

Disentangling the Drivers of Grading Bias: Evidence from a Natural Experiment

Teacher-awarded grades can act as crucial signals about academic skills, which affects students' choice of secondary and tertiary education (Holm, Hjorth-Trolle, and Jæger 2019; Rangvid 2015). However, on average, boys, ethnic/racial minorities, and students with low parental education receive lower grades from their teacher than in external assessments (e.g., Burgess and Greaves 2013; De Benedetto and De Paola 2023; Lavy 2008; Rangvid 2015; Triventi 2020). These lower grades may then misinform students about their academic potential and cause them to lower their ambitions, leading to social inequality in educational attainment (Rangvid 2015) with implications for social mobility.

In this study, we exploit a natural experiment to estimate the role for teacher bias of statistical discrimination, which arises because teachers use observable group membership as a proxy for unobserved skills (Arrow 1973; Phelps 1972). In 2010, the Danish government introduced a mandatory, external, and standardized test in the 8th grade, the results of which the teacher receives. If grading disparities are mainly driven by statistical discrimination, standardized tests will reduce teacher bias because the test results reveal otherwise unobserved skills, eliminating the need to use group membership as a proxy.

We conform to the existing literature on teacher bias by estimating the following model:

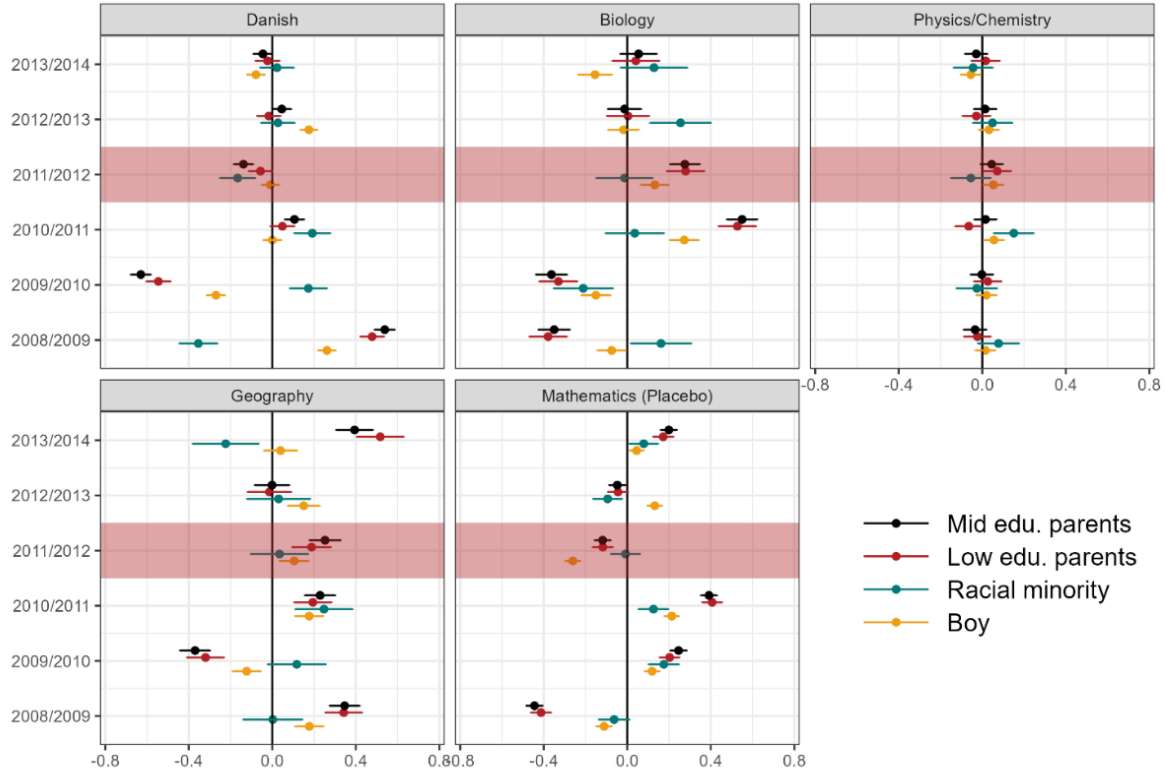
$$TG_{ij} = \beta_0 + \beta_1 Male_{ij} + \beta_2 Minority_{ij} + \beta_3 Edu_{ij} + \beta_4 EG_{ij} + \varepsilon_{ij} \quad (1)$$

where the outcome, TG_{ij} , is the teacher-awarded grade for student i in subject j . $Male_{ij}$ and $Minority_{ij}$ measure gender and migration background and Edu_{ij} is a vector of two dummies measuring whether parents have either no or some secondary education (the reference group is parents with tertiary education). EG_{ij} measures exam grades. We estimate the model for five subjects: Danish, biology, Physics/Chemistry, Geography, and mathematics. The tests gave teachers information about student skills in each of these subjects, except mathematics, which acts as a

placebo. If β_1 is negative, boys on average receive lower teacher-awarded grades compared to girls with the same exam grade, which we interpret as indication of teacher bias. The interpretation of β_2 and β_3 is similar. To disentangle statistical discrimination, we estimate equation 1 for each pair of years between 2008-2014 and test for significant changes in β_1 , β_2 , and β_3 between 2010 and 2011 when the test was introduced. We use Danish administrative registers on teacher-awarded grades and exam grades, both awarded in 9th grade in the years 2008-2014. We link these grades to information on students' sex, migration history, and parental education ($n \sim 190,000 - 440,000$).

Figure 1 shows that for the year-pair 2010/2011, when the tests were introduced, the changes in grading disparities are indiscernible from (a) the other years where nothing changed and (2) from mathematics, where no 8th grade test was introduced. We interpret this result as evidence that statistical discrimination is unlikely to be the main driver of teacher bias.

Figure 1: Yearly changes in grading disparities by subject and social background



References

- Arrow, Kenneth J. 1973. "The Theory of Discrimination." Pp. 3–33 in *Discrimination in Labor Markets*, edited by O. Ashenfelter and A. Rees. Princeton University Press.
- Burgess, Simon, and Ellen Greaves. 2013. "Test Scores, Subjective Assessment, and Stereotyping of Ethnic Minorities." *Journal of Labor Economics* 31(3):535–76. doi: 10.1086/669340.
- De Benedetto, Marco Alberto, and Maria De Paola. 2023. "Immigration and Teacher Bias towards Students with an Immigrant Background." *Economic Policy*. doi: 10.1093/epolic/eiac068.
- Holm, A., Hjorth-Trolle, A., Jæger, M. 2019. "Signals, Educational Decision-Making, and Inequality", *European Sociological Review*, 35:4, pp447-460
- Lavy, Victor. 2008. "Do Gender Stereotypes Reduce Girls' or Boys' Human Capital Outcomes? Evidence from a Natural Experiment." *Journal of Public Economics* 92(10):2083–2105. doi: 10.1016/j.jpubeco.2008.02.009.
- Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62(4):659–61.
- Rangvid, Beatrice Schindler. 2015. "Systematic Differences across Evaluation Schemes and Educational Choice." *Economics of Education Review* 48:41–55. doi: 10.1016/j.econedurev.2015.05.003.
- Rangvid, Beatrice Schindler. 2019. "Gender Discrimination in Exam Grading? Double Evidence from a Natural Experiment and a Field Experiment." *The B.E. Journal of Economic Analysis & Policy* 19(2). doi: 10.1515/bejeap-2018-0084.
- Triventi, Moris. 2020. "Are Children of Immigrants Graded Less Generously by Their Teachers than Natives, and Why? Evidence from Student Population Data in Italy." *International Migration Review* 54(3):765–95. doi: 10.1177/0197918319878104.