lower admission ratings, driven by their lower SAT scores. In addition, they subsequently score worse on a first-year QR test.

A natural follow-up question is whether more attractive students end up performing better in college than their less attractive counterparts. In other words, is there evidence that the beauty advantage develops during college? We should note right away that such better performance could occur for a number of reasons: skill, bias on behalf of instructors or classmates, endogenous effort, and selection. We try to directly assess some of these factors below.

Throughout the analysis, we use the math and verbal SAT scores as well as the admission rating as ability controls. Again, we do not control for the writing SAT score because it was not offered until 2005, and including it would have significantly reduced our sample size. The benefit of including the admission rating is that it captures a broader range of skills than SAT scores and it appears unaffected by attractiveness. However, our results are unchanged if we omit the admission rating from the set of controls.[12]

Table 5 shows the relationship between attractiveness and GPA, with and without controlling for test scores and the admission rating. We consider both first-year and overall GPA. While the latter is a better reflection of overall student performance, the former might be more relevant for our attractiveness measure, which reflects student appearance in their first year.[13]

[12] Our results are also generally robust to including state of high school/international student fixed effects and to not controlling for the anti-grade-inflation policy. A full set of estimates is available upon request.

[13] Although we allow for the possibility that our attractiveness measure is specific to the first year, von Bose (2012) finds that attractiveness is highly correlated within an individual over time.

Overall, there is no significant relationship between a student's attractiveness rating and her first-year GPA, although the quintile specifications indicate that students in the second quintile have marginally lower GPAs than students in the first quintile. Even though the estimated beauty premium in Column 6 is significantly different from zero, the significance level is marginal and the magnitude is not large. The 95% upper bound for this estimate is 0.21 points or about 6% of the mean GPA per one standard deviation of attractiveness. The admission rating is highly predictive of GPA, demonstrating that it is a useful measure of ex-ante student ability.

We next examine whether there is heterogeneity by area of study. Specifically, we consider GPA separately for five major study areas: sciences, social sciences, humanities, area studies, and economics.[14] There may be less room for instructor discretion in the sciences. Thus, any difference in GPA in this area is more likely to reflect performance differences or selection rather than instructor bias. In addition, more attractive students may select into study areas in which they have a comparative advantage.

Table 6 shows the results. More attractive students have a marginally higher GPA in the sciences, but there is no significant difference between more and less attractive students in any of the other study areas. However, the standard errors on the point estimates are fairly large, and we cannot rule out the possibility that the point estimates in all five areas of study are equal to each other.

We next consider the relationship between course-level grades and attractiveness. Because of the detailed nature of the course-level data, we are able to include many controls to eliminate the selection channel, such as course type (humanities, sciences, social sciences, area studies,

[14] For a discussion of how we classify majors and courses into study areas, see Appendix B.

economics, and other) and department (e.g., math, French, English). We also include course-type-by-semester fixed effects, year-of-enrollment fixed effects, race fixed effects, and financial aid amounts as controls. Standard errors are clustered at the student level.[15]

The results are shown in Table 7 and indicate that there is no significant correlation between attractiveness and course-level grades. The point estimates are positive and similar in magnitude to those shown in Table 5. The inclusion of SAT scores and admissions ratings increases the magnitude of our point estimates, but they remain statistically insignificant. The results are robust to excluding the course-level controls listed above, to not controlling for the anti-grade-inflation policy, and to using only QR test scores or admission ratings as ability controls.[16]

We next test for heterogeneity in the beauty premium between small and large courses and between male and female instructors. We use two measures of course size—an indicator for below- and above-median enrollment (18 or fewer v. 19 or more students) and indicators for enrollment size quartiles. We might expect males to be more responsive to female attractiveness than females (e.g., Landry et al. 2006). We might also expect the beauty premium to be larger in smaller courses because the appearance of individual students is easier to observe.[17]

---

[15] Including course fixed effects in the course-level regressions does not alter our results. However, due to the small number of observations per course (mean of 8, median of 5), we do not use course fixed effects in our preferred specification.

[16] For space reasons, we do not show these specifications. Results are available upon request.

[17] Note that "large" courses have 33 students enrolled on average, with 121 students being the largest class size in our sample.

The results are shown in Table 8. Overall, it appears that there is a modest and marginally significant beauty premium in courses taught by female instructors and in courses with above-median enrollment. There is also a marginally significant beauty premium in the smallest course size quartile (13 students or fewer), with more attractive students receiving grades that are 0.027 points higher. However, as the p-values from the test of equality show, we cannot reject the hypothesis that the rating coefficients in each specification are equal to each other. Moreover, the point estimate for courses with below-median enrollment is actually *larger* than that for courses with above-median enrollment.

Overall, we find little evidence for a meaningful beauty advantage for college grades: while more attractive women have a marginally higher GPA overall, they do not receive significantly higher grades once we control for a rich set of course characteristics. There is also some evidence of a marginal beauty premium in small courses and in courses taught by female instructors, but we cannot rule out the null hypothesis of no heterogeneity in these course characteristics. Our interpretation of these results is that the role of attractiveness in college grades is small and economically insignificant.

iii. Sorting

More attractive students do not begin college with better credentials than their less attractive peers. They appear to earn marginally better grades during college, all else remaining equal, but the differences largely disappear once we introduce extensive course-level controls. This suggests that there may be some beauty-based selection, another channel through which the beauty premium may occur if more attractive people specialize in areas in which they have a comparative advantage.

Our final goal then is to explicitly assess whether more attractive students make systematically different choices in terms of course and major selection. First, we consider the propensity of more attractive students to take courses in five main subject areas: humanities, social sciences, science, area studies, and economics. The dependent variable is the percent of courses taken by the student in that particular subject area.

The results are shown in Table 9. Conditional on their test scores and admission ratings, more attractive students take 1.59 percentage points more economics courses and 1.92 percentage points fewer science courses. There is no selection into other social sciences, humanities, or area studies courses. The pattern of selection by attractiveness quintile suggests that the most attractive women select out of science courses and into economics courses almost one-for-one: women in the fourth and fifth quintile of attractiveness take 3.3 and 4.5 percentage points fewer science courses, respectively, and 3.0 and 4.4 more economics courses, respectively, than the least attractive women. Thus, there is substantial beauty-based selection into course areas.

A natural follow-up question is whether more attractive students are also less likely to major in sciences and more likely to major in economics. We investigate this sorting hypothesis by regressing an indicator variable for whether the student is in a particular major on her attractiveness rating and various controls, using a probit specification. Some students have multiple majors and may thus appear in multiple categories.

The results are shown in Table 10. The estimated coefficients have been scaled by 100 to make them easier to read. As expected given the results in Table 9, more attractive students are significantly less likely to major in the sciences and significantly more likely to major in economics. The marginal effects at the mean indicate that a one s.d. increase in attractiveness is associated with a 5.4% decrease in the probability of majoring in science and a 3.5% increase in the probability of majoring in economics. There is no significant selection into humanities, other social sciences, or area studies majors. Other majors that make up a significant fraction of the sample, namely psychology, English, and political science, likewise show no beauty-based selection (results not shown). The selection out of the sciences and into economics is again driven by the top two quintiles of attractive women, although the likelihoods appear to change monotonically with the quintile.

Finally, we test for beauty-based sorting into occupations, using a probit model. The estimated coefficients, scaled by 100, are shown in Table 11. Because occupation choice occurs shortly before or after final GPA is known, we include it as a control. However, our results are robust to not controlling for GPA.

We find that more attractive women are much more likely to become consultants or managers and much less likely to enter technical or scientific fields.[18] Specifically, a one s.d. increase in attractiveness is associated with a 6.4% increase in the probability of becoming a consultant or manager and a 2.2% decrease in the probability of becoming a scientist or a technician. This is consistent with our earlier results on major choice. There is no significant beauty-based selection into administrative fields, art and advertising, or teaching. Similarly, we find no selection into the medical or legal professions or into non-profit/government jobs (results

[18] Our results are robust to considering “consultant” and “manager” separately, combining “lawyer” and “doctor,” combining “doctor” and “other medical,” and considering “art” and “advertising” separately.

not shown). Although we lack occupation information for almost half of the alumnae in our sample, the fact that these findings mirror those in Tables 8 and 9 makes us more certain that they are not driven by selective reporting. Moreover, the response rate in our data is similar to (and, if anything, slightly higher than) the response rate of the alumnae pool from which we draw our sample.

It is worth considering whether our findings translate into earnings. Although we cannot examine the relationship between earnings and attractiveness directly, we can refer to prior literature. There is some evidence that SAT scores are positively correlated with post-college earnings (Dale and Krueger 2011), as is college GPA (Loury 1997, Arcidiacono 2004, Hershbein 2013).[19] Because none of these papers considers attractiveness specifically and many restrict their sample to men, we are hesitant to argue that they extend to our setting.

There is however more evidence available on the relationship between earnings and choice of college major. The existence of earnings differentials across majors is well-documented (see, e.g., Daymont and Andrisani 1984, Grogger and Eide 1995, Loury 1997, and Arcidiacono 2004). Overall, researchers find that students majoring in fields such as business/economics, science, and engineering generally earn more than those majoring in humanities, education, and other social sciences. These differences persist even after controlling for selection on observables (Andrews et al., 2012). Because we find that more attractive women are less likely to major in sciences but more likely to major in economics, this previous research has ambiguous implications for our findings.

[19] Hershbein (2013) finds that this relationship holds only for less selective colleges.

In experimental settings, both Arcidiacono et al. (2011) and Wiswal and Zafar (2012) show that students' perceptions of expected earnings and ability are significant predictors of major choice. Likewise, Berger (1988) finds that individuals choose majors that they perceive as being more likely to provide a larger stream of earnings. Thus, it is likely that some of the beauty-based selection in our sample is driven by earnings expectations.

Alternatively, differential preferences could be driving the observed sorting in our data. For example, if more attractive individuals are also more extroverted, they may prefer to enter jobs where they are more likely to work with others, even if this does not result in higher wages. To our knowledge, however, there is no work showing the existence of a correlation between attractiveness and personality traits that would lead a more attractive person to choose one career over another, all else (including wages) remaining equal.

V. Conclusion

The issue of beauty-based discrimination has gained increasing attention in recent years. Prior literature has found that more attractive people earn more on average. However, much remains unknown about the origins and evolution of the beauty premium, including whether there are differences in academic capability between more and less attractive individuals, and the extent of bias and sorting that occurs. We contribute to the literature by considering whether there is a beauty advantage before and during college and by estimating the extent to which beauty-based sorting occurs.

We find that more attractive women do not appear more academically capable at the point of college admissions. On the contrary, they receive lower admission ratings, even though the application readers never directly observe applicant appearance. This is because more attractive women receive lower SAT scores. Although previous researchers have found that standardized test scores are positively correlated with measures of cognitive ability (see, e.g., Frey and Detterman 2004, Beaujean et al. 2006, Rohde and Thompson 2006, and Koenig et al. 2008), these findings do not necessarily contradict earlier findings that more attractive people have higher IQs. SAT scores are likely to be a function of both innate ability and effort. It may be that more attractive people rationally exert less effort on the SAT because the other advantages available to them make it optimal to invest fewer resources into scoring higher on the SATs. Testing this hypothesis is beyond the scope of this paper, but should be a worthwhile avenue for future research.

We find substantial beauty-based sorting into areas of study, with more attractive women being significantly less likely to major in the sciences and much more likely to major in economics. They are also subsequently less likely to work in science-related or technical fields and more likely to become consultants, analysts, or managers. Overall, our findings show that the main difference between more and less attractive people during college appears to lie not in the grades they receive but rather in the major and career choices they make.

Unfortunately, we cannot determine whether the observed sorting into majors is socially optimal, which is important for estimating welfare effects and deriving policy implications. However, even if inefficient sorting is present, policy tools capable of addressing it would be controversial and perhaps impossible to implement.

The results suggest several directions for future research. First, reproducing the analysis with a mixed-gender group of college graduates would enhance our understanding of gender differences in the role of appearance in undergraduate education. Second, studying more postgraduation outcomes, such as labor force status, earnings, and history of promotions would shed light on how the beauty premium for college graduates evolves later, after they have entered the labor market.

Figures and Tables

Figure 1. Density plot of admissions ratings

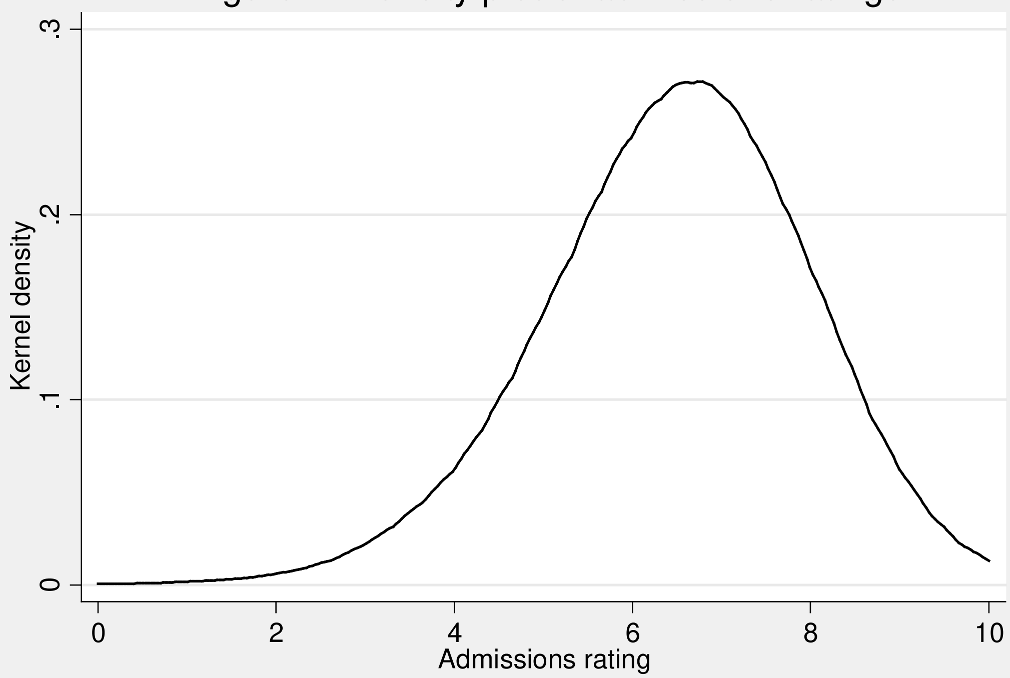

Table 1: Summary Statistics

	Above-median attractiveness rating					Below-median attractiveness rating				
	Mean	Std. dev.	Min	Max	Obs	Mean	Std. dev.	Min	Max	Obs
Attractiveness rating	0.70***	0.54	-0.03	2.42	397	-0.69	0.47	-2.69	-0.03	397
Admission rating	6.34***	1.36	0	10	397	6.62	1.29	1.67	10	395
Total GPA	3.48	0.28	2.5	3.98	396	3.48	0.29	2.30	4	396
Humanities GPA	3.50	0.29	2.11	4	397	3.48	0.32	1.34	4	394
Social Science GPA	3.5	0.29	2.44	4	395	3.5	0.29	2.00	4	393
Science GPA	3.24	0.49	1.56	4	395	3.22	0.52	1.40	4	395
Area studies GPA	3.52	0.38	2.33	4	125	3.48	0.43	2.00	4	140
Economics GPA	3.23	0.53	1.00	4	223	3.24	0.57	1.00	4	204
Math SAT score	678***	62	510	800	387	689	57	490	800	378
Verbal SAT score	696***	61	490	800	387	712	59	450	800	378
Writing SAT score	699*	67	490	800	274	710	66	500	800	277
QR test score	13.08*	2.65	2	18	397	13.42	2.55	4.5	18	397
Passed QR test	0.93	0.25	0	1	397	0.95	0.22	0	1	397
Asian	0.21	0.41	0	1	397	0.22	0.41	0	1	397
Black	0.03	0.16	0	1	397	0.04	0.2	0	1	397
White	0.61	0.49	0	1	397	0.64	0.48	0	1	397
Hispanic	0.05**	0.21	0	1	397	0.02	0.14	0	1	397
Latina	0.07**	0.26	0	1	397	0.03	0.18	0	1	397
Need-based loans ($)	1,759	3,438	0	15,795	397	1,492	3,080	0	17,675	397
Grants ($)	47,524	54,836	0	202198	397	51,202	57,598	0	199368	397
Other loans ($)	891**	3,424	0	24,700	397	500	1,864	0	14,500	397

Stars indicate significant differences in means from the "below-median" group. Significance levels: *10 percent, ** 5 percent, *** 1 percent. Social science GPA excludes economics.

Table 2: difference between the general student population and those giving consent

	(1) Math SAT	(2) Verbal SAT	(3) Writing SAT	(4) QR test score	(5) Admission rating	(6) Year enrolled
All students	674	684	683	12.94	6.22	2002
Consenting minus all	9.13***	20.08***	21.36***	0.31***	0.26***	0.53***
Observations	5,894	5,894	4,544	6,158	6,155	6,160

Significance levels: *10 percent, ** 5 percent, *** 1 percent.

Table 3: Attractiveness and admission ratings

	(1)	(2)	(3)	(4)
Attractiveness rating	-0.131***		-0.038	
	(0.049)		(0.042)	
Attractiveness quintile = 2		0.01		0.046
		(0.146)		(0.128)
Attractiveness quintile = 3		0.005		0.131
		(0.138)		(0.121)
Attractiveness quintile = 4		-0.146		-0.027
		(0.146)		(0.128)
Top attractiveness quintile		-0.409***		-0.128
		(0.150)		(0.132)
Math SAT score			0.432***	0.429***
			(0.050)	(0.050)
Verbal SAT score			0.408***	0.409***
			(0.045)	(0.045)
Dep. var. mean	6.48	6.48	6.51	6.51
Observations	791	791	762	762
R-squared	0.13	0.13	0.35	0.35

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. All regressions include year of enrollment and race fixed effects, as well as controls for the amount of financial aid received.

Table 3: Attractiveness and test scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Math SAT		Verbal SAT		Writing SAT		QR test	
Attractiveness rating	-0.10***		-0.14***		-0.45*		-0.20**	
	(0.03)		(0.03)		(0.26)		(0.09)	
Attractiveness quintile = 2		0.04		-0.14		-1.37		0.00
		(0.10)		(0.11)		(0.86)		(0.27)
Attractiveness quintile = 3		-0.12		-0.27**		-1.08		-0.20
		(0.10)		(0.11)		(0.83)		(0.29)
Attractiveness quintile = 4		-0.08		-0.30***		-1.43		-0.31
		(0.10)		(0.11)		(0.89)		(0.27)
Top attractiveness quintile		-0.29***		-0.40***		-1.22		-0.55**
		(0.10)		(0.11)		(0.86)		(0.27)
Dep. var. mean	11.37	11.37	11.59	11.59	70.44	70.44	13.25	13.25
Observations	764	764	764	764	551	551	793	793
R-squared	0.22	0.22	0.11	0.11	0.12	0.12	0.12	0.12

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. All regressions include year-of-enrollment and race fixed effects, as well as controls for the amount of financial aid received.

Table 5: Attractiveness and GPA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	First year GPA				Overall GPA			
Attractiveness rating	-0.003	0.012			0.004	0.016*		
	(0.012)	(0.012)			(0.010)	(0.010)		
Attractiveness quintile = 2			-0.076*	-0.077*			-0.069**	-0.072**
			(0.040)	(0.040)			(0.032)	(0.032)
Attractiveness quintile = 3			0.009	0.013			-0.001	-0.002
			(0.035)	(0.035)			(0.028)	(0.029)
Attractiveness quintile = 4			-0.016	-0.002			-0.003	0.007
			(0.037)	(0.036)			(0.030)	(0.028)
Top attractiveness quintile			-0.041	0.003			-0.020	0.015
			(0.037)	(0.036)			(0.031)	(0.029)
Math SAT score		0.034**		0.035**		0.017		0.019
		(0.015)		(0.015)		(0.012)		(0.012)
Verbal SAT score		0.000		-0.001		0.005		0.004
		(0.013)		(0.013)		(0.011)		(0.011)
Admission rating		0.072***		0.071***		0.059***		0.060***
		(0.010)		(0.010)		(0.009)		(0.009)
Dep. var. mean	3.40	3.41	3.40	3.41	3.48	3.48	3.48	3.48
Observations	793	762	793	762	791	760	791	760
R-squared	0.12	0.20	0.12	0.20	0.10	0.18	0.11	0.19

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. All regressions include year-of-enrollment fixed effects, race fixed effects, and financial aid amounts.

Table 6: Attractiveness and GPA by area of study

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Sciences		Social Sciences		Humanities		Area Studies		Economics	
Attractiveness rating	0.028*		0.016		0.016		0.031		0.022	
	(0.017)		(0.010)		(0.010)		(0.025)		(0.027)	
Attractiveness quintile = 2		-0.073		-0.024		-0.059		-0.054		-0.124
		(0.056)		(0.031)		(0.036)		(0.082)		(0.087)
Attractiveness quintile = 3		0.036		-0.011		0.028		-0.028		-0.023
		(0.053)		(0.031)		(0.029)		(0.076)		(0.085)
Attractiveness quintile = 4		0.035		0.039		0.013		-0.026		0.046
		(0.051)		(0.030)		(0.030)		(0.076)		(0.077)
Top attractiveness quintile		0.036		0.028		0.014		0.077		-0.014
		(0.052)		(0.031)		(0.030)		(0.077)		(0.077)
Math SAT score	0.105***	0.107***	0.000	0.000	0.011	0.012	0.045	0.051	0.128***	0.131***
	(0.022)	(0.022)	(0.012)	(0.012)	(0.012)	(0.013)	(0.034)	(0.034)	(0.031)	(0.031)
Verbal SAT score	-0.017	-0.018	0.013	0.012	-0.002	-0.003	0.010	0.004	0.007	0.007
	(0.020)	(0.020)	(0.011)	(0.011)	(0.011)	(0.011)	(0.032)	(0.032)	(0.027)	(0.026)
Admission rating	0.072***	0.071***	0.055***	0.056***	0.057***	0.057***	0.048**	0.051**	0.079***	0.076***
	(0.017)	(0.017)	(0.010)	(0.010)	(0.010)	(0.010)	(0.021)	(0.022)	(0.023)	(0.023)
Dep. var. mean	3.24	3.24	3.50	3.50	3.49	3.49	3.52	3.52	3.24	3.24
Observations	759	759	756	756	760	760	251	251	412	412
R-squared	0.16	0.17	0.19	0.19	0.18	0.19	0.19	0.19	0.24	0.24

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. All specifications include year and race fixed effects, as well as financial aid controls and controls for math SATs, verbal SATs, and admission rating.

Table 7: Attractiveness and course-level grades

	(1)	(2)	(3)	(4)	(5)	(6)
Attractiveness rating	0.003		0.012		0.015	
	(0.011)		(0.010)		(0.010)	
Attractiveness quintile = 2		-0.044		-0.045		-0.046
		(0.032)		(0.032)		(0.031)
Attractiveness quintile = 3		-0.006		-0.001		-0.003
		(0.030)		(0.030)		(0.029)
Attractiveness quintile = 4		-0.001		0.003		0.009
		(0.030)		(0.029)		(0.028)
Top attractiveness quintile		-0.016		0.010		0.021
		(0.031)		(0.030)		(0.029)
Math SAT score			0.059***	0.060***	0.035***	0.035***
			(0.012)	(0.012)	(0.012)	(0.012)
Verbal SAT score			0.034***	0.033***	0.007	0.006
			(0.011)	(0.011)	(0.011)	(0.011)
Admission rating					0.061***	0.061***
					(0.009)	(0.009)
Dep. var. mean	3.44	3.44	3.45	3.45	3.45	3.45
Observations	19,525	19,525	18,872	18,872	18,832	18,832
R-squared	0.12	0.13	0.14	0.14	0.15	0.15

Standard errors clustered by student in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. All regressions include fixed effects for: department, course level, semester-by-course type, year of enrollment, and race. In addition, controls include the gender of the instructor, total course enrollment (log), and the amount of financial aid received by the student. Course level is either beginning, intermediate, or advanced. Course type is humanities, social sciences, economics, area studies, sciences, or other. Department fixed effects represent a specific department code, such as English, Economics, Physics, etc.

Table 8: Attractiveness and course-level grades heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)
Female prof x rating	0.007	0.022*				
	(0.013)	(0.012)				
Male prof x rating	-0.002	0.013				
	(0.013)	(0.012)				
Below median size x rating			0.007	0.026		
			(0.016)	(0.016)		
Above median size x rating			0.004	0.020*		
			(0.012)	(0.011)		
Bottom quartile x rating					0.008	0.025*
					(0.013)	(0.013)
2nd quartile x rating					0.000	0.015
					(0.014)	(0.013)
3rd quartile x rating					0.005	0.019
					(0.016)	(0.015)
Top quartile x rating					-0.001	0.012
					(0.015)	(0.014)
Ability controls		Y		Y		Y
Test of equality p-value	0.35	0.29	0.77	0.57		
1st = 2nd quartile p-value					0.48	0.39
1st = 3rd quartile p-value					0.82	0.66
1st = 4th quartile p-value					0.46	0.30
Dep. var. mean	3.44	3.45	3.44	3.45	3.44	3.45
Observations	19,525	18,832	19,433	18,741	19,433	18,741
R-squared	0.12	0.15	0.12	0.15	0.12	0.15

Standard errors clustered by student in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. All regressions include fixed effects for: department, course level, semester-by-course type, year of enrollment, and race. In addition, controls include the gender of the instructor, total course enrollment (log), and the amount of financial aid received by the student. Course level is either beginning, intermediate, or advanced. Course type is humanities, social sciences, economics, area studies, sciences, or other. Department fixed effects represent a specific department code, such as English, Economics, Physics, etc.

Table 9: Selection into subject areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Sciences		Social Science		Humanities		Area studies		Economics	
Attractiveness rating	-1.92***		0.36		-0.05		-0.01		1.59***	
	(0.62)		(0.67)		(0.63)		(0.21)		(0.45)	
Attractiveness quintile = 2		1.95		-4.64**		1.35		0.67		0.84
		(1.95)		(2.10)		(1.96)		(0.65)		(1.42)
Attractiveness quintile = 3		-1.14		-2.72		1.04		0.19		2.11
		(1.94)		(2.08)		(1.95)		(0.64)		(1.41)
Attractiveness quintile = 4		-3.30*		-1.64		2.45		-0.01		2.96**
		(1.94)		(2.09)		(1.96)		(0.65)		(1.42)
Top attractiveness quintile		-4.48**		-0.82		0.51		0.17		4.37***
		(1.96)		(2.11)		(1.97)		(0.65)		(1.43)
Math SAT score	4.30***	4.25***	-4.34***	-4.29***	-3.72***	-3.74***	0.57**	0.56**	3.23***	3.24***
	(0.75)	(0.75)	(0.81)	(0.81)	(0.75)	(0.76)	(0.25)	(0.25)	(0.54)	(0.55)
Verbal SAT score	-1.86***	-1.81**	1.16	1.06	2.26***	2.32***	0.18	0.18	-1.84***	-1.87***
	(0.70)	(0.70)	(0.76)	(0.76)	(0.71)	(0.71)	(0.23)	(0.23)	(0.51)	(0.51)
Admission rating	1.04*	0.99*	-0.69	-0.63	-0.25	-0.26	-0.09	-0.09	0.08	0.10
	(0.57)	(0.57)	(0.61)	(0.61)	(0.57)	(0.57)	(0.19)	(0.19)	(0.41)	(0.42)
Dep. var. mean	0.23	0.23	0.32	0.32	0.31	0.31	0.03	0.03	0.09	0.09
Observations	762	762	762	762	762	762	762	762	762	762
R-squared	0.12	0.12	0.08	0.09	0.15	0.16	0.10	0.10	0.14	0.14

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. Dependent variable is fraction of courses taken in a particular subject area. All regressions include year of enrollment and race fixed effects, as well as controls for the amount of financial aid received.

Table 10: Selection into majors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Sciences		Social Science		Humanities		Area Studies		Economics	
Attractiveness rating	-18.17***		3.84		-5.78		10.91		16.49***	
	(5.51)		(4.88)		(4.97)		(7.00)		(6.05)	
Attractiveness quintile = 2		-9.35		-26.57*		5.03		9.99		8.48
		(15.73)		(15.28)		(15.42)		(22.43)		(19.92)
Attractiveness quintile = 3		-24.64		-5.02		-1.14		-9.25		17.52
		(15.79)		(14.88)		(15.27)		(23.52)		(19.78)
Attractiveness quintile = 4		-53.37***		-1.51		9.40		28.61		35.76*
		(16.68)		(14.96)		(15.32)		(20.35)		(18.97)
Top attractiveness quintile		-45.28***		0.64		-21.36		33.15		45.08**
		(16.69)		(15.19)		(16.09)		(20.42)		(19.13)
Math SAT score	30.84***	31.37***	-28.58***	-28.51***	-17.69***	-18.22***	-5.43	-5.64	47.69***	47.83***
	(6.56)	(6.58)	(5.89)	(5.88)	(5.87)	(5.90)	(8.04)	(8.10)	(7.41)	(7.49)
Verbal SAT score	-13.01**	-13.04**	5.43	5.15	4.39	4.83	9.20	9.01	-20.09***	-20.51***
	(6.07)	(6.09)	(5.34)	(5.33)	(5.70)	(5.68)	(7.86)	(7.82)	(6.55)	(6.55)
Admission rating	4.29	4.11	-3.88	-3.66	-2.03	-2.31	1.60	2.07	-1.69	-1.36
	(4.70)	(4.70)	(4.45)	(4.47)	(4.69)	(4.71)	(6.41)	(6.32)	(5.52)	(5.52)
Dep. var. mean	0.24	0.24	0.40	0.40	0.34	0.34	0.09	0.09	0.17	0.17
Observations	760	760	760	760	760	760	760	760	760	760

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. Regression specification is a probit. Dependent variable is an indicator for majoring in a given subject area. All regressions include year-of-enrollment and race fixed effects, as well as controls for the amount of financial aid received.

Table 11: Attractiveness and career choice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Consultant/manager		Administrator		Art/advertising		Teacher		Technical		Scientist	
Attractiveness rating	20.42***		0.42		-1.37		-13.20		-22.56**		-28.19**	
	(7.24)		(7.63)		(8.42)		(10.42)		(9.40)		(11.53)	
Attractiveness quintile = 2		4.65		-53.31**		-2.26		-38.58		23.68		12.11
		(23.28)		(22.79)		(29.20)		(29.05)		(27.98)		(27.35)
Attractiveness quintile = 3		8.78		-10.30		-0.21		-30.31		31.16		-43.64
		(22.90)		(21.36)		(29.08)		(28.15)		(28.13)		(36.21)
Attractiveness quintile = 4		26.61		-12.09		27.76		-34.78		-85.87**		-48.02
		(22.14)		(21.91)		(28.04)		(27.71)		(42.42)		(33.80)
Top attractiveness quintile		44.08*		-10.37		-22.55		-41.12		-44.57		-62.57*
		(23.12)		(22.49)		(30.07)		(30.65)		(36.25)		(36.45)
Math SAT score	18.38**	17.80**	-18.14**	-17.67**	-17.65*	-19.04*	-12.92	-12.46	9.53	9.08	4.14	3.06
	(8.92)	(8.94)	(8.27)	(8.29)	(9.96)	(10.11)	(11.17)	(11.24)	(12.75)	(12.84)	(12.58)	(12.58)
Verbal SAT score	-5.34	-5.89	-3.16	-3.10	14.43	15.99	5.25	4.49	5.82	8.40	-5.06	-4.44
	(8.61)	(8.54)	(8.56)	(8.45)	(10.05)	(10.30)	(10.67)	(10.65)	(10.23)	(10.76)	(10.65)	(10.81)
Admission rating	-7.12	-6.44	5.30	5.33	-12.63	-12.83	0.82	1.26	-4.12	-8.02	-1.32	-0.27
	(7.37)	(7.26)	(6.68)	(6.64)	(8.10)	(8.13)	(7.47)	(7.54)	(7.89)	(8.03)	(7.29)	(6.81)
GPA	41.00	39.89	-25.96	-34.41	12.90	17.16	26.44	24.41	7.40	16.12	-58.41	-50.27
	(28.36)	(28.54)	(27.18)	(27.63)	(31.27)	(32.64)	(37.78)	(38.80)	(43.80)	(44.14)	(44.79)	(43.70)
Dep. var. mean	0.27	0.27	0.27	0.27	0.14	0.14	0.09	0.09	0.06	0.06	0.06	0.06
Observations	413	413	413	413	349	349	413	413	413	413	383	383

Robust standard errors in parentheses. Significance levels: *10 percent, ** 5 percent, *** 1 percent. Regression specification is a probit. Dependent variable is an indicator for reporting an occupation in the given area. All regressions include year of enrollment and race fixed effects, as well as controls for the amount of financial aid received.

References

Andreoni, James and Ragan Petrie. 2008. "Beauty, Gender and Stereotypes: Evidence From Laboratory Experiments." *Journal of Economic Psychology*, 29(1): 73–93.

Andrews, Rodney J, Jing Li, and Michael F. Lovenheim. 2012. "Quantile Treatment Effects of College Quality on Earnings: Evidence from Administrative Data in Texas." NBER Working Paper 18068.

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

Arcidiacono, Peter, V. Joseph Hotz, and Songman Kang, 2011. "Modeling College Major Choice using Elicited Measures of Expectations and Counterfactuals." *Journal of Econometrics*, 166(1): 3-16.

Bates, Timothy C. 2007. "Fluctuating Asymmetry and Intelligence." *Intelligence*, 35: 41–46.

Beaujean, A. Alexander, Michael W. Firmin, Andrew J. Knoop, Jared D. Michonski, Theodore P. Berry, and Ruth E. Lowrie. 2006. "Validation of the Frey and Detterman (2004) IQ Prediction Equations Using the Reynolds Intellectual Assessment Scales." *Personality and Individual Differences*, 41(2): 353-357.

Berger, Mark C. 1988. "Predicted Future Earnings and Choice of College Major." *Industrial and Labor Relations Review*, 41(3): 418–429.

Berggren N., H. Jordahl, and P. Poutvaara. 2010. "The Looks of a Winner: Beauty, Gender, and Electoral Success." *Journal of Public Economics*, 94(1-2): 8:15.

Berri, David J., Rob Simmons, Jennifer Van Gilder, and Lisle O'Neill. 2011. "What Does it Mean to Find the Face of the Franchise? Physical Attractiveness and the Evaluation of Athletic performance." *Economics Letters*, 111(3): 200–202.

Biddle, J. and D. Hamermesh. 1998. "Beauty, Productivity, and Discrimination: Lawyers' Looks and Lucre." *Journal of Labor Economics,* 16(1): 172-201.

Blundell Richard, Lorraine Dearden, Alissa Goodman, and Howard Reed. 2000. "The Returns to Higher Education in Britain: Evidence from a British Cohort." *Economic Journal*, 110(461): 82–99.

von Bose, Caroline. 2012. "Child Stars vs. Ugly Ducklings: Does Adolescent Attractiveness Contribute to the Beauty Premium?" Working paper.

Butcher, McEwan, and Weerapana. 2013. "A Quasi-Experimental Analysis of the Impact of an Anti-Grade Inflation Policy On Students and Instructors," Working paper.

Corbett, W. R. 2011. "Hotness Discrimination: Appearance Discrimination as a Mirror for Reflecting on the Body of Employment-Discrimination Law." *Catholic University Law Review*, 60 Cath. U.L. Rev. 615

Dale, Stacy and Alan B. Krueger. 2011. "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data." NBER Working Paper No. 17159.

Daymont, Thomas N. and Paul J. Andrisani. 1984. "Job Preferences, College Major, and the Gender Gap in Earnings." *Journal of Human Resources*, 19(3): 408–428.

Deryugina, T. and O. Shurchkov. 2013. "Are Appearances Deceiving? The Nature and Evolution of the Beauty Premium." Working paper.

Fletcher, J. 2009. “Beauty vs. Brains: Early Labor Market Outcomes of High School Graduates.” *Economic Letters*, 105 (3): 321-325.

Frey, Meredith C. and Douglas K. Detterman. 2004. “Scholastic Assessment or G? The Relationship Between the SAT and General Cognitive Ability.” *Psychological Science*, 15(6): 373–378.

Hamermesh, D. and J. Biddle. 1994. “Beauty and the Labor Market.” *American Economics Review,* 84(5): 1174-1194.

Hamermesh, Daniel S., Xin Meng, and Junsen Zhang. 2002. “Dress for success – does primping pay?” *Labour Economics*, 9(3): 361-373.

Hamermesh, D. 2011. “Beauty Pays: Why Attractive People Are More Successful.” Princeton University Press, Princeton, NJ, USA.

Kanazawa, Satoshi and Jody L. Kovar. 2004. “Why Beautiful People are More Intelligent.” *Intelligence*, 32: 227–243.

Kelly, Elish, Philip J. O’Connell, and Emer Smyth. 2010. “The Economic Returns to Field of Study and Competencies among Higher Education Graduates in Ireland.” *Economics of Education Review* 29, 650–657.

Koenig, Katherine A., Meredith C. Frey, and Douglas K. Detterman. 2008. “ACT and General Cognitive Ability.” *Intelligence*, 36(2): 153-160.

Landry, Craig E., Andreas Lange, John A. List, Michael K. Price, and Nicholas G. Rupp. 2006. “Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment.” *Quarterly Journal of Economics*, 121(2): 747-782.

Loury, Linda D. and David Garman. 1995. “College Selectivity and Earnings.” *Journal of Labor Economics*, 13(2): 289-308.

Loury, Linda D. 1997. “The Gender-Earnings Gap Among College-Educated Workers.” *Industrial and Labor Relations Review*, 50(4): 580–593.

Mobius, Markus M., and Tanya S. Rosenblat. 2006. "Why Beauty Matters." *American Economic Review*, 96(1): 222–235.

Mocan, Naci and Erdal Tekin. 2010. "Ugly Criminals." *The Review of Economics and Statistics*, 92(1): 15-30.

Ravina, E. 2009. “Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets.” Working Paper.

Rhodes, Gillian, Fiona Proffitt, Jonathon M. Grady, and Alex Sumich. 1999. “Facial Symmetry and the Perception of Beauty.” *Psychonomic Bulletin & Review*, 5(4): 659-669.

Rhodes, Gillian. 2006. “The Evolutionary Psychology of Facial Beauty.” *Annual Review of Psychology*, 57: 199–226.

Rohde, Treena E. and Lee A. Thompson. 2006. “Predicting Academic Achievement with Cognitive Ability.” *Intelligence*, 35(1):83-92.

Scholz, John Karl and Kamil Sicinski. 2011. “Facial Attractiveness and Lifetime Earnings: Evidence from a Cohort Study.” Working paper.

Wilson, R. K., and C. C. Eckel. 2006. "Judging a Book by its Cover: Beauty and Expectations in a Trust Game." *Political Research Quarterly*, 59(2): 189-202.

Wiswal, Matthew and Basit Zafar, 2012. “Determinants of College Major Choice: Identification using an Information Experiment.” Working paper.

Appendix A. Rating procedure and instructions

All alumna pictures were rated by at least 25 female and 25 male raters. Raters were students at a college in a different state and were pre-screened to ensure that they were not familiar with students from the college of interest.

Raters were shown pictures of each student and asked to rate her physical appearance on a 1–10 point scale. Five of the numbers had descriptions describing the level of attractiveness corresponding to that number (see experimental instructions on the next page). Raters were instructed to choose the numbers without descriptions if they felt the student's appearance fell between the two descriptions.

Each rater was shown four sets of about 100 photos. The order of the photos within each set was randomized for each rater. In early stages of the experiment, we compared the mean and standard deviation of ratings across photo sets to see if having subjects rate 400 pictures led to fatigue. There was no significant difference in either the mean or standard deviation of ratings for earlier and later sets, which led us to conclude that 400 pictures was not an excessive number. We did not use data from three raters who chose 1's 40% or more of the time. The "1" option was the closest to the "Next" button. Thus, these subjects were most likely trying to complete the experiment as quickly as possible.

Instructions for the experiment

You are about to participate in an experiment involving the perception of appearance. Once the experiment begins, you will see a photograph of an individual along with the following prompt:

> Rate this person's physical appearance using the following scale:
>
> 10 strikingly handsome or beautiful
> 9
> 8 good-looking (above average for age and sex)
> 7
> 6 average looks for age and sex
> 5
> 4 quite plain (below average for age and sex)
> 3
> 2 homely
> 1

Choose the number that best corresponds to your evaluation. Choose the numbers without descriptive text (1, 3, 5, 7, and 9) if you feel the person's appearance falls between the descriptions found in the adjacent numbers.

After you have chosen a number, click "Next." You will then see another photograph and be asked to repeat the procedure. Continue selecting the number you feel best reflects your assessment of the individual's appearance until you are told to stop.

Appendix B. Major, course, and occupation classifications

Major and course classifications

Humanities	Social Sciences
Art - History	Africana Studies
Art - Studio	Anthropology
Chinese	Environmental Studies
Cinema and Media Studies	History
Classical Civilization	International Relations
Comparative Literature	Peace & Justice Studies
English	Philosophy
French	Political Science
German	Psychology
Greek	Religion
Italian Studies	Sociology
Japanese	Women's Studies
Latin	Women's and Gender Studies
Media Arts and Sciences	
Medieval/Renaissance Studies	
Music	
Russian	
Spanish	
Theater Studies	

Area Studies	Science
American Studies	Astronomy
Chinese Studies	Astrophysics
Classical & Nr Eastern	Biological Chemistry
East Asian Studies	Biological Sciences
French Cultural Studies	Chemical Physics
German Studies	Chemistry
Jewish Studies	Cognitive & Linguistic Sciences
Latin American Studies	Computer Science
Middle Eastern Studies	Geology
Russian Area Studies	Geosciences

South Asia Studies	Mathematics
	Neuroscience
	Physics
Other	**Economics**
Archeology	Economics
Architecture	
Education	
Engineering	
Linguistics	
Military science	
Physical Education	
Urban studies	

Occupation classifications

Advertising/art	**Scientist**
Advertising/marketing	Researcher (except economics)
Political (non-government employee)	Earth sciences
Design	Chemistry
Non-technical writing	Biology
Architect	Astronomy
Performing	Physics
Publishing/broadcasting	Mathematics
Museums/galleries	Engineering

Technician	**Consultant/manager**
Technician	Manager
Paralegal/legal assistant	Consultant
Technical writer	Analysis
Computer-related work	Finance
	Economist

Administrator/retail	**Lawyer**
Administrative/human resources	Lawyer
Retail	

Teacher	**Physician**
Elementary, middle, high school	Physician/doctor
	Dentist
	Psychologist

Other medical
Nurse
Health worker
Physical therapist
Veterinarian