Economics Letters 219 (2022) 110782

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Student beauty and grades under in-person and remote teaching[☆]

Adrian Mehic

Department of Economics, Lund University, Sweden

ARTICLE INFO

Article history:
Received 23 June 2022
Received in revised form 25 July 2022
Accepted 1 August 2022
Available online 6 August 2022

JEL classification:
D91
I23
J16
Z13

Keywords:
Attractiveness
Beauty
COVID-19
Discrimination

ABSTRACT

This paper examines the role of student facial attractiveness on academic outcomes under various forms of instruction, using data from engineering students in Sweden. When education is in-person, attractive students receive higher grades in non-quantitative subjects, in which teachers tend to interact more with students compared to quantitative courses. This finding holds both for males and females. When instruction moved online during the COVID-19 pandemic, the grades of attractive female students deteriorated in non-quantitative subjects. However, the beauty premium persisted for males, suggesting that discrimination is a salient factor in explaining the grade beauty premium for females only.

1. Introduction

It is well-known that physical appearance is an important predictor for success in life. Attractive people are more satisfied with their lives, earn higher wages and grades, and are less likely to engage in criminal activity (Mocan and Tekin, 2010; Hamermesh, 2011). However, the explanation for the beauty premium is subject to debate, where the traditional viewpoint according to which it is a consequence of taste-based discrimination (Hamermesh and Biddle, 1994; Scholz and Sicinski, 2015) is increasingly challenged by findings suggesting that beauty is a productive attribute (Cipriani and Zago, 2011; Stinebrickner et al., 2019). As an example of the latter, attractive individuals are likely to be more self-confident, which can positively affect human capital formation (Mobius and Rosenblat, 2006).

In this paper, I use data from mandatory courses within a Swedish engineering program to examine the role of student facial attractiveness on university grades. I first consider academic outcomes when education is in-person, and the faces of students are readily available to teachers. The results suggest that beauty is positively related to academic outcomes, however, the results are only significant in non-quantitative courses, which to a greater extent rely on interactions between teachers and students. The beauty premium on grades in non-quantitative subjects hold for both male and female students. Then, using the COVID-19 pandemic as a natural experiment, and utilizing a difference-in-difference framework, I show that switching to full online teaching resulted in deteriorated grades in non-quantitative courses for attractive females. However, there was still a significant beauty premium for attractive males.

Taken together, these findings suggest that the return to facial beauty is likely to be primarily due to discrimination for females, and the result of a productive trait for males. The former result in line with the findings by Hernández-Julián and Peters (2017), while the latter is new to the literature. An advantage with the empirical strategy of this paper is that the switch to online teaching during the pandemic enables us to more credibly isolate the effect of appearance. This is because only the mode of instruction changed, and not the structure of the courses. Additionally, my identification strategy removes the problem of self-selection into courses.

The reminder of the paper is structured as follows. Section 2 describes the setting. Section 3 describes the data, while Section 4 presents the empirical strategy, and the results. The paper concludes with Section 5.

2. Setting

The Industrial Engineering Program (denoted I) at Lund University is a five-year program, leading to an MA in Engineering.

☆ I am grateful to the Editor, as well as one anonymous referee. I thank Annette Brondin, Ivan Premovski, Hanna M. Stålhammar, and Sara Swartling for providing assistance in recruiting raters. This work has benefited from discussions with Niclas Berggren, Andreas Bergh, Joakim Westerlund, and numerous seminar participants. I gratefully acknowledge financial support from the Arne Ryde Foundation.

E-mail address: adrian.mehic@nek.lu.se.

https://doi.org/10.1016/j.econlet.2022.110782

Table 1
Course structure.

First year				
	Period	Subject	Gender of instructor	Mode of examination
Calculus in One Variable	Sep–Dec	Mathematics	M	Final exam
Industrial Engineering	Sep–Dec	Business	W	Final exam, seminars, oral presentations, group assignments
Linear Algebra	Nov–Dec	Mathematics	M	Final exam
Multivariable Calculus	Jan–Mar	Mathematics	M	Final exam
Classical Mechanics	Apr–May	Physics	M (15–16) W (17–)	Final exam
Energy and Environmental Physics	Apr–May	Physics	W (15–17) M (18–)	Final exam, group assignments
Second year				
	Term	Subject	Gender of instructor	Mode of examination
Microeconomic Theory	Sep–Oct	Economics	M	Final exam, group assignments
Supply Chain Management	Sep–Oct	Business	W	Final exam, seminars, oral presentations, group assignments
Mathematical Statistics	Sep–Dec	Mathematics	M	Final exam
Marketing	Nov–Mar	Business	M	Final exam, seminars, oral presentations, group assignments
Programming	Nov–Mar	Programming	W	Final exam, seminars, oral presentations, group assignments
Complex Analysis	Jan–Mar	Mathematics	M	Final exam
Industrial Engineering, advanced course	Apr–May	Business	M	Final exam, seminars, oral presentations, group assignments
Systems and Transforms	Apr–May	Mathematics	M	Final exam

Note. The table shows the course structure for the Industrial Engineering program.

The number of students admitted each Fall is about 100. The first two years consist of a total of 15 mandatory courses in mathematics, physics, computer science, business, and economics, after which students choose one specialization track. Thus, to avoid selection bias, I restrict the sample to include the first two years of the program. To evaluate heterogeneous effects, I classify courses as either quantitative or non-quantitative; all mathematics and physics courses are classified as quantitative, and the reminder are considered non-quantitative. Non-quantitative courses have a higher share of group assignments, seminars, and oral presentations, whereas mathematics and physics courses rely almost exclusively on final written exams. Thus, in non-quantitative subjects, teachers are more likely to interact with and ''get to know'' students, making it reasonable to expect that the beauty premium is higher in non-quantitative courses. Table 1 outlines the course structure for the first two years of the program, and provides additional details.

Towards the end of the 2019–20 academic year, all Swedish universities switched to online teaching to mitigate the spread of COVID-19. The start date for these measures was March 17, 2020, and the measures were in place until the end of the 2020–21 academic year, that is, in May 2021. Consequently, students who started the program in 2018 had two online courses in their second year, whereas students starting in 2019 had two online courses in their first year, and eight online courses in their second year. At the time of the switch, the first part of the Spring semester had just finished, and the second part of the Spring semester had not yet started. Thus, there were no courses in which both on-campus and online teaching was used.

During online teaching, the course structures as presented in Table 1 were left unchanged, however, written individual exams were conducted via Zoom with students required to have two cameras turned on. Similarly, lectures and seminars were conducted remotely, and although students were encouraged to have their cameras turned on, there was no formal requirement to do so. Hence, online learning significantly reduced teacher–student interaction. Taken together, the identification strategy allows us to causally distinguish between the part of beauty premium due to taste-based discrimination, and the part of beauty premium that is due to productivity.

3. Data

For all courses, passing grades are given by 3, 4, and 5, where 5 is the top grade.[1] The grading scale is absolute, meaning that the cutoff level for each grade is determined before the start of the course, and is not affected by the relative performance of students. I use data from five cohorts, namely from students starting their studies in 2015, 2016, 2017, 2018, and 2019. The cohorts are denoted I15, I16, I17, I18, and I19, respectively. In total, the full sample includes 307 students. To facilitate interpretation, I standardize all grade data so that the sample mean is equal to zero, and the sample standard deviation is equal to unity.

To quantify beauty, I recruit a jury consisting of 74 individuals. Due to the large number of students in the sample, each jury member rates one-half of the sample only. Thus, each face receives an average of 37 independent ratings. By using publicly available pictures of all students, I let each juror grade the faces using a scale from 1 to 10, where 1 is extremely unattractive,

[1] The failing mark is ''U'' (Swed. *underkänd*, meaning ''failed''). If a student did not take an exam, that particular course is treated as a missing observation. If the student did take the exam, but failed, I assign the value 1 to the course in question.

Table 2
Grades and attractiveness: Pre-pandemic estimates.

Outcome variable: Standardized grades	All courses (1)	(2)	Quantitative courses only (3)	(4)	Non-quantitative courses only (5)	(6)
Grade_{t-1}	0.097	0.104	−0.518*	−0.533**	−0.276	−0.401
	[0.551]	[0.511]	[0.055]	[0.043]	[0.283]	[0.140]
Attractiveness	0.034	0.032	0.047	0.061	0.076***	0.054***
	[0.161]	[0.120]	[0.329]	[0.130]	[0.006]	[0.008]
Course FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Method	GMM	GMM	GMM	GMM	GMM	GMM
Observations	3,085	3,065	1,814	1,801	1,271	1,264
Mean dep. var.	0.000	0.000	0.000	0.000	0.000	0.000
Hansen J test p-value	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
AR(2) test p-value	[0.20]	[0.20]	[0.15]	[0.12]	[0.63]	[0.98]

Note. Outcome variable: Standardized grades. Controls: The student's gender, age, average parental income, median income of the student's home municipality, and the gender of the instructor. Standard errors are clustered by the $G = 5$ cohorts. To adjust for the small number of clusters, p-values are calculated using a t-distribution with $G - 1$ degrees of freedom (Cameron and Miller, 2015). *, **, and *** denote significance at the 10&, 5&, and 1% level, respectively.

and 10 is extremely attractive. Intercoder reliability was excellent (Cronbach's alpha = 0.94). Again, I standardize the jury ratings, so that the mean and standard deviation of the full sample is equal to zero and one, respectively.

As control variables, I include the student's age, gender, and for each course, the gender of the professor. To account for socioeconomy, I include the average taxable income of both parents, and the median income of the student's home municipality. Figure A.1 of Online Appendix A illustrates the histogram of estimated beauty over each of the five cohorts, whereas Table A.1 presents the summary statistics. Online Appendix B presents the data sources for all variables, and provides additional definitions.

4. Empirical strategy and results

4.1. Pre-pandemic estimates

I begin by examining the link between appearance and grades when teaching is fully in-person. This relationship can be estimated using the AR(1) dynamic panel model

$$y_{ict} = \alpha_{ic} + \phi y_{ic,t-1} + \beta X_{ic} + \boldsymbol{\gamma}' \boldsymbol{W}_{ic} + \omega_t + \varepsilon_{ict} \tag{1}$$

Here, y_{ict} is the grade of student i in cohort c in the course (subject) t, α_{ic} are student fixed effects, X_{ic} is the beauty of student i in cohort c, $\boldsymbol{W}_{ic}$ is a vector of student-level controls, ω_t are course fixed effects, and ε_{ict} is an idiosyncratic error term. In this specification, the coefficient of interest is β.[2] Since OLS estimates of the dynamic panel model are biased and inconsistent, I use the system GMM of Blundell and Bond (1998) to estimate β.

Table 2 presents the results. When all courses in the program are considered, there is a positive, albeit statistically insignificant relationship between attractiveness and grades. However, when using the division of courses into quantitative and non-quantitative, the coefficient for attractiveness is highly significant for the non-quantitative courses. The results suggest that one standard deviation higher beauty is associated with around 0.08σ higher grades. The magnitude of the estimated coefficient is slightly lower when the full set of controls is included. Concomitantly, there is no significant relationship between attractiveness and grades for the non-quantitative courses. Given the lower weight on teacher–student interactions in mathematics and physics teaching outlined previously, this finding is expected.[3]

As noted in Table 1, the first year physics course Energy and Environmental Physics is classified as a quantitative course, although it has some element of group assignments. Hence, we can expect somewhat more teacher–student interaction in this course compared to other quantitative courses. Table A.2 of Online Appendix A presents the main results with this course instead classified as a non-quantitative course. The magnitude of the beauty effect in the non-quantitative courses decreases slightly, although the coefficient estimates are still highly significant, and the conclusions outlined previously are robust to this change. As an additional robustness check, it is possible to show that the exclusion of the lagged grade in (1) does not impact the coefficient estimates.[4]

4.2. Difference-in-difference estimates

Having established that beauty is significantly related to grades when the faces of students are visible to teachers, I now re-visit this relationship when teaching is fully online, as was the case during the COVID-19 pandemic. Figure A.2 of Online Appendix A illustrates the parallel trends plots, one comparing the I15–I17 cohorts, who had finished their first two years before the pandemic, with I18, and one comparing I15–I17 with I19. The parallel trends plot show no indications of pre-trends.

Having established the absence of pre-trends, I now continue with the difference-in-difference estimates. Let the binary variable *Online* be equal to zero for all courses taken before March 17, 2020, and unity after this date. Table A.3 of Online Appendix A shows the results when interacting *Online* with standardized beauty without differentiating between course types. The results suggest that the switch to online learning did not result in an overall deterioration of the grades of high-attractive students. In most of the quantitative courses, teachers are not likely to interact much with students, so this finding does not come as a major surprise. Building on this, Table 3 presents the results when

2 The coefficient ϕ measures the impact of the lagged grade, that is, the grade obtained by the student in the previous course. It is included to account for the significant one-period autocorrelation present in the data. However, I will show that the main findings of the paper are not sensitive to the inclusion of the lagged dependent variable.

3 An additional regression, in which (1) is augmented to include the interaction between student gender and attractiveness, suggests no heterogeneous effects with respect to student gender.

4 Overall, the coefficient estimates are slightly lower when excluding the autoregressive term. Exclusion of the AR(1) term yields a coefficient estimate $\hat{\beta}$ of 0.059 ($p = 0.009$) for the non-quantitative courses (with no controls included), and 0.037 ($p = 0.361$) for the quantitative courses. These results are very similar to those obtained previously.

Table 3
Grades and attractiveness: Difference-in-difference estimates.

Outcome variable: Standardized grades	All students (1)	(2)	Male students (3)	(4)	Female students (5)	(6)
Grade_{t-1}	0.001	−0.028	−0.042	−0.070	−0.084	−0.100
	[0.993]	[0.730]	[0.837]	[0.737]	[0.772]	[0.737]
Online	0.005	−0.081	0.015	−0.101	0.053	−0.023
	[0.933]	[0.181]	[0.866]	[0.224]	[0.254]	[0.730]
Non-quantitative course	−0.336**	−0.520	0.127	1.090*	−0.062	0.173
	[0.045]	[0.357]	[0.587]	[0.078]	[0.730]	[0.822]
Attractiveness	0.032	0.033	0.082	0.079	0.000	0.008
	[0.453]	[0.387]	[0.479]	[0.490]	[1.000]	[0.888]
Online × Attractiveness	−0.002	0.020	0.002	0.009	−0.017	0.019
	[0.970]	[0.643]	[0.985]	[0.925]	[0.744]	[0.779]
Online × Non-quantitative course	−0.069	0.018	−0.071	0.003	0.048	0.064
	[0.419]	[0.815]	[0.612]	[0.985]	[0.534]	[0.474]
Attractiveness × Non-quantitative course	0.008	0.006	−0.038	−0.040	−0.003	−0.008
	[0.866]	[0.895]	[0.689]	[0.663]	[0.881]	[0.723]
Online × Attractiveness × Non-quantitative course	0.043	0.042	0.091	0.102	−0.061**	−0.060***
	[0.387]	[0.383]	[0.396]	[0.183]	[0.029]	[0.005]
Course FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes
Method	GMM	GMM	GMM	GMM	GMM	GMM
Observations	3,992	3,950	2,428	2,400	1,564	1,550
Mean dep. var.	0.000	0.000	0.000	0.000	0.000	0.000
Hansen J test p-value	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]
AR(2) test p-value	[0.12]	[0.16]	[0.82]	[0.87]	[0.63]	[0.67]

Note. Outcome variable: Standardized grades. Controls: The student's age, average parental income, median income of the student's home municipality, and the gender of the instructor. Standard errors are clustered by the $G = 5$ cohorts. To adjust for the small number of clusters, p-values are calculated using a t-distribution with $G - 1$ degrees of freedom (Cameron and Miller, 2015). *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.

augmenting the model by including the indicator variable for non-quantitative course. Now, the triple interaction between *Online*, attractiveness, and the indicator for non-quantitative course is highly significant for female students. This finding suggests that the grades of female students deteriorated in non-quantitative subjects, with grades declining with attractiveness. There is no equivalent relationship for males.

Table A.4 of Online Appendix A re-estimates (1) for the non-quantitative courses when education is online, interacting standardized beauty with gender. While the coefficient for beauty is positive, the interaction coefficient between beauty and female gender is negative. This finding suggests that the beauty premium is present only for males, and is consistent with the difference-in-difference results. Taken together, these results suggest that the beauty premium in education is due to discrimination for females, whereas for male students, it is primarily the result of a productivity-enhancing attribute.

Why is beauty a productivity-enhancing attribute for males in non-quantitative subjects? Generally, it is difficult to disentangle the reasons behind why beauty improves productivity (Hamermesh and Parker, 2005). However, relative to other students, attractive men are more successful in peer influence, and are more persistent, a personality trait positively linked to academic outcomes (Dion and Stein, 1978; Alan et al., 2019). In addition, attractive individuals are more socially skilled, have more open social networks, and are more popular vis-à-vis physically unattractive peers (Feingold, 1992). Importantly, possession of these traits is significantly linked to creativity (Soda et al., 2021). In our setting, the tasks faced by students in non-quantitative subjects, for instance in marketing and supply chain management, are likely to be seen as more "creative", and significantly contrast the more traditional book-reading and problem-solving in mathematics and physics courses, the latter presumably perceived as more monotonous. Together with the large use of group assignments in non-quantitative courses, these theoretical results imply that socially skilled individuals are likely to have a comparative advantage in non-quantitative subjects.

5. Concluding remarks

This paper has shown that students' facial attractiveness impact academic outcomes when classes are held in-person. As education moved online following the onset of the pandemic, the grades of attractive female students deteriorated. This finding implies that the female beauty premium observed when education is in-person is likely to be chiefly a consequence of discrimination. On the contrary, for male students, there was still a significant beauty premium even after the introduction of online teaching. The latter finding suggests that for males in particular, beauty can be a productivity-enhancing attribute.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econlet.2022.110782.

References

Alan, Sule, Boneva, Teodora, Ertac, Seda, 2019. Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. Q. J. Econ. 134 (3), 1121–1162.
Blundell, Richard, Bond, Stephen, 1998. Initial conditions and moment restrictions in dynamic panel data models. J. Econometrics 87 (1), 115–143.
Cameron, A. Colin, Miller, Douglas L., 2015. A practitioner's guide to cluster-robust inference. J. Hum. Resour. 50 (2), 317–372.

Cipriani, Giam Pietro, Zago, Angelo, 2011. Productivity or discrimination? Beauty and the exams. Oxf. Bull. Econ. Stat. 73 (3), 428–447.
Dion, Karen K., Stein, Steven, 1978. Physical attractiveness and interpersonal influence. J. Exp. Soc. Psychol. 14 (1), 97–108.
Feingold, Alan, 1992. Good-looking people are not what we think. Psychol. Bull. 111 (2), 304–341.
Hamermesh, Daniel S., 2011. Beauty Pays: Why Attractive People are more Successful. Princeton University Press, Princeton, NJ.
Hamermesh, Daniel S., Biddle, Jeff E., 1994. Beauty and the labor market. Amer. Econ. Rev. 84 (5), 1174–1194.
Hamermesh, Daniel S., Parker, Amy, 2005. Beauty in the classroom: Professors' pulchritude and putative pedagogical productivity. Econ. Educ. Rev. 24 (4), 369–376.
Hernández-Julián, Rey, Peters, Christina, 2017. Student appearance and academic performance. J. Hum. Cap. 11 (2), 247–262.
Mobius, Markus M., Rosenblat, Tanya S., 2006. Why beauty matters. Amer. Econ. Rev. 96 (1), 222–235.
Mocan, Naci, Tekin, Erdal, 2010. Ugly criminals. Rev. Econ. Stat. 92 (1), 15–30.
Scholz, John Karl, Sicinski, Kamil, 2015. Facial attractiveness and lifetime earnings: Evidence from a cohort study. Rev. Econ. Stat. 97 (1), 14–28.
Soda, Giuseppe, Vittorio, Mannucci Pier, Burt, Ronald S., 2021. Networks, creativity, and time: Staying creative through brokerage and network rejuvenation. Acad. Manage. J. 64 (4), 1164–1190.
Stinebrickner, Ralph, Stinebrickner, Todd, Sullivan, Paul, 2019. Beauty, job tasks, and wages: A new conclusion about employer taste-based discrimination. Rev. Econ. Stat. 101 (4), 602–615.