NBER WORKING PAPER SERIES

THE GENDER CONCEALMENT GAP

Christine L. Exley
Raymond Fisman
Judd B. Kessler
Louis-Pierre Lepage
Xiaomeng Li
Corinne Low
Xiaoyue Shan
Mattie Toma
Basit Zafar

Working Paper 32350
http://www.nber.org/papers/w32350

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2024, revised August 2025

This paper previously circulated as "Information-Optional Policies and the Gender Concealment Gap". This paper supersedes material previously included in "The Transparency Gap" and "Anticipated Discrimination and Grade Disclosure". We acknowledge the generous cooperation of the Boston University Registrar in obtaining the anonymized student transcript data. We thank Annabelle Finlayson, Tomer Mangoubi, John-Henry Pezzuto, and Emma Ronzetti for excellent research assistance. We thank numerous seminar and conference participants for valuable feedback and comments. This project was supported by the Harvard Business School, the Wharton School, the Wharton Analytics Initiative, and the Wharton Behavioral Lab. It was also supported through a Quartet Pilot Research award funded by the Boettner Center at the University of Pennsylvania. The content is solely the responsibility of the author(s) and does not necessarily represent the official views of Boston University, the University of Pennsylvania, the National Institutes of Health, or the National Bureau of Economic Research.
At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w32350

The Gender Concealment Gap
Christine L. Exley, Raymond Fisman, Judd B. Kessler, Louis-Pierre Lepage, Xiaomeng Li, Corinne Low, Xiaoyue Shan, Mattie Toma, and Basit Zafar
NBER Working Paper No. 32350
April 2024, revised August 2025
JEL No. D82, J16, J71

ABSTRACT

We identify and explore a gender concealment gap when individuals have the opportunity to hide information about their performance from others. In data from two universities that allowed students to replace letter grades with "credit" on their transcripts, we find that men are substantially more likely than women to conceal grades that will harm their GPAs. The gender concealment gap persists across student traits and course features and generates inequity: the option to conceal leads to GPA gains that are 50% larger for men than for women. University data and complementary experimental evidence suggests that women may conceal less because they expect others will make worse inferences about their concealed grades.

Christine L. Exley
University of Michigan
Department of Economics
clexley@gmail.com

Raymond Fisman
Boston University
Department of Economics
and NBER
rfisman@bu.edu

Judd B. Kessler
University of Pennsylvania
The Wharton School
and NBER
judd.kessler@wharton.upenn.edu

Louis-Pierre Lepage
Stockholm University
louis-pierre.lepage@sofi.su.se

Xiaomeng Li
University of Michigan
Department of Economics
lixiaom@umich.edu

Corinne Low
University of Pennsylvania
The Wharton School
and NBER
corlow@wharton.upenn.edu

Xiaoyue Shan
National University of Singapore
x.shan@nus.edu.sg

Mattie Toma
University of Warwick
Mattie.Toma@wbs.ac.uk

Basit Zafar
University of Michigan
Department of Economics
and NBER
basitak@gmail.com

1 Introduction

There exist persistent gender gaps in pay and labor market representation (Goldin, 2014; Blau and Kahn, 2017). These gaps have given rise to a rich literature exploring the factors that contribute to them—ranging from gender differences in willingness to negotiate or compete (Niederle and Vesterlund, 2007; Hernandez-Arenaz and Iriberri, 2019*a*; Biasi and Sarsons, 2022; Roussille, 2024) to occupational sorting based on amenity tastes or other preferences (Fernandez and Freidrich, 2011; Petersen, Penner and Hogsnes, 2011; Zafar, 2013; Wiswall and Zafar, 2018).[1] Some of these gaps may be due to labor market discrimination, which research has shown is influenced by the amount of information available to employers.[2]

An important source of information about personal attributes comes from the individuals themselves. College applicants decide whether to share standardized test scores and what personal matters to discuss in their essays. Job applicants decide what information to include on their resumes and what to share with potential employers in cover letters and interviews. Workers decide what to reveal to managers in performance reviews. Employees decide what information about their performance to disclose to their teammates in group production. Professionals attempting to secure new business decide what information about their past performance to share with potential clients. Some of the information that individuals must decide whether or not to reveal is potentially harmful, so it is worth investigating whether men and women differ systematically in their likelihood to disclose harmful information about themselves when not obligated to do so.

A key limitation of research on concealing information—and whether decisions to conceal differ by gender—is that concealed information is typically unobserved by the researcher. To identify whether there is a gender gap in concealing poor performance, researchers would need an environment in which they could see information about underlying performance, identify men and women with equally poor performance, and then observe whether that poor performance was revealed to others. Ideally, the environment would be a high-stakes setting in which concealment decisions could have significant consequences.

We identify such an environment. Specifically, we leverage natural experiments at Boston

[1] Such preferences could relate to cultural beliefs and ideals (Correll, 2001; Charles and Bradley, 2009; Cech, 2013; Burbano, Padilla and Meier, Forthcoming). Additional factors contributing to gender gaps in pay and representation that have been discussed in the literature include gender differences in risk and social preferences (Croson and Gneezy, 2009); parenting responsibilities and other reproductive differences (Adda, Dustmann and Stevens, 2017; Kleven, Landais and Sogaard, 2019; Low, 2022); gendered expectations (Bursztyn, Fujiwara and Pallais, 2017); and psychological gender differences (Hyde, 2014).

[2] A number of empirical studies show that constraints on information (e.g., "ban the box" policies) can increase discrimination (Wozniak, 2015; Bartik and Nelson, 2016; Agan and Starr, 2017; Doleac and Hansen, 2017; Agan and Starr, 2018; Doleac and Hansen, 2020). More generally, how the availability of information influences discrimination is a frequent topic in the literature (Lundberg and Startz, 1983; Altonji and Pierret, 2001; Reuben, Sapienza and Zingales, 2014; Blair and Chung, 2021; Lepage, 2024).

University (BU) and a large flagship public school in the Midwest (hereafter the "Midwestern flagship"), which both introduced opportunities for students to conceal information on their grades. For one or two semesters at each school, students had the option of choosing—for each of their courses—whether or not to conceal a passing letter grade by deciding whether to have it officially recorded as "Credit."[3] When students concealed a passing letter grade, the grade no longer impacted their grade point average (GPA).

A key component of the policies at both universities is that students observed the assigned final letter grade they received for each class *before* deciding whether to conceal it. We are particularly interested in these policies since they allow us to analyze men and women who observe their performance with certainty and simply have to decide whether or not to reveal that information to others.[4] We study responses to this policy for the one semester it was in place at BU and for the two semesters it was in place at the Midwestern Flagship. Both schools provided us with information on assigned letter grades and whether each grade was converted to Credit—allowing us to observe the performance of students even when they decide to conceal their performance from others—as well as additional information on students and their courses.

As expected, students conceal a substantial fraction of grades that would lower their GPA; 52 percent of these "harmful" grades are concealed at BU and 33 percent of them are concealed at the Midwestern Flagship. For these harmful grades, we document a large and robust *gender concealment gap*: women are 8.9 percentage points (15%) less likely to conceal such grades than men at BU, and they are 8.3 percentage points (22%) less likely to do so at the Midwestern Flagship.

We estimate the consequences of this gender concealment gap. Because women conceal harmful grades less often than men do, the policies shift GPA distributions in favor of men. On average, due to the policies, GPAs of men improve by 0.07 points more than the GPAs of women over one semester at BU and also by 0.07 points over the two semesters at the Midwestern Flagship.[5] This relative GPA shift is substantial, eliminating more than half of the GPA advantage earned by female students in a typical semester prior to the introduction of the policies. This relative GPA shift in favor of men has the potential to impact a number

[3]These policies emerged unexpectedly in 2020, in recognition of the challenges faced by many students arising during the Covid-19 pandemic.

[4]This contrasts with policies where individuals must decide whether to conceal information before knowing with certainty what that information will be, allowing gender differences in beliefs about that information (e.g., gender differences in confidence) to also influence decisions to conceal (Kolb et al., 2023; Trost and Wooten, 2023). While these other settings are interesting and important, they introduce additional motives for concealing that we see as confounds in attempting to answer our research question.

[5]At BU, the policy improved men's GPAs by 0.23 points and women's GPAs by 0.16 points; at the Midwestern Flagship, the policy improved men's GPAs by 0.22 points and women's GPAs by 0.15 points.

of outcomes, such as qualification for awards, internships, majors, and jobs; future career and educational decisions; and others' perceptions about students' abilities.

Having established the existence of the gender concealment gap and its GPA impact, we seek to better understand what could be driving it. Features of the programs allow us to rule out some candidate explanations for the gap. For example, the gender concealment gap cannot be driven by gender differences in confidence or risk aversion about a particular course grade, since students learn their grade before deciding whether or not to conceal it. The robustness of the gap also shows that other explanations—such as those that predict the gap to arise, or to be dramatically larger, in only a subset of courses or among a subset of students—cannot fully explain the gap. Within each school, we find that the concealment gap is present across a range of student-level traits, majors, and course-level attributes. For example, we observe a strong and statistically significant concealment gap across years of study (from first-years to seniors); across prior GPA levels (from those with low to high GPAs); throughout a wide swath of the grade distribution (from B+ to C); across different potential impacts on overall GPA (from a GPA reduction of just 0.01 point to reductions even larger than 0.15 points); during different academic terms (Spring 2020, Fall 2020, and Spring 2021); and across classes that vary in terms of factors such as their size, measures of difficulty, demand for seats, and gender balance.

To further explore the potential drivers of this gap, we note that students may value what others—such as employers, graduate school admission committees, faculty considering them for research assistant positions, or faculty writing reference letters for internships or jobs—will infer about their performance if they conceal a grade. Since beliefs about the inferences others will draw from concealed grades cannot be observed from the observational data, we conduct a pair of surveys: one run with undergraduate students from the Midwestern Flagship and one run with a broader set of U.S. adults. Respondents consider a man and a woman with identical academic records who both concealed a grade for the same course. They are asked whether prospective employers or graduate school admission committees would infer that the man and the woman had the same or different grades. We then present a similar scenario in which the man and woman have identical resumes without a GPA listed. Respondents are asked whether others will infer that they have the same or different GPAs. In both scenarios, for both populations, we find that both male and female respondents believe that a more positive inference will be made about the man (for the latter exercise, responses are incentivized, as explained in Section 6). These findings suggest that students may believe others will infer that women who conceal a grade performed worse than men who conceal a grade. Such beliefs could help to explain why women conceal less often than men do. Even when concealing grades would be beneficial to their overall GPA, women

may conceal less often because they expect others will make worse inferences about their concealed grades.

In considering how our findings relate to gender gaps documented in prior work, earlier results could lead one to expect that women would be more likely—rather than less likely—to conceal performance information. Here, it is worth noting that our results focus on a setting in which women make decisions about whether to actively conceal performance while prior literature typically focuses on settings in which women make decisions about whether to actively reveal performance. For instance, women actively reveal their performance less than men do by negotiating less often (Hernandez-Arenaz and Iriberri, 2019*b*), competing less often (Niederle and Vesterlund, 2007, 2011), speaking up less often (Coffman, 2014), volunteering salary information less often (Agan, Cowgill and Gee, 2020), reporting their skills on resumes less often (Murciano-Goroff, 2022), and applying for challenging work less often (Coffman, Collis and Kulkarni, 2023).[6] In addition, the degree of ambiguity in our setting—in which it is clear whether or not a grade falls below one's GPA—may differ from settings explored in prior work in which women may lack performance information, including how it compares to relevant benchmarks.[7]

Our results also speak to the potential unintended consequences of policies that allow individuals to conceal information. Because men are more likely to conceal harmful grades than women in our context, these policies shift GPA distributions in favor of men, which might exacerbate labor market inequities. Our findings thus add support to the growing body of evidence on the possibly unintended effects of information-optional or information-restricted policies more generally.[8] This discussion is growing in importance as many universities have considered (and reconsidered) their information-optional policies, including test-optional ad-

[6]These prior findings could help to explain the results of an expert survey we conducted in which 64 economists (primarily experts in labor or education, more than half of whom work directly on gender-related topics) were asked to predict whether men or women would be more likely to conceal grades. Only 22% of experts correctly reported that female students would be less likely to conceal grades while 44% reported that men would be less likely to do so. In our survey, experts were asked the following: "Consider one male and one female student with identical cumulative GPAs and grades in a given course. Which student do you think would be more likely to mask their grade for the course?" Appendix Section E provides details on our expert survey.

[7]For instance, consistent with the importance of ambiguity in these decisions, Agan, Cowgill and Gee (2020) shows that workers report that knowing how a disclosure would affect their job outcomes is critical in helping them feel more comfortable with disclosing salary information. For additional work on how ambiguity affects gender gaps, see also Bowles and McGinn (2008) and Leibbrandt and List (2015).

[8]Agan, Cowgill and Gee (2020) provides a theoretical framework showing another information-optional policy (salary bans, in which it is still possible for workers to volunteer salary information) may have unintended equity consequences in the case of differential compliance by gender. Our work also relates to the unintended equity consequences of ostensibly *gender-neutral* policies more generally; for instance, Hirshman and Willén (2022) finds that a policy change that affects the risk of requesting a regrade augments a gender gap in regrade requests because of gender differences in risk perceptions.

missions and broader pass-fail policies.[9] Our findings also suggest that it will be valuable to further explore the presence and implications of concealment gaps in other settings, including, for example, decisions on what is included on a resume or job application, what is revealed in self-evaluations and performance reviews, and whether to report negative information to a supervisor.

Finally, our results from the belief studies contribute to the nascent but growing literature on expected gender discrimination (Alston, 2019; Dustan, Koutout and Leo, 2022; Aksoy, Chadd and Koh, 2023; Ruebeck, 2023), which may contribute to labor market outcomes (Gagnon, Bosmans and Riedl, 2022; Koutout, 2022) and educational choices (Ugalde, 2022; Lepage, Li and Zafar, 2025). Indeed, improving our understanding of when men and women expect to face gender discrimination may prove essential to understanding gender differences in disclosure decisions, negotiation strategies, educational tracks, and career choices.

2 Data

2.1 Grade-Optional Policy

In recognition of Covid-19's "disruptive, stressful, and unconventional" impact on both coursework and life more generally, the two institutions we study adopted special grading policies (Morrison, 2020). Specifically, for each course taken for a letter grade, students could choose to conceal any passing grade (any letter grade from A to D at Boston University

[9]For example, in 2019, 55 percent of colleges required standardized test scores for admission; by 2023, this figure had fallen to 4 percent. More recently, some schools have reinstated test requirements. A primary rationale for the initial shift to test-optional policies was to make college more inclusive to students who otherwise might be disadvantaged or deterred by standardized tests. But, as highlighted in recent work such as Conlin, Dickert-Conlin and Chapman (2023), the equity implications of these policies can depend on how different groups respond to them, a message that is reinforced by our own findings. Chetty, Deming and Friedman (2023) finds that relying more on test scores may actually benefit low- and middle-income students, who face a relative disadvantage in terms of other admissions-relevant factors (non-academic credentials, legacy preferences, and athlete recruitment). Along these lines, several recent studies have examined the effect of test-optional policies on student characteristics and test scores. Belasco, Rosinger and Hearn (2015) and Saboe and Terrizzi (2019) find that test-optional policy adoption has no effect on diversity in student enrollment. Bennett (2022) identifies a statistically-significant but small increase in the share of students who are Pell grant recipients, underrepresented minorities, and women, while Borghesan (2022) identifies a small increase in college attendance among low-income students. A recent theoretical literature also relates to optional-information policies. This strand of research identifies the conditions under which it is advantageous for policymakers, university officials, or other decision-makers to offer the option to reveal less information, for instance as a tool to reduce their "disagreement cost" with society (Hancart, 2023; Dessein, Frankel and Kartik, 2023).

and any grade from A to C- at the Midwestern Flagship[10]) by converting it into a grade of "Credit" (at BU) or "Pass" (at the Midwestern Flagship).[11] If a student chooses to conceal a letter grade in a course, the grade of "Credit" or "Pass" was recorded on their transcript so that the letter grade would not be visible to anyone reviewing their transcript. In addition, while credits from a course with a concealed grade would count towards a student's major and their progress through their undergraduate program in the same way as letter-graded credits, the concealed grade would not carry a point value, so it would be excluded from the student's GPA calculation. The grading policies were explained on each student's transcript.

Crucially for our analysis, at both schools and for each course, each student was assigned a final letter grade that they could observe before deciding whether or not to conceal it. This feature of the policies allows us to investigate the decision to conceal a particular grade from the transcript without any uncertainty about what the particular grade might be.[12]

We observe two semesters in which students were informed about the policy unexpectedly during these semesters. At BU, the policy was implemented in the Spring semester of 2020: it was announced on March 27 2020, and the last day of exams was May 9 2020. At the Midwestern Flagship, the policy was implemented in the Fall semester of 2020: it was announced on November 13 2020 and the last day of Fall exams was December 18 2020.

We also observe one semester in which the policy was announced prior to the start of the semester. Specifically, prior to the Spring semester of 2021, the Midwestern Flagship announced that the policy would continue that semester. The last day of classes that semester was April 30 2021.

We always analyze the data separately for BU and the Midwestern Flagship (and Section 5.7 further shows the results are robust to separately examining each semester at the Midwestern Flagship).

[10]At the Midwestern Flagship, there are two additional features of the grading policy relevant to our analysis. First, an A+ outside of the business school equals 4.0 GPA points and is thus equivalent to an A that also equals 4.0 GPA points. An A+ in the business school equals 4.4 GPA points, but we note that no students concealed a grade of an A+ in the business school. Thus, for our analyses, we simply treat A and A+ grades as equivalent. Second, students who conceal grades ranging from A+ to C- receive a "P" on their transcripts. However, for students who conceal a C-, a "PS" is noted in internal records to comply with degree audit requirements for courses that require a minimum grade of C. Excluding C- from our analysis does not change the results; the concealment gap in Table 1 is -0.083 when C- is included and -0.084 when C- is excluded.

[11]Our analyses exclude cases in which students did not have the opportunity to change their letter grade to a grade of "Credit." At BU, this applies to grades of F, which could only be changed to a "No Credit." At the Midwestern Flagship, this applies to grades of either D or F. For these grades at the Midwestern Flagship, the default is set to "No Record Covid," and students are only presented with an opportunity to change the "No Record Covid" to their letter grade of D or F, respectively. Fewer than 2% of students received such grades at either institution.

[12]This adds to prior work in which confidence about grades may be relevant since students do not know their letter grade before deciding whether or not to conceal it (Kolb et al., 2023; Trost and Wooten, 2023).

2.2 Data

We obtained administrative student-term level transcript records from both institutions. We received information on course enrollments, credits attempted, credits earned, and grades obtained. Importantly, the datasets record the original letter grades and the concealment decisions for each course. We also observe student demographics (including gender, race, and whether they are a first-generation college student) as well as some additional academic information (including year of study, major, and cumulative GPA). In total, we have information on over 15,000 students at BU and over 35,000 at the Midwestern Flagship, corresponding to roughly 60,000 and 260,000 passing grades at each institution that could potentially be concealed by students.

Appendix Table A.1 shows summary statistics of students at BU and the Midwestern Flagship split by student gender. Nearly 60% of students at BU are female and slightly above 50% of students at the Midwestern Flagship are female. At both schools, female students take slightly more credits and have GPAs that are about 0.10 to 0.13 points higher. We also see that GPAs are higher at the Midwestern Flagship than at BU.

3 Documenting the Concealment Gap

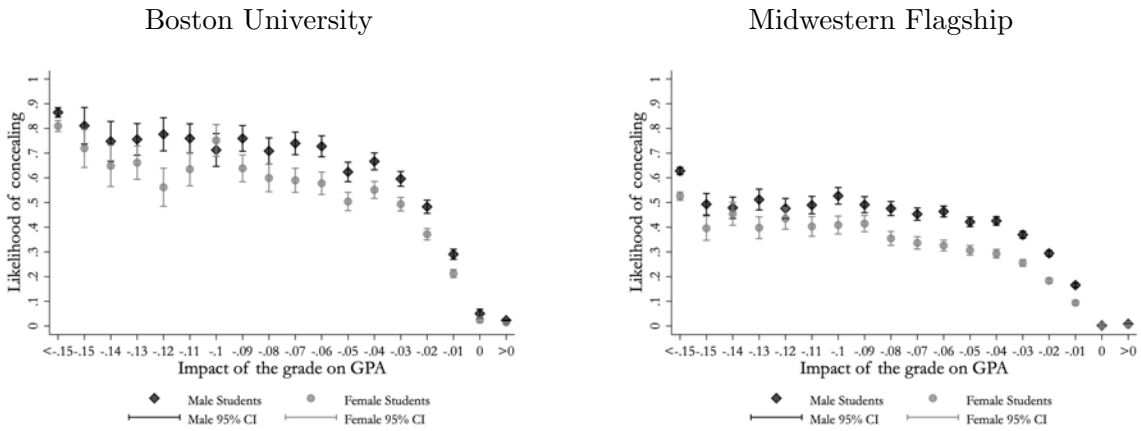
Students faced with the decision of whether to conceal a grade in a particular course may consider how the letter grade compares to their GPA. Grades above a student's GPA would improve their average (we call these "helpful" grades). Grades that are the same as a student's GPA (possible only when a student's GPA is exactly equal to the point value of a letter grade) would not change their average (we call these "neutral" grades). Grades below a student's GPA would pull down their average (we call these "harmful" grades).[13]

We begin by examining how students make their concealment decisions as a function of the decision's impact on the student's GPA, and whether those decisions differ by gender. Figure 1 shows the rates at which male and female students conceal their letter grades at Boston University (in the left panel) and at the Midwestern Flagship (in the right panel) by the "Impact of the grade on GPA" shown on the x-axis.

The rate of concealing helpful grades is shown on the far right of each panel of Figure 1, labeled with "> 0" (indicating that the grade increases the student's GPA). At both schools, the concealment rate is less than 2% for these helpful grades. The rate of concealing neutral

[13]When defining course grades in this way, we consider the student's GPA at the start of the semester as the relevant GPA. One might instead consider the GPA after accounting for courses that are not concealed by a student at the end of the semester. This complicates the analysis as the definition of whether a grade is helpful, neutral, or harmful (as well as its quantitative impact on GPA) might then depend on the student's decision of whether or not to conceal their other grades in that semester.

Figure 1: Decision to conceal by the grade's impact on GPA



This figure plots the likelihood of concealing a grade as a function of its impact on a student's GPA if the grade remained unconcealed, by male (diamond) and female (circle) students. "Impact of the grade on GPA" collapses ranges into specific points on the graph. The tick at -.01 includes course grades that would decrease the student's GPA by (0, -0.01], the tick at -.02 by (-0.01, -0.02], and so forth. The tick at <-.15 includes grades that would decrease the student's GPA by more than 0.15 points. The tick at 0 includes grades that would leave GPA unchanged and the one at >0 includes all grades that would increase GPA. Grades A to D are considered for Boston University, while grades A to C- are considered for the Midwestern Flagship. Error bars show 95% confidence intervals. This figure includes every grade obtained for each term covered by the policy at either institution, which involved 65,090 grades at BU from the decisions made by 15,690 students and 264,131 grades at the Midwestern Flagship from the decisions made by 37,574 students.

grades is shown next to the rate for helpful grades in Figure 1, labeled with "0" (indicating no impact on GPA). At both schools, the concealment rate is less than 5% for neutral grades. Given these low concealment rates of helpful and neutral grades, it is thus not surprising that there we see very similar rates of concealment by gender.[14]

By contrast, evidence for gender differences in concealment rates is substantial for harmful grades, which are much more likely to be concealed. Figure 1 shows—by plotting the concealment rates across a range of harmful grades—that the concealment gap arises for grades that are just slightly harmful and for grades that are more substantially harmful. Beginning with the slightly harmful grades, labeled with "−.01" (indicating that the grade decreases the student's GPA by $(0, 0.01]$ points), the concealment gap arises at both universities. At BU, men conceal 29% of these slightly harmful grades while women conceal 22% of them. At the Midwestern Flagship, men conceal 17% of slightly harmful grades while women conceal 10% of them. The concealment gap persists as the grades become more harmful and as the concealment rates themselves generally increase. Thus, there is a robust concealment gap across harmful grades: women are less likely than men to conceal harmful grades that

[14]While the point estimates are very small in these cases, there are still statistically significant gender differences in concealment, a point we return to in Section 6.1.

have various negative impacts on their GPA, including grades that would lower their GPA by more than 0.15 (shown at label "$< -.15$").

We note several further patterns across the two panels of Figure 1. First, concealment rates are more precisely estimated at the Midwestern Flagship, where we have roughly four times as much data as at BU. Second, there is a substantial level difference in the rates of concealment: students at the Midwestern Flagship are much less likely to conceal harmful grades than students at BU. Rather than focus on any number of factors that may contribute to level differences in the rates of concealment *across* these two universities, we believe the robustness of the concealment gap across these two universities is particularly notable given their differences. We will further document the robustness of the concealment gap to factors such as student demographics, student GPAs, and variations in how the grade-optional policies were announced.

Table 1 analyzes the same data in a regression framework. To compare men and women who are faced with an opportunity to conceal a grade that would have approximately the same impact on their GPA, our regressions include Year$\times\Delta$GPA FEs. These fixed effects include an indicator for the amount by which a student's GPA would change if the student took a class for a letter grade rather than concealing it, discretized into 1,000 intervals separately for each collegiate year. Allowing these fixed effects to vary by collegiate year is intended to capture the fact that the impact on GPAs is mechanically larger for first-year students than seniors because the former group has taken fewer courses. That said, our results remain substantial and statistically significant when we include GPA fixed effects but do not allow them to vary by collegiate year (as well as having no controls, adding in course and student level controls, adding in major fixed effects, and adding in course fixed effects).[15] Since each student makes multiple decisions in our data (i.e., one for each course), our regressions cluster standard errors at the student level.

Panel A of Table 1 presents results from BU and Panel B presents results from the Midwestern Flagship. Column 1 reveals a significant concealment gap among all grades that a student could conceal: relative to their male peers, female students are 3.4 percentage points less likely to conceal grades at BU and 2.8 percentage points less likely to conceal grades at the Midwestern Flagship. However, as evident from the concealment rates observed in Figure 1, pooling across all grades masks substantial and important heterogeneity. For neutral and helpful grades, since the rates of concealment are small—around 4% and 2% for BU and always less than 1% for the Midwestern Flagship—it is not surprising that only

[15]Building on the specification in Column 2 in Table 1, Appendix Table A.2 shows the robustness of our main result when we have no controls, add in course and student level controls, add in major fixed effects, and add in course fixed effects. See also the robustness of our results when examining whether the concealment gap persists across various groups of students and types of courses.

Table 1: Gender gap in concealing grades

Panel A: Boston University				
	(1)	(2)	(3)	(4)
	All	Harmful Grades	Neutral Grades	Helpful Grades
Female	-0.034***	-0.089***	-0.030*	-0.007***
	(0.003)	(0.009)	(0.016)	(0.002)
Observations	61,211	18,679	1,248	41,283
Conceal mean	0.172	0.522	0.043	0.018
Year × ΔGPA FEs	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship				
	(1)	(2)	(3)	(4)
	All	Harmful Grades	Neutral Grades	Helpful Grades
Female	-0.028***	-0.083***	-0.000	-0.003***
	(0.002)	(0.005)	(0.001)	(0.001)
Observations	227,533	70,775	14,362	142,396
Conceal mean	0.107	0.329	0.002	0.007
Year × ΔGPA FEs	Yes	Yes	Yes	Yes

This table shows estimates from a linear probability model of whether a student chooses to conceal a letter grade in a course. Here we consider cases where students have the opportunity to change a letter grade to a grade of credit by concealing it in Column 1, a harmful letter grade (i.e., a grade that would pull down their GPA) to a grade of credit in Column 2, a neutral letter grade (i.e., a grade that would leave their GPA unaffected) to a grade of credit in Column 3 and a helpful letter grade (i.e., a grade that would pull up their GPA) to a grade of credit in Column 4. Female is a binary indicator that equals one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year (because we include these controls, we do not analyze students who are in their first semester, because they do not have a prior GPA). Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

a small concealment gap arises in these cases and that this gap is not always statistically significant (see Columns 3 and 4). For harmful grades, which are concealed around 52% of the time at BU and 33% of the time at the Midwestern Flagship, the concealment gap is substantial and statistically significant for harmful grades (see Column 2). At BU, female students are 8.9 percentage points (15%) less likely to conceal harmful grades than their male peers. At the Midwestern Flagship, female students are 8.3 percentage points (22%) less likely to conceal harmful grades than their male peers.

To conclude this section, we present two sets of robustness analyses. First, as also discussed in Section 5, the concealment gap persists across the distribution of letter grades (Appendix Table A.3); does not depend on the gender composition of a class (Appendix

Table A.4); does not vary systematically with course size, demand for the course, or course difficulty (Appendix Tables A.5 and A.6); is present for both relatively high-performing and low-performing students (Appendix Table A.7); persists for students across school years (Appendix Table A.8); is present across categories of majors (Appendix Table A.9); does not vary by whether the class part of a student's major (Appendix Table A.10); and is broadly observed across minority status, income levels, and whether a student is the first generation in their family to attend college (Appendix Table A.11).

Table 2: Student-level analysis: gender gap in concealing harmful grades

Panel A: Boston University			
	(1)	(2)	(3)
Dep. variable:	Ever concealed	# concealed	# concealed | >0
Female	-0.114***	-0.272***	-0.146***
	(0.010)	(0.021)	(0.023)
Constant	0.616***	1.075***	1.746***
	(0.007)	(0.016)	(0.017)
Observations	10,576	10,576	5,838
Conceal mean	0.552	0.922	1.671
Panel B: Midwestern Flagship			
	(1)	(2)	(3)
Dep. variable:	Ever concealed	# concealed	# concealed | >0
Female	-0.113***	-0.249***	-0.186***
	(0.005)	(0.010)	(0.014)
Constant	0.449***	0.741***	1.649***
	(0.004)	(0.007)	(0.010)
Observations	37,574	37,574	14,823
Conceal mean	0.395	0.620	1.572

This table shows estimates at the level of a student rather than a letter grade. Column 1 reports estimates from a linear probability model of whether a student ever chooses to conceal a harmful letter grade in a course. Column 2 reports estimates from an OLS regression where the outcome variable equals the number of harmful letter grades a student concealed. Column 3 reports estimates from an OLS regression where the outcome variable equals the number of harmful letter grades a student concealed, conditional on concealing at least one. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Robust standard errors are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Second, the concealment gap is robust to considering alternative measures that exploit student-level data rather than student-course-level data. When we look at whether students conceal *any* harmful grades in Table 2, Column 1 reveals a significant concealment gap: female students are 11 percentage points less likely than their male peers to conceal any

harmful grades at both universities. When considering the total number of harmful grades concealed, Column 2 also reveals a significant concealment gap: female students conceal 0.27 fewer grades at BU and 0.26 fewer grades at the Midwestern Flagship. Finally, Column 3 shows that, even conditional on concealing at least one grade, female students conceal 0.15 fewer grades at BU and 0.19 fewer grades at the Midwestern Flagship.

4 The Impact of the Concealment Gap on GPA

At both schools, students use the policies to conceal harmful grades that fall below their GPA. These policies thus make it possible for students to end up with substantially higher GPAs than they would have earned if they were forced to keep letter grades for all of their classes. However, students also frequently choose *not* to conceal harmful grades and sometimes, albeit very rarely, conceal helpful grades that would improve their GPA. In this section, we investigate the overall effect of these policies on students' GPAs and whether, given the gender concealment gap, there are significant gender differences in the impact of these policies on students' GPAs.

Table 3: Gender gap in the GPA impact of concealment

	Dep. variable: GPA impact of concealment	
	(1)	(2)
	BU	Midwestern Flagship
Female	-0.068***	-0.074***
	(0.005)	(0.004)
Constant	0.227***	0.228***
	(0.004)	(0.003)
Observations	15,690	61,194

This table shows estimates from an OLS regression of the GPA impact of concealment. The GPA impact of concealment reflects the difference between the actual GPA calculated for the term compared with the GPA students would have received if the policy was not in place and all grades were revealed. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Standard errors are reported in parentheses: robust in column (1) and clustered at the individual level in column (2). ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table 3 reports the GPA increase that students receive due to being able to conceal their grades over one semester at BU, shown in Column 1, and across two semesters at the Midwestern Flagship, shown in Column 2. The dependent variable is the GPA impact from the policy calculated by comparing the GPA that students actually received to the GPA they would have received had they not been able to conceal any of their letter grades. The

constant shows the increase in GPA earned by men (i.e., 0.227 points at BU and 0.228 points at the Midwestern Flagship). The coefficient on *Female* reveals a significant shift in GPA distributions in favor of men: female students gained roughly 0.07 fewer GPA points than men due to the policy. This shift in GPA distributions is substantial; it eliminates over half of the typical GPA advantage held by female students in a given semester (see also Appendix Figure B.2 for a graphical representation of this shift in GPA distributions).

In summary, the concealment gap has a significant impact on the GPAs of men relative to the GPAs of women. In addition to being an important academic outcome on its own, students' GPAs may also affect other important educational and labor market outcomes such as class rank, eligibility for select majors or extracurricular activities, and how employers consider students for internships or job opportunities sort their applications.[16]

5 Robustness of the Concealment Gap

Our data allows us to investigate the concealment gap across a variety of grade, course, and student characteristics. In this section, we show that the concealment gap is robust across all of these dimensions.

5.1 Does the concealment gap persist across various GPA benefits from concealment?

To maximize their GPAs, students need to recognize the benefits of concealing harmful grades. Both men and women appear quite sophisticated in the decision to conceal grades. As previously shown in Figure 1, both male and female students rarely conceal helpful grades and are more likely to conceal grades that are more harmful to their GPAs.[17] In addition, Figure 1 shows that the concealment gap persists across various ranges of GPA benefits from concealment.

5.2 Does the concealment gap persist across various letter grades?

The gender concealment gap arises across various specific letter grades. Appendix Table A.3 presents our results by letter grade that could be changed to a grade of "Credit" if

[16]Indeed, additional data from an incentivized experiment that we further discuss later reveals that employers recruiting from an elite undergraduate program are significantly more likely to select job candidates for interviews when they have higher GPAs. In particular, as shown in Appendix Figure D.1 Panel A, job candidates with a one-point higher GPA are 18 percentage points more likely to receive an interview.

[17]While the point estimates are tiny, men are statistically significantly more likely than women to conceal helpful grades, which is the opposite of what we would expect if men were engaging in more strategic efforts to maximize their GPA.

concealed.[18] While the concealment rates for a grade of A- are very low (the concealment rate is 7% at BU and 4% at the Midwestern Flagship), the concealment rates increase as the grades decrease; the concealment rate is over 70% for grades of C+, C, and C- at both universities. The gender gap in concealment is at least directionally present at A- and C-, and we observe a large and robust gender gap in concealment that is statistically significant—at both universities—for grades of B+, B, B-, C+, and C.

5.3 Does the concealment gap persist among both low-performing and high-performing students?

Appendix Table A.7 splits the data for each university based on a student's GPA at the start of the semester. The concealment gap arises among students across the GPA distribution, including among students with high GPAs.

5.4 Does the concealment gap persist across courses that vary in size, difficulty, and gender balance?

Motivated by prior literature on the relevance of group composition (Eckel and Grossman, 2001; Solnick, 2001; Gneezy, Niederle and Rustichini, 2003; Bowles, Babcock and Lai, 2007; Sutter et al., 2009; Hernandez-Arenaz and Iriberri, 2023) and competitiveness (for a review, see Niederle and Vesterlund (2011)) one might be interested in how the concealment gap various across various courses. Appendix Tables A.4–A.6 and A.10 show that the concealment gap persists across courses that vary in their gender composition, number of students enrolled, demand for enrollment, difficulty level (as captured by the average and standard deviation of grades in a course), and whether the course falls within or outside a student's major. This persistence of the concealment gap across different types of courses suggests that any set of explanations for the concealment gap must account its presence across a wide variety of courses with different features.

5.5 Does the concealment gap persist across majors?

Motivated by prior work on how gender differences often arise in male-stereotyped domains (Coffman, 2014; Bordalo et al., 2016; Coffman, Collis and Kulkarni, 2019; Exley and Kessler, 2022), it is natural to consider whether the concealment gap varies depending on whether a student is a STEM major or a major in another field. In Appendix Table A.9, we show that the concealment gap is observed for STEM majors (Column 5), but also is present (and

[18] Recall that this excludes letter grades of F at BU and letter grades of D or F at the Midwestern Flagship, as detailed in Footnote 11.

statistically significant) across the other five major categories at the Midwestern Flagship; for BU, we observe results that are directionally consistent across all major categories (although the coefficient on *Female* is statistically significant for only three of six major groupings).

5.6 Does the concealment gap persist absent concerns—including confidence and risk aversion—about specific grades?

Suppose students had been required to make their concealment decisions *before* learning their grades. Given the well-documented gender gap in confidence (Lundeberg, Fox and Punćcohaŕ, 1994; Niederle and Vesterlund, 2007; Niederle, 2016; Bordalo et al., 2019) and self-evaluations (Exley and Kessler, 2022), one might then expect that women would choose to conceal grades *more* often because, relative to men, they underestimate the grade they would earn in a class. One might also expect that women would conceal grades more often because women may be more risk averse about the potential outcome of a low grade (Eckel and Grossman, 2008; Croson and Gneezy, 2009; Niederle, 2016). Indeed, when Trost and Wooten (2023) look at whether students opt into a credit/no-credit grading scheme *before* the final exam (and hence before their class grade is known with certainty), they find that women are more likely to choose the credit/no-credit grading option, which could—as those authors note—be reflective of gender differences in confidence or risk aversion concerns about what grades they will receive.

In our setting, however, students make their concealment decision *after* learning their grades and hence after any uncertainty about the grade they have earned in each course has been resolved. This fact suggests that gender differences in confidence or risk aversion about specific grades cannot drive the concealment gap in our setting.

It is still possible, however, that our results could in part be driven by gender differences in beliefs and concerns about *future* grades. For example, women may believe a grade that is harmful this semester could be above their final collegiate GPA (perhaps because of a gender gap in confidence in future course performance) or because women are risk averse about earning particularly low GPAs and so are willing to reveal grades below their current GPA but high enough to mitigate against a theoretically possible low final collegiate GPA.

To investigate if the concealment gap persists even when uncertainty about future grades is mostly or entirely resolved, we examine whether the concealment gap persists among students in their senior year. Appendix Table A.8 reveals that it does: at both universities, the concealment gap is present and statistically significant when only considering senior students. In addition, recall from Figure 1 that the concealment gap arises for grades well below a student's current GPA, including grades that are almost certainly below a student's final collegiate GPA (i.e., settings in which students should always conceal their grades if

they are primarily driven by concerns about their final collegiate GPA). Taken together, these results suggest that that uncertainty about future grades is unlikely to be driving the concealment gap we observe.

5.7 Does the concealment gap persist when there is less scope for gender differences that could arise due to awareness or attention?

Students had to opt into concealing grades, so the gender concealment gap could theoretically reflect lower awareness or less attentiveness toward the policy among female students.[19] But features of the setting and our results argue against this being a dominant explanation for the observed gender concealment gap.

In terms of the setting, the grading policies were prominently announced and advertised at each institution. In terms of results, recall that the concealment gap arises even when we examine the number of grades students conceal conditional on concealing at least one grade, which means that it is observed among students who are clearly aware of the policy and responding to it (see Table 2).[20]

In addition, Appendix Table A.12 shows that the concealment gap arises when we separately consider different policy implementation periods that varied in terms of how much time students had to become aware and attend to the policy (recall details in Section 2.1). Indeed, these results show that the concealment gap persists when the policy is announced before a semester begins and unexpectedly late during a semester, so that the concealment gap persists when there is more and less room for any potential impact it has on effort during class or on class selection. At the Midwestern Flagship, Appendix Figure B.1 further shows that we observe no significant change in the gender gap in course performance going from Fall 2020 (the year before the policy) to Fall 2021 (the year of the policies). This suggests that men and women are not differentially changing their effort as a result of the policies. (We lack data for a corresponding analysis at BU.)

[19] We are not aware of prior evidence of women being less attentive than men. In contrast, there exists evidence that men are more likely than women to procrastinate (Cortés et al., forthcoming). To the extent that procrastination would lead individuals to miss the opportunity to conceal grades, this phenomenon would push against observing the gender concealment gap we document.

[20] The concealment gap also persists when we only consider data from the Spring 2021 semester at the Midwestern Flagship, restricting to the set of students who previously concealed at least one grade in the Fall 2020 semester; see Column 3 of Appendix Table A.12.

5.8 Does the concealment gap persist even when students do not need to reveal certain grades (e.g., for graduate school admissions)?

There are well-documented gender differences in selection into industries and even selection into graduate school. Given that women are more likely to attend graduate school, a gender concealment gap could arise if women are compelled to leave certain grades visible on their transcripts for graduate school admissions committees.[21] There are several patterns in the data which suggest that the concealment gap persists even in cases in which graduate school considerations are unlikely to be relevant.

First, as previously discussed in Section 5.3 and as shown in Table A.7, the concealment gap arises across the distribution of GPAs, including students who have a relatively low GPA in the 2.51–3.00 range. While there is no fixed cutoff for graduate school applications, common advice is that 3.00 or higher is strongly advisable for graduate school applicants to all but the lowest-ranked schools.

Second, as previously discussed in Section 5.1 and as shown in Appendix Table A.3, the concealment gap persists even when considering courses with low letter grades of C or $C-$. It is unlikely that (female) students are choosing to reveal these low grades on their transcript in an effort to get into their graduate program of interest.

Third, the concealment gap arises and is of comparable magnitude both when considering courses within a student's major *and* outside of a student's major (see Appendix Table A.10), the latter of which are likely less relevant for graduate school.

Fourth, the concealment gap arises for introductory-level courses (see Appendix Table A.13), and even when excluding more advanced courses that might be more likely to be used as screening devices for some competitive graduate programs.

Thus, the concealment gap persists even when it is unlikely that students need to reveal certain grades in particular courses (e.g., for graduate school admissions).

6 Potential Drivers of the Concealment Gap

The prior section highlighting the robustness of the concealment gap provides evidence against a number of potential drivers for the gap. In this section, we highlight three potential drivers of the concealment gap that we believe are plausible given our data.

[21] For national trends on graduate school admissions by gender, see Appendix Table B.3 at https://cgsnet.org/wp-content/uploads/2023/10/2022-Graduate-Enrollment-and-Degrees-Tables-and-Appendices.pdf, last accessed July 9, 2025.

6.1 Could gender differences in a preference for transparency drive the concealment gap?

While the literature on gender differences in lying and deception is mixed, a recent meta-analysis provides support for the view that men have a higher propensity to lie than women do (Capraro, 2018).[22]

While lying is conceptually distinct from concealing grades in our context—in our settings, given how the policies were advertised and implemented, choosing to conceal grades is neither a lie by omission nor by commission—one could still wonder whether gender differences in preferences for transparency account for the gender concealment gap.

Our results are supportive of women generally being more transparent about academic performance than men, particularly given the persistence of the concealment gap across different types of students, different courses, and two different universities. In addition, women appear more likely than men to be transparent about their neutral and helpful grades as well, although the rates of concealment are very low in these cases and hence not the focus of our analyses.

6.2 Could gender differences in a preference for the status quo drive the concealment gap?

In academic settings, grades are typically revealed. Even when the policies we study are introduced, most grades are revealed and students need to take active steps to conceal information on their performance. If women were more reluctant than men to deviate from the status quo or to deviate from what is typically done, this might make them less likely than men to conceal information in our setting.

Because we only explore an environment in which the default is revealing information, and revealing information is the typical course of action, a gender difference in preferences for the status quo will be indistinguishable from a preference for transparency as discussed above in Section 6.1. We also note that, since prior work often has focused on active decisions to reveal performance—rather than active decisions to conceal performance—we hope future work explores more settings that examine active concealment decisions as well.

[22] See also, e.g., Dreber and Johannesson (2008), Erat and Gneezy (2012), and Abeler, Falk and Kosse (2021).

6.3 Could beliefs about what others will infer when information is concealed drive the concealment gap?

While gender differences in a preference for transparency, for adhering to the status-quo, or for doing what is typically done could drive the gender concealment gap, additional evidence suggests that these explanations are unlikely to fully explain our findings.

As shown in Appendix Table A.3, we do not see a statically significant concealment gap for relatively poor letter grades such as C–. In addition, we see a reverse concealment gap for grades of "D" at BU—grades for which a pure preference for transparency argument should still apply.[23]

That women are more likely than men to conceal a grade of D (i.e., the lowest possible passing grade) suggests that women are relatively willing to conceal poor performance when there is no worse inference an observer could make about performance than observing the grade that was actually earned.

Indeed, when assessing the potential costs and benefits of concealing a harmful grade, students may consider what others—such as employers, graduate school admission committees, faculty considering them for research assistant positions, or faculty writing reference letters for internships or jobs—will infer about their performance in a course when they have concealed the letter grade they received. As a result, a concealment gap might arise due to a gender difference in beliefs about how concealment will be viewed by others.

Because this potential driver of the gender concealment gap relies on students' beliefs about others' inferences, we are unable to assess whether it is a plausible driver with our existing data. In order to speak to this potential explanation, we ran additional studies to gather information on such beliefs.

To explore beliefs about what others will infer from concealed grades, we recruited 407 undergraduate students from the Midwestern Flagship to complete our *student belief study*, and an additional 399 adults from Prolific to complete our *general public belief study* (which closely follows the *student belief study* and is detailed in Appendix C.4). To incentivize one of set of beliefs that these participants provide, we also rely on additional data from our *employer study*, an experiment with actual employers.

In the *student belief study*, all 407 recruited participants are rising juniors and seniors at

[23] In addition, at BU we also do not see a gender gap in the rates at which men and women change an "F" to a grade of "No Credit" (which may be relevant to the extent that an "F" is more transparent than a grade of "No Credit").

the Midwestern Flagship.[24] In addition to receiving a $10 Amazon gift card for completing the study, students had a chance to receive a $20 bonus payment.

In the study, student subjects were asked two questions about the inferences that employers and graduate school admissions committees would make when reviewing applications from candidates who concealed a grade. They were also asked three questions about the inferences employers would make when reviewing resumes that did not provide GPA information.

We asked the inference questions about concealing letter grades given that we documented robust concealment gaps of letter grades in the BU and Midwestern Flagship data. We asked the inference questions about concealing GPA from a resume for two reasons. One key reason was because we were able to incentivize some of these questions by leveraging data from our *employer study*. An additional reason was to generate insights into what people believe will be inferred when academic performance is concealed in other settings, such as in job applications where GPA is often the main indicator of academic performance. We see exploring such beliefs as a fruitful area for future work.

Beginning with the beliefs about inferences related to concealed letter grades, students in the *student belief study* are asked to consider a man and a woman from the same large university in the United States who have exactly the same GPA, the same transcript, and the same resume. We then elicit beliefs about inferences when both students received a grade of "credit" or "pass" in the same relevant course. In one question, we ask what employers will infer when both students apply to exactly the same job by selecting from three options: "employers probably expect both to have earned similar grades," "employers probably expect the woman to have earned a worse grade," or "employers probably expect the man to have earned a worse grade" (see Appendix Figure C.6). In a second question, we ask what graduate school admission committees will infer when both students apply to exactly the same graduate school program by selecting from three similar options (see Appendix Figure C.7).

In addition, we ask one question of the same form but about a concealed GPA. Subjects are asked to consider a man and a woman from the same large university in the United States who have exactly the same GPA, the same transcript, and the same resume. We ask subjects what employers will infer when both students apply to exactly the same job and provide resumes that do not include GPA information by selecting from three options: "employers probably expect both to have earned similar GPAs," "employers probably expect

[24] We invited 2,500 randomly selected juniors and seniors to participate in July 2024. We closed the survey when we successfully recruited 400 students (our pre-determined cutoff). We allowed in-progress surveys to be completed and ended up with a sample of 407 students. Students who completed the survey are similar in observable characteristics to the broader student population. Details and screenshots of the experiment are presented in Appendix Section C.

the woman to have earned a worse GPA," or "employers probably expect the man to have earned a worse GPA" (see Appendix Figure C.5).

Panels A and B of Figure 2 reveal expected inference gaps related to concealed letter grades: students believe employers and graduate school admissions committees will hold more negative inferences about a woman than a man with a grade of credit. In addition, Panel C of Figure 2 shows that an expected inference gap also arises when we ask about students concealing their GPAs from resumes. While some students report that employers and graduate admissions committees will have equal inferences about men and women (33%–47%), the remaining students are 1.9–2.4 times more likely to believe that men will be viewed more favorably than women (36–47%) than the opposite (17–21%). Appendix Figures C.11 and C.12 reveal that the beliefs about more negative inferences for women are more pronounced for female students; male students also directionally expect more negative inferences for women, but the gap is not statistically significant in two of three cases for male students.

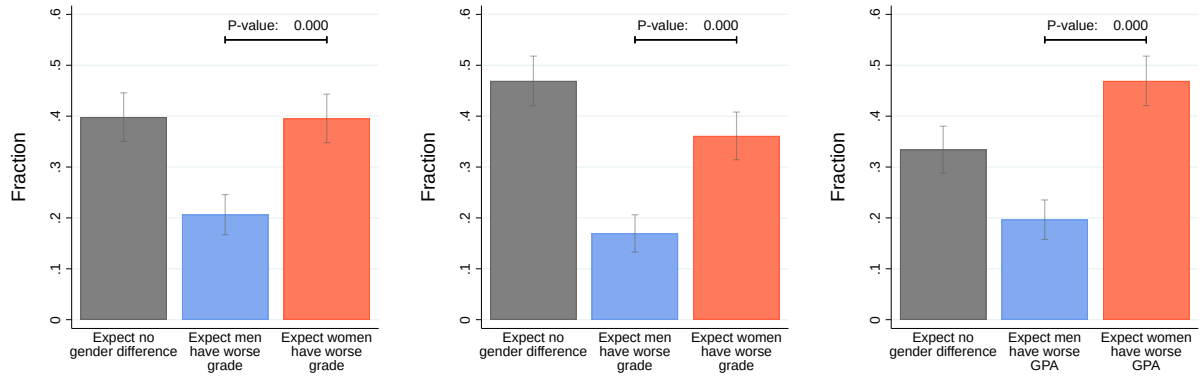
While these three belief questions were asked without incentives, results are very similar when students are asked *incentivized* questions about the inferences that employers will draw when performance information is concealed. Beliefs about what employers infer when GPAs are omitted from resumes are incentivized using evaluations of job candidates in our *employer study* which involves actual employers engaged in on-campus recruiting at the University of Pennsylvania. Students learn about the employers' evaluation decisions in detail (see Appendix C for more details) and then are told that their chance of receiving a $20 bonus is higher if they answer each question correctly.

As shown in Appendix Figure C.10, students believe that employers will treat male students with a concealed GPA as if they had a GPA of 3.34 but instead treat female students only as if they had a GPA of 3.22 ($p < 0.01$). This expected inference gap also arises when comparing the beliefs men and women hold about their own gender: Men believe that employers will treat male students as if they had a GPA of 3.33 while women believe that employers will treat female students as if they have a GPA of 3.21.

Examining data from our *employer study*—which leverages an Incentivized Resume Rating experiment (Kessler, Low and Sullivan, 2019)—we find that employer inferences are indeed consistent with students' beliefs.[25] In the absence of GPA information, men are substantially more likely to be selected for an interview than women are: employers treat men missing GPA information as equivalent to candidates with substantially higher GPAs than they treat women missing GPA information (see Appendix D for more details).

[25]This approach builds from classic audit studies that have also provided important insight into how employers respond to resumes, e.g., see Bertrand and Mullainathan (2004) and Deming et al. (2016).

Figure 2: Expectations of relative inferences by gender, for employers and admissions committees



(a) Expected grade, employers

(b) Expected grade, admissions committees

(c) Expected GPA, employers

This figure plots the percent of students who expected the same inference to be made about men and women, worse inferences for men, and worse inferences for women, respectively, for three different scenarios: Panel (a): employers assessing a man and woman applying to the same job who received a grade of "credit" or "pass" in the same relevant course (and therefore have missing grade information); Panel (b): graduate school admissions committees assessing a man and woman applying to the same graduate program who received a grade of "credit" or "pass" in the same relevant course (and therefore have missing grade information); and Panel (c): employers assessing a man and woman applying to the same job who have missing GPA information on their resumes; gray bars reflect 95% confidence intervals. The sample includes 407 predictions for each scenario.

In addition, we also document the expected inference gaps among 399 online participants in our *general public study* (as detailed in Appendix C.4). We thus add to prior work that documents similar findings with undergraduate student populations and online participant populations, particularly when comparative statistics are involved (Snowberg and Yariv, 2021); both students in the *student belief study* and adults in the *general public belief study* display the expected inference gaps when considering beliefs about concealed grades and concealed GPAs.

These expected inference gaps lend credence to the idea that the gender concealment gap may be driven by beliefs about what inferences others will draw when information is concealed. Women may expect employers and admissions committees to make worse inferences about them if they conceal their performance, increasing the costs of concealing grades for women relative to men.

This result suggests that the gender concealment gap might differ in situations in which women expect that others will give them the "benefit of the doubt" about their performance if information is concealed. We hope future work explores whether this is in fact the case.

7 Conclusion

In this paper, we document a robust gender gap in concealment of grades at two large, selective universities. Women are substantially less likely to conceal grades below their GPA than men are, a pattern that is observed across a range of grades, student traits, and course-level characteristics. In particular, we observe a persistent concealment gap across: the distribution of letter grades; male- and female-dominated courses; all quintiles of course sizes; high- and low-demand courses; courses with above- and below-median grades; courses with above- and below-median variance in grades; all course levels; high- and low-performing students; all school years; categories of majors; courses within and outside a student's major; student minority status; students from high- and low-income backgrounds; and students who are the first generation in their family to attend college or not.

This gap has consequences for gender equity, since it results in lower GPA gains for women relative to men. There is also an expectation, held by both students and the general public, that women are more likely to have negative inferences made about their performance when they conceal their grades, suggesting a rationale for why women might be more hesitant than men about concealing harmful grades.

An important question for future work relates to *when* gender gaps in concealment are likely to arise. Our results make clear the robustness of the concealment gap, particularly when considering performance information that corresponds to a *known failure* because we define harmful grades at the individual-level (i.e., as a grade that falls below a student's GPA). Our focus on *known failures* is notably different from prior literature that typically focuses on gender differences when the performance outcome is *unknown* to the decision maker. This prior literature often finds that women are more reluctant than men to reveal information on their unknown performance, either directly or indirectly by engaging in some interaction or task. Since we instead find that women are less reluctant than men to reveal known failures, a natural question for future work is to explore how gender differences in performance transparency may depend on whether the performance involves a known success, a known failure, or an ambiguous or unknown outcome. A closely related question is whether there are systematic gender differences in how men and women view outcomes as either a failure or success in educational and workplace settings. In addition, since prior work has often focused on active decisions to reveal or promote one's performance, we view more work that instead focuses on active decisions to conceal performance as particularly important to our understanding of what information is available on the performance of men and women.[26]

[26]Indeed, examining how gender differences are influenced by the default option is an important avenue for the broader gender literature. For instance, He, Kang and Lacetera (2021) show that changing the default option substantially affects gender differences in willingness to enter a competition.

Future work may also investigate how concealment gaps are influenced by the degree to which people expect discrimination and the type of discrimination that is expected. Recent work highlights the importance of (often inaccurate) belief-based sources of discrimination (see, e.g., Coffman, Exley and Niederle (2021); see also Bohren et al. (2025) for an excellent discussion of the literature and related evidence). Our belief studies highlight that people expect worse inferences about the performance of women than men when performance information is concealed or not provided. Educational choices, negotiation choices, and career choices may all be influenced by the extent to which men and women expect to face (accurate or inaccurate) belief-based discrimination as well as taste-based discrimination.

Finally, given the paucity of work on gender differences relating to the disclosure of known failures, future work may examine the presence and implications of concealment gaps in the choice to report clearly negative information in other settings, such as in the workplace.[27] Building up this literature should further make clear the need to carefully consider equity implications when deciding whether to allow individuals to conceal information.[28] Indeed, just because a policy introduces flexibility does not mean it is good for equity. The option value of a choice to reveal information may be more valuable for those who can benefit from positive inference in the absence of information, such as groups that are viewed more favorably when relatively little information is available.

[27]Future work may also investigate if the gender concealment gap depends on the gender stereotypes, norms and beliefs (Bordalo et al., 2016, 2019; Bursztyn, González and Yanagizawa-Drott, 2020; Bursztyn, Cappelen and Tungodden, 2023).

[28]Consistent with the findings in Exley and Nielsen (2024), one could be concerned that—even if others (e.g., employers and graduate school admissions committees) became aware of men and women taking advantage of these policies differentially—it may be difficult for them to account for this fact in their decisions. That said, simply making clear the unequal impact on GPA gains for men and women that result from this grade-optional policy may affect institutions' preferences, as beliefs about biases in test scores affect preferences to use test score information (Liang and Xu, 2024).

References

Abeler, Johannes, Armin Falk, and Fabian Kosse. 2021. "Malleability of Preferences for Honesty." *IZA Discussion Paper No. 14304.*

Adda, Jerome, Christian Dustmann, and Katrien Stevens. 2017. "The Career Costs of Children." *Journal of Political Economy*, 125(2): 293–337.

Agan, Amanda, and Sonja Starr. 2017. "The effect of criminal records on access to employment." *The American Economic Review*, 107(5): 560–64.

Agan, Amanda, and Sonja Starr. 2018. "Ban the box, criminal records, and racial discrimination: A field experiment." *The Quarterly Journal of Economics*, 133(1): 191–235.

Agan, Amanda, Bo Cowgill, and Laura Gee. 2023*a*. "Salary History and Employer Demand: Evidence from a Two-Sided Audit." *Working Paper.*

Agan, Amanda, Bo Cowgill, and Laura Gee. 2023*b*. "The Tradeoffs of Deception: Measuring Discrimination When Subjects Are Told They Are in an Experiment." *Working Paper.*

Agan, Amanda, Bo Cowgill, and Laura Katherine Gee. 2020. "Do Workers Comply with Salary History Bans: A Survey of Voluntary Disclosure, Adverse Selection, and Unraveling." *American Economic Association Papers and Proceedings*, 110: 215–219.

Aksoy, Billur, Ian Chadd, and Boon Han Koh. 2023. "Sexual identity, gender, and anticipated discrimination in prosocial behavior." *European Economic Review*, 154(104427).

Alston, Mackenzie. 2019. "The (Perceived) Cost of Being Female: An Experimental Investigation of Strategic Responses to Discrimination." *Working Paper.*

Altonji, Joseph G, and Charles R Pierret. 2001. "Employer learning and statistical discrimination." *The Quarterly Journal of Economics*, 116(1): 313–350.

Bartik, Alexander, and Scott Nelson. 2016. "Deleting a signal: Evidence from pre-employment credit checks."

Belasco, Andrew, Kelly Rosinger, and James Hearn. 2015. "The Test-Optional Movement at America's Selective Liberal Arts Colleges: A Boon for Equity or Something Else?" *Educational Evaluation and Policy Analysis*, 7(2): 206–223.

Bennett, Christopher. 2022. "Untested admissions: Examining changes in application behaviors and student demographics under test-optional policies." *American Educational Research Journal*, 59: 180–216.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*, 94(4): 991–1013.

Biasi, Barbara, and Heather Sarsons. 2022. "Flexible Pay, Bargaining, and The Gender Gap." *The Quarterly Journal of Economics*.

Blair, Peter Q, and Bobby W Chung. 2021. "A Model of Occupational Licensing and Statistical Discrimination." Vol. 111, 201–05.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends. and Explanations." *Journal of Economic Literature*, 55(3).

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2025. "Inaccurate Statistical Discrimination." *Review of Economics and Statistics*, 1–16.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2019. "Beliefs about Gender." *American Economic Review*, 109(3): 739–73.

Borghesan, Emilio. 2022. "The Heterogeneous Effects of Changing SAT Requirements in Admissions: An Equilibrium Evaluation." *Working Paper*.

Bowles, Hannah Riley, and Kathleen L. McGinn. 2008. "Gender in job negotiations: a two-level game." *Negotiation Journal*, 24(4): 393–410.

Bowles, Hannah Riley, Linda Babcock, and Lei Lai. 2007. "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask." *Organizational Behavior and human decision Processes*, 103(1): 84–103.

Burbano, Vanessa, Nicolas Padilla, and Stephan Meier. Forthcoming. "Gender Differences in Preferences for Meaning at Work." *American Economic Journal: Economic Policy*.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott. 2020. "Misperceived social norms: Women working outside the home in Saudi Arabia." *American Economic Review*, 110(10): 2997–3029.

Bursztyn, Leonardo, Alexander W Cappelen, and Bertil Tungodden. 2023. "How are gender norms perceived?" *Working Paper*.

Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais. 2017. "'Acting Wife': Marriage Market Incentives and Labor Market Investments." *The American Economic Review*, 107(11): 3288–3319.

Capraro, Valerio. 2018. "Gender differences in lying in sender-receiver games: A meta-analysis." *Judgment and Decision Making*, 13(4): 345–355.

Cech, Erin. 2013. "The Self-Expressive Edge of Occupational Sex Segregation." *American Journal of Sociology*, 119(3): 747–789.

Chan, Alex. 2022. "Discrimination Against Doctors: A Field Experiment." *Working Paper*.

Charles, Maria, and Karen Bradley. 2009. "Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries." *The American Journal of Sociology*, 14: 924–976.

Chetty, Raj, David Deming, and John Friedman. 2023. "Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges." *Working paper*.

Coffman, Katherine Baldiga. 2014. "Evidence on Self-Stereotyping and the Contribution of Ideas." *The Quarterly Journal of Economics*, 129(4): 1625–1660.

Coffman, Katherine B., Christine L. Exley, and Muriel Niederle. 2021. "The Role of Beliefs in Driving Gender Discrimination." *Management Science*, 67(6): 3551–3569.

Coffman, Katherine B., Manuela R. Collis, and Leena Kulkarni. 2023. "Whether to Apply?" *Management Science*, 70(7).

Coffman, Katherine, Manuela Collis, and Leena Kulkarni. 2019. "Stereotypes and Belief Updating." *Working Paper*.

Conlin, Michael, Stacy Dickert-Conlin, and Gabrielle Chapman. 2023. "Voluntary disclosure and the strategic behavior of colleges." *Journal of Economic Behavior & Organization*, 96: 48–64.

Correll, Shelley. 2001. "Gender and the Career Choice Process: The Role of Biased Self Assessments." *American Journal of Sociology*, 106(6).

Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar. forthcoming. "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab." *Quarterly Journal of Economics*.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature*, 47(2): 448–474.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2007. "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness." *Economic Theory*, 33: 67–80.

Deming, David J, Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F Katz. 2016. "The value of postsecondary credentials in the labor market: An experimental study." *American Economic Review*, 106(3): 778–806.

Dessein, Wouter, Alexander P. Frankel, and Navin Kartik. 2023. "Test-Optional Admissions." *Working Paper*.

Di Tella, Rafael, Ricardo Perez-Truglia, Andres Babino, and Mariano Sigman. 2015. "Conveniently Upset: Avoiding Altruism by Distorting Beliefs about Others' Altruism." *The American Economic Review*, 105(11): 3416–42.

Doleac, Jennifer L, and Benjamin Hansen. 2017. "Moving to Job Opportunities? The Effect of" Ban the Box" on the Composition of Cities." *The American Economic Review*, 107(5): 556–59.

Doleac, Jennifer L., and Benjamin Hansen. 2020. "The Unintended Consequences of Ban the Box." *Journal of Labor Economics*, 38(2): 321–374.

Dreber, Anna, and Magnus Johannesson. 2008. "Gender differences in deception." *Economics Letter*, 99(1).

Dustan, Andrew, Kristine Koutout, and Greg Leo. 2022. "Second-order beliefs and gender." *Journal of Economic Behavior & Organization*, 200: 752–781.

Eckel, Catherine C, and Philip J Grossman. 2001. "Chivalry and solidarity in ultimatum games." *Economic inquiry*, 39(2): 171–188.

Eckel, Catherine C., and Philip J. Grossman. 2008. "Men, Women and Risk Aversion: Experimental Evidence." In *Handbook of Experimental Economics Results*. 1061–1073.

Erat, Sanjiv, and Uri Gneezy. 2012. "White Lies." *Management Science*, 50(4): 723–733.

Exley, Christine L. 2016. "Excusing Selfishness in Charitable Giving: The Role of Risk." *Review of Economic Studies*, 83(2): 587–628.

Exley, Christine L., and Judd B. Kessler. 2022. "The Gender Gap in Self-Promotion." *The Quarterly Journal of Economics*, 137(3): 1345–1381.

Exley, Christine L., and Judd B. Kessler. Accepted. "Motivated Errors." *American Economic Review*.

Exley, Christine L., and Kirby Nielsen. 2024. "The Gender Gap in Confidence: Expected But Not Accounted For." *The American Economic Review*, 114(3): 851–885.

Fernandez, Roberto, and Colette Freidrich. 2011. "Gender Sorting at the Application Interface." *Industrial Relations*, 50(4): 591–609.

Gagnon, Nickolas, Kristof Bosmans, and Arno Riedl. 2022. "The Effect of Gender Discrimination on Labor Supply." *Available at SSRN 3519540*.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in competitive environments: Gender differences." *The quarterly journal of economics*, 118(3): 1049–1074.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *The American Economic Review*, 104(4): 1091–1119.

Grossman, Sanford J. 1981. "The informational role of warranties and private disclosure about product quality." *The Journal of Law and Economics*, 24(3): 461–483.

Haisley, Emily C., and Roberto A. Weber. 2010. "Self-serving interpretations of ambiguity in other-regarding behavior." *Games and Economic Behavior*, 68: 614–625.

Hancart, Nathan. 2023. "Designing the Optimal Menu of Tests." *Working Paper*.

He, Joyce, Sonia Kang, and Nicola Lacetera. 2021. "Leaning In or Not Leaning Out? Opt-Out Choice Framing Attenuates Gender Differences in the Decision to Compete." *Proceedings of the National Academy of Sciences*, 118(42).

Hernandez-Arenaz, Inigo, and Nagore Iriberri. 2019*a*. "A Review of Gender Differences in Negotiation." *Oxford Research Encyclopedia of Economics and Finance*.

Hernandez-Arenaz, Iñigo, and Nagore Iriberri. 2019*b*. "A review of gender differences in negotiation." *Oxford Research Encyclopedia of Economics and Finance*.

Hernandez-Arenaz, Iñigo, and Nagore Iriberri. 2023. "Gender differences in alternating-offer bargaining: An experimental study." *Experimental Economics*, 1–36.

Hirshman, Samuel D, and Alexander Willén. 2022. "Does Increasing Risk Widen Gender Gaps?" *NHH Dept. of Economics Discussion Paper*.

Hyde, Janet Shibley. 2014. "Gender Similarities and Differences." *Annual Review of Psychology*, 65: 373–398.

Kessler, Judd B, Corinne Low, and Colin D Sullivan. 2019. "Incentivized Resume Rating: Eliciting Employer Preferences without Deception." *The American Economic Review*, 109(11): 3713–44.

Kessler, Judd, Corinne Low, and Xiaoyue Shan. 2023. "Lowering the Playing Field: Discrimination through Sequential Spillover Effects." *Working Paper*.

Kleven, Henrik, Camille Landais, and Jakob Egholt Sogaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, 11(4): 181–209.

Kolb, Aaron, Marilyn Pease, Daniel W. Sacks, and Joshua Quick. 2023. "Blind Disclosure." *American Economic Journal: Microeconomics*, 15(2): 41–79.

Koutout, Kristine. 2022. "Gendered Beliefs and the Job Application Decision: Evidence from a Large-Scale Field and Lab Experiment." *Working Paper*.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin*, 108(3): 480–498.

Leibbrandt, Andreas, and John A. List. 2015. "Do women avoid salary negotiations? Evidence from a large-scale natural field experiment." *Management Science*, 61(9): 2016–2024.

Lepage, Louis-Pierre. 2024. "Experience-based discrimination." *American Economic Journal: Applied Economics*, 16(4): 288–321.

Lepage, Louis-Pierre, Xiaomeng Li, and Basit Zafar. 2025. "Anticipated Discrimination and Major Choice." *Working Paper*.

Liang, Yucheng, and Wenzhuo Xu. 2024. "Why Exclude Test Scores from Admission Criteria?" *Working paper*.

Low, Corinne. 2022. "The Human Capital - "Reproductive Capital" Tradeoff in Marriage Market Matching." *Manuscript, Wharton School of Business*.

Lundberg, Shelly J, and Richard Startz. 1983. "Private discrimination and social intervention in competitive labor market." *The American Economic Review*, 73(3): 340–347.

Lundeberg, Mary A, Paul W Fox, and Judith Punćcohaŕ. 1994. "Highly confident but wrong: Gender differences and similarities in confidence judgments." *Journal of Educational Psychology*, 86(1).

Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics*, 380–391.

Morrison, Jean. 2020. "University Grading Policies for Spring Semester 2020."

Murciano-Goroff, Raviv. 2022. "Missing women in tech: The labor market for highly skilled software engineers." *Management Science*, 68(5): 3262–3281.

Niederle, Muriel. 2016. "Gender." In *Handbook of Experimental Economics*. Vol. 2, , ed. John Kagel and Alvin E. Roth, 481–553. Princeton University Press.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women shy away from competition? Do men compete too much?" *The Quarterly Journal of Economics*, 122(3): 1067–1101.

Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and Competition." *Annual Review of Economics*, 3: 601–630.

Petersen, Trond, Andrew Penner, and Geir Hogsnes. 2011. "The Male Marital Wage Premium: Sorting Versus Differential Pay." *ILR Review*, 64(2): 283–304.

Reuben, Ernesto, Paola Sapienza, and Luigi Zingales. 2014. "How stereotypes impair women's careers in science." *Proceedings of the National Academy of Sciences*, 111(12): 4403–4408.

Roussille, Nina. 2024. "The central role of the ask gap in gender pay inequality." *Quarterly Journal of Economics*, 139(3): 1557–1610.

Ruebeck, Hannah. 2023. "Perceived discrimination at work."

Saboe, Matt, and Sabrina Terrizzi. 2019. "SAT optional policies: Do they influence graduate quality, selectivity or diversity?" *Economics Letters*, 174: 13–17.

Snowberg, Erik, and Leeat Yariv. 2021. "Testing the waters: Behavior across participant pools." *American Economic Review*, 111(2): 687–719.

Snyder, Melvin L, Robert E Kleck, Angelo Strenta, and Steven J Mentzer. 1979. "Avoidance of the handicapped: an attributional ambiguity analysis." *Journal of personality and social psychology*, 37(12): 2297–2306.

Solnick, Sara J. 2001. "Gender differences in the ultimatum game." *Economic Inquiry*, 39(2): 189–200.

Sutter, Matthias, Ronald Bosman, Martin G Kocher, and Frans van Winden. 2009. "Gender pairing and bargaining—Beware the same sex!" *Experimental Economics*, 12: 318–331.

Trost, Steve, and Jadrian Wooten. 2023. "Gender, risk aversion, and the "COVID" grading option in a principles of economics course." *Applied Economics Letters*.

Ugalde, Paola. 2022. "Gender, Grade Sensitivity, and Major Choice."

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender." *The Quarterly Journal of Economics*, 133(1): 457–507.

Wozniak, Abigail. 2015. "Discrimination and the effects of drug testing on black employment." *Review of Economics and Statistics*, 97(3): 548–566.

Zafar, Basit. 2013. "College Major Choice and the Gender Gap." *Journal of Human Resources*, 48(3): 545–595.

ONLINE APPENDIX

A Additional tables

Table A.1: Summary statistics

	Boston University		Midwestern Flagship	
	Female	Male	Female	Male
Eligible students	9,148	6,542	16,022	15,888
First-years	2,099	1,468	1,399	1,289
Sophomores	2,582	1,921	3,378	3,142
Juniors	2,091	1,573	4,003	3,837
Seniors	2,376	1,580	7,242	7,620
Student Performance				
Cumulative GPA	3.33	3.20	3.56	3.45
Credits earned	14.87	15.28	13.97	13.57
Number of grades	37,899	27,191	136,410	127,721
Student Characteristics				
URM	0.16	0.14	0.15	0.13
Black or African American	0.04	0.03	0.06	0.05
Hispanic	0.12	0.11	0.07	0.07
Native American/Hawaiian or Other Pacific Islander	0.00	0.00	0.02	0.02
First-gen	0.14	0.15	0.16	0.13
Low-income	0.18	0.16	0.54	0.51
Concealment Policy Statistics				
Pr(Concealed grade)	0.149	0.215	0.076	0.129
Pr(Concealed grade | grade below GPA)	0.482	0.597	0.271	0.379
Pr(Concealed all grades)	0.037	0.019	0.019	0.008
Likelihood of concealing any grade	0.374	0.493	0.229	0.331
GPA for not concealed grades	3.65	3.56	3.73	3.66
GPA for concealed grades	2.53	2.43	2.66	2.67

The table reports summary statistics for Boston University and the Midwestern Flagship, respectively, by student gender as identified in administrative data. Eligible students is the number of students eligible for the policy. First-years, sophomores, juniors, and seniors refer to the number of students in each program year standing. For the Midwestern flagship, the numbers are based on the Winter 2021 term. The first panel reports summary statistics relevant to student performance: Cumulative GPA is the average Cumulative GPA prior to the term(s) with the information-optional policy, where both institutions use a 0–4 GPA scale. Credits earned is the average number of credits earned per student. Number of grades is the number of grades assigned eligible for the policy. The second panel reports summary statistics relevant to student characteristics: URM includes U.S. Citizens or U.S. Permanent Residents who have self-identified as belonging to any of the following race/ethnicity categories: Hispanic, Black or African American, Native American, Native Hawaiian, or Other Pacific Islander. First-gen students are those who are the first in their family to attend college. Low-income students include those with Pell grant status at program entry for Boston University and students whose parental income is in the bottom half at the Midwestern Flagship. The third panel presents summary statistics on take-up and performance related to the concealment policy: Pr(Concealed grade) is the likelihood of concealing a grade. Pr(Concealed grade | grade below GPA) is the likelihood of concealing a grade, conditional on the grade being below the prior GPA. Pr(Concealed all grades) is the likelihood of concealing all grades. Likelihood of concealing any grade is the percent of students choosing to conceal at least one grade, calculated at the student level. GPA for not concealed grades is the average grade (mapped onto the GPA scale) for classes taken for a letter grade. GPA for concealed grades is the average grade for classes taken for credit or no credit.

Table A.2: Alternative specifications for the gender gap in concealing grades

Panel A: Boston University					
	(1)	(2)	(3)	(4)	(5)
Female	-0.112***	-0.089***	-0.087***	-0.050***	-0.050***
	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)
Observations	18,686	18,679	18,671	18,671	18,671
Conceal mean	0.522	0.522	0.522	0.522	0.522
Year × ΔGPA FEs	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Major FE	No	No	No	Yes	Yes
Course level FE	No	No	No	No	Yes
Panel B: Midwestern Flagship					
	(1)	(2)	(3)	(4)	(5)
Female	-0.108***	-0.083***	-0.087***	-0.061***	-0.061***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	70,882	70,775	69,113	69,113	69,112
Conceal mean	0.329	0.329	0.329	0.329	0.329
Year × ΔGPA FEs	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Major FE	No	No	No	Yes	Yes
Course level FE	No	No	No	No	Yes

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful letter grade in a course. Female is a binary indicator equal to one when the student is listed as such in the administrative data. Conceal mean is the proportion of letter grades concealed for each column. Year×ΔGPA FEs are indicators for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Controls includes controls for number of courses, number of letter-graded courses, low-income status, first-generation status, race, and whether the student is an international student. Major FEs are indicators for each college major category, encompassing fields such as Engineering, Humanities, Natural Science, and Social Science. Course level FEs are indicators as to whether the course is a 100, 200, 300, or 400 level course. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.3: Gender gap in concealing harmful grades by letter grade

Panel A: Boston University								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A-	B+	B	B-	C+	C	C-	D
Female	-0.010	-0.046**	-0.107***	-0.110***	-0.068***	-0.075***	-0.010	0.040*
	(0.014)	(0.019)	(0.020)	(0.024)	(0.025)	(0.022)	(0.029)	(0.022)
Observations	2,143	3,994	4,028	2,550	1,540	1,578	719	862
Conceal mean	0.072	0.319	0.465	0.616	0.813	0.833	0.915	0.934
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
	A-	B+	B	B-	C+	C	C-	
Female	-0.020***	-0.081***	-0.119***	-0.103***	-0.058***	-0.064***	-0.020	
	(0.004)	(0.008)	(0.010)	(0.013)	(0.015)	(0.015)	(0.023)	
Observations	15,399	17,984	16,703	7,917	4,994	4,698	2,058	
Conceal mean	0.042	0.228	0.346	0.499	0.709	0.726	0.730	
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful letter grade in a course, by the letter grade the student received in a class. Female is a binary indicator equal to one when the student is listed as such in the administrative data. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.4: Gender gap in concealing by female ratio in classes

Panel A: Boston University			
	(1)	(2)	(3)
	Male dominated	Female dominated	Gender balanced
Female	-0.037***	-0.081***	-0.126
	(0.012)	(0.011)	(0.090)
Observations	7,908	10,282	165
Conceal mean	0.626	0.442	0.558
Year × ΔGPA FEs	Yes	Yes	Yes
p-value Col(1) vs. Col(2)		0.006	
Panel B: Midwestern Flagship			
	(1)	(2)	(3)
	Male dominated	Female dominated	Gender balanced
Female	-0.041***	-0.091***	-0.058
	(0.007)	(0.007)	(0.062)
Observations	41,635	28,551	281
Conceal mean	0.385	0.248	0.295
Year × ΔGPA FEs	Yes	Yes	Yes
p-value Col(1) vs. Col(2)		0.000	

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, by the gender ratio in a class. Column 1 includes classes that have more men that women enrolled. Column 2 includes classes that have more women that men enrolled. Column 3 includes classes that have the exact same number of men and women enrolled. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.5: Gender gap in concealing by course size

Panel A: Boston University					
	(1)	(2)	(3)	(4)	(5)
	Lowest	2nd quintile	3rd quintile	4th quintile	Highest
Female	-0.046**	-0.079***	-0.097***	-0.085***	-0.130***
	(0.020)	(0.018)	(0.018)	(0.018)	(0.019)
Observations	3,388	3,615	3,558	3,655	3,494
Conceal mean	0.434	0.515	0.529	0.596	0.517
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship					
	(1)	(2)	(3)	(4)	(5)
	Lowest	2nd quintile	3rd quintile	4th quintile	Highest
Female	-0.065***	-0.083***	-0.112***	-0.109***	-0.047***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
Observations	14,382	13,703	13,998	14,541	13,621
Conceal mean	0.297	0.318	0.350	0.342	0.341
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, by the quintiles of class size. Column 1 includes classes in the 0–19.99th percentile of class size, Column 2 includes classes in the 20–39.99th percentile of class size, Column 3 includes classes in the 40–59.99th percentile of class size, Column 4 includes classes in the 60–79.99th percentile of class size, and Column 5 includes classes in the 80–100th percentile of class size. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.6: Gender gap in concealing grades by class characteristics

Panel A: Boston University						
	(1)	(2)	(3)	(4)	(5)	(6)
	Low demand	High demand	Low average	High average	Low stand. dev.	High stand. dev.
Female	-0.074***	-0.107***	-0.096***	-0.077***	-0.079***	-0.099***
	(0.011)	(0.012)	(0.011)	(0.012)	(0.012)	(0.011)
Observations	10,310	8,304	9,379	9,233	9,494	9,071
Conceal mean	0.533	0.508	0.594	0.448	0.474	0.572
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes
p-value	0.024		0.188		0.164	
Panel B: Midwestern Flagship						
	(1)	(2)	(3)	(4)	(5)	(6)
	Low demand	High demand	Low average	High average	Low stand. dev.	High stand. dev.
Female	-0.084***	-0.081***	-0.070***	-0.101***	-0.104***	-0.065***
	(0.005)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	47,394	23,346	34,938	35,790	36,505	34,174
Conceal mean	0.336	0.314	0.367	0.292	0.291	0.369
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes
p-value	0.710		0.000		0.000	

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, by class characteristics. Columns 1–2 split the data according to whether a class is among the ten most popular classes in its major or not, determined by the ratio of students in each major taking that specific course. Columns 3–4 do a median split of the data according the average grades received in classes. Columns 5–6 do a median split of the data according to the standard deviation of class grades. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.7: Gender gap in concealing by GPA at the start of the policy

Panel A: Boston University					
	(1)	(2)	(3)	(4)	(5)
	≤2.50	2.51–3.00	3.01–3.50	3.51–3.75	>3.75
Female	0.003	-0.043*	-0.089***	-0.072***	-0.052**
	(0.050)	(0.024)	(0.013)	(0.020)	(0.020)
Observations	738	2,536	7,951	3,729	2,706
Conceal mean	0.768	0.730	0.534	0.448	0.286
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship					
	(1)	(2)	(3)	(4)	(5)
	≤2.50	2.51–3.00	3.01–3.50	3.51–3.75	>3.75
Female	-0.100*	-0.078***	-0.091***	-0.090***	-0.051***
	(0.054)	(0.024)	(0.011)	(0.011)	(0.006)
Observations	838	3,696	17,373	16,321	31,867
Conceal mean	0.399	0.525	0.411	0.354	0.244
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, by the student's GPA at the start of the term. Column 1 includes students with a GPA at or below 2.50 at the start of the term, Column 2 includes students with a GPA ranging from 2.51 to 3.00 at the start of the term, Column 3 includes students with a GPA ranging from 3.01 to 3.50 at the start of the term, Column 4 includes students with a GPA ranging from 3.51 to 3.75 at the start of the term, and Column 5 includes students with a GPA higher than 3.75 at the start of the term. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.8: Gender gap in concealing grades across program years

Panel A: Boston University				
	(1)	(2)	(3)	(4)
	First-year	Sophomore	Junior	Senior
Female	-0.084***	-0.114***	-0.097***	-0.045**
	(0.017)	(0.015)	(0.018)	(0.018)
Observations	4,719	5,801	4,342	3,817
Conceal mean	0.498	0.519	0.558	0.516
Year × ΔGPA FEs	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship				
	(1)	(2)	(3)	(4)
	First-year	Sophomore	Junior	Senior
Female	-0.043***	-0.073***	-0.091***	-0.089***
	(0.013)	(0.008)	(0.009)	(0.008)
Observations	4,502	17,163	22,243	26,867
Conceal mean	0.224	0.285	0.346	0.361
Year × ΔGPA FEs	Yes	Yes	Yes	Yes

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, by the student's year in school noted in the column header. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.9: Gender gap in concealing grades

Panel A: Boston University						
	(1)	(2)	(3)	(4)	(5)	(6)
	Arts+Humanities	Business+Economics	Health+Medicine	Social science	STEM	Undecided
Female	-0.014	-0.023	-0.120***	-0.002	-0.102***	-0.147***
	(0.031)	(0.018)	(0.018)	(0.025)	(0.014)	(0.039)
Observations	1,911	4,398	4,556	2,912	6,685	1,309
Conceal mean	0.419	0.668	0.597	0.494	0.526	0.521
Female Mean	0.685	0.424	0.530	0.687	0.436	0.621
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship						
	(1)	(2)	(3)	(4)	(5)	(6)
	Arts+Humanities	Business+Economics	Health+Medicine	Social science	STEM	Undecided
Female	-0.035*	-0.082***	-0.166***	-0.083***	-0.055***	-0.069***
	(0.016)	(0.014)	(0.014)	(0.017)	(0.008)	(0.009)
Observations	6,719	12,004	11,513	5,626	27,999	15,443
Conceal mean	0.258	0.438	0.371	0.298	0.355	0.276
Female Mean	0.647	0.367	0.472	0.564	0.375	0.522
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, by the student's major noted in the column header. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.10: Gender gap in concealing by whether the course is within or outside of the student's major

Panel A: Boston University		
	(1)	(2)
	Outside major	Within major
Female	-0.086***	-0.091***
	(0.009)	(0.018)
Observations	14,872	3,641
Conceal mean	0.526	0.505
Year × ΔGPA FEs	Yes	Yes
p-value	0.802	
Panel B: Midwestern Flagship		
	(1)	(2)
	Outside major	Within major
Female	-0.078***	-0.096***
	(0.005)	(0.010)
Observations	58,224	12,406
Conceal mean	0.344	0.260
Year × ΔGPA FEs	Yes	Yes
p-value	0.083	

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful letter grade in a course, by whether the course is within or outside of their major. Female is a binary indicator equal to one when the student is listed as such in the administrative data. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.11: Gender gap in concealing grades across demographic subgroups

Panel A: Boston University						
	(1)	(2)	(3)	(4)	(5)	(6)
	URM	Non URM	Low-income	Non low-income	First gen	Non first gen
Female	-0.056**	-0.095***	-0.116***	-0.089***	-0.037	-0.098***
	(0.025)	(0.009)	(0.024)	(0.009)	(0.024)	(0.009)
Observations	2,586	15,847	3,033	15,434	2,763	15,677
Conceal mean	0.526	0.522	0.541	0.517	0.557	0.516
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes
p-value	0.111		0.237		0.009	
Panel B: Midwestern Flagship						
	(1)	(2)	(3)	(4)	(5)	(6)
	URM	Non URM	Low-income	Non low-income	First gen	Non first gen
Female	-0.098***	-0.084***	-0.085***	-0.078***	-0.087***	-0.081***
	(0.014)	(0.005)	(0.008)	(0.008)	(0.013)	(0.005)
Observations	10,178	55,488	28,783	24,067	10,354	58,673
Conceal mean	0.349	0.318	0.330	0.327	0.335	0.327
Year × ΔGPA FEs	Yes	Yes	Yes	Yes	Yes	Yes
p-value	0.354		0.591		0.672	

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course, across demographic subgroups. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Column 1 examines U.S. Citizens or U.S. Permanent Residents who have self-identified as belonging to any of the following race/ethnicity categories: Hispanic, Native American, Black or African American, Native Hawaiian or Other Pacific Islander; Column 2 examines Non-URM students; Column 3 examines low-income students, corresponding to students with Pell grant status at program entry for Boston University and students whose parental income is in the bottom half at the Midwestern Flagship; Column 4 examines students who do not qualify as low-income; Column 5 examines students who are the first in their family to attend college; Column 6 examines non-first generation students. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A.12: Gender gap in concealing by each term at the Midwestern Flagship

	(1)	(2)	(3)
	Fall 2020	Spring 2021	Spring 2021|concealed before
Female	-0.097***	-0.074***	-0.056***
	(0.007)	(0.005)	(0.010)
Observations	30,543	40,135	15,179
Conceal mean	0.351	0.312	0.495
Year × ΔGPA FEs	Yes	Yes	Yes
p-value		0.001	

This table shows estimates from a linear probability model of whether a student chooses to conceal a harmful grade in a course. Column 1 presents data from Fall 2020 at the Midwestern Flagship, while Column 2 presents data from Spring 2021 at the Midwestern Flagship. Column 3 also presents data from Spring 2021 but is further restricted to students who had concealed at least one grade in Fall 2020. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. No results are shown for BU because we only have data for one semester at BU. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

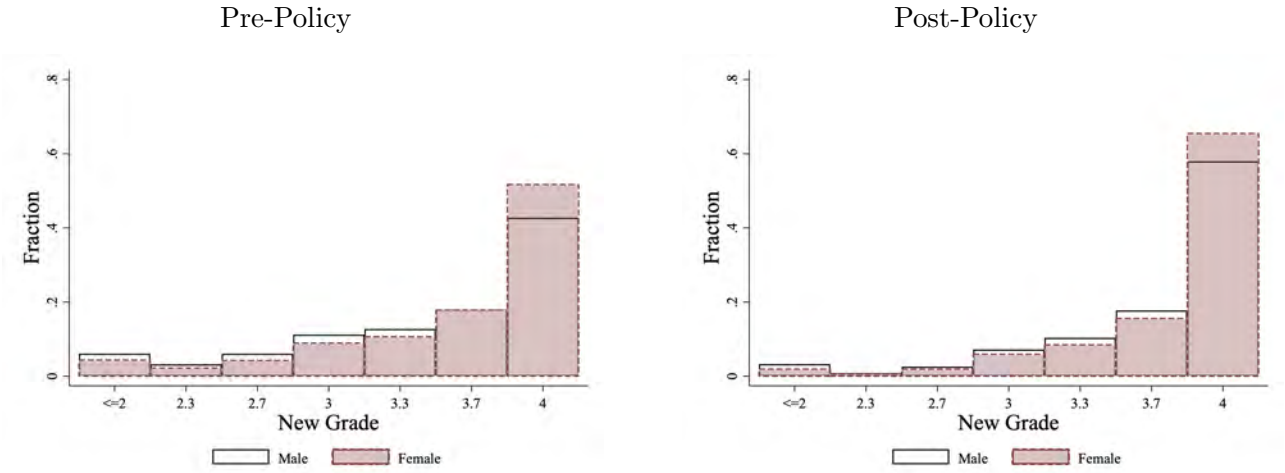
Table A.13: Gender gap in concealing by course level

Panel A: Boston University				
	(1)	(2)	(3)	(4)
	100s	200s	300s	400s and above
Female	-0.099***	-0.097***	-0.064***	-0.089***
	(0.015)	(0.015)	(0.019)	(0.018)
Observations	5,214	5,452	3,380	3,992
Conceal mean	0.547	0.509	0.560	0.475
Year × ΔGPA FEs	Yes	Yes	Yes	Yes
Panel B: Midwestern Flagship				
	(1)	(2)	(3)	(4)
	100s	200s	300s	400s and above
Female	-0.054***	-0.099***	-0.085***	-0.077***
	(0.009)	(0.007)	(0.008)	(0.008)
Observations	10,734	19,236	19,856	20,591
Conceal mean	0.309	0.308	0.355	0.334
Year × ΔGPA FEs	Yes	Yes	Yes	Yes

This table presents estimates from a linear probability model examining whether a student conceals a harmful grade in a course, stratified by course level. Column 1 includes 100-level courses, Column 2 includes 200-level courses, Column 3 includes 300-level courses, and Column 4 includes 400-level and above courses. Female is a binary indicator equal to one when the student is listed as such in the administrative data; the data identify students as male or female. Conceal mean is the proportion of grades concealed for each column. Year × ΔGPA fixed effects include an indicator for every possible change in GPA, discretized into 1,000 intervals separately for each program year. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

B Additional figures

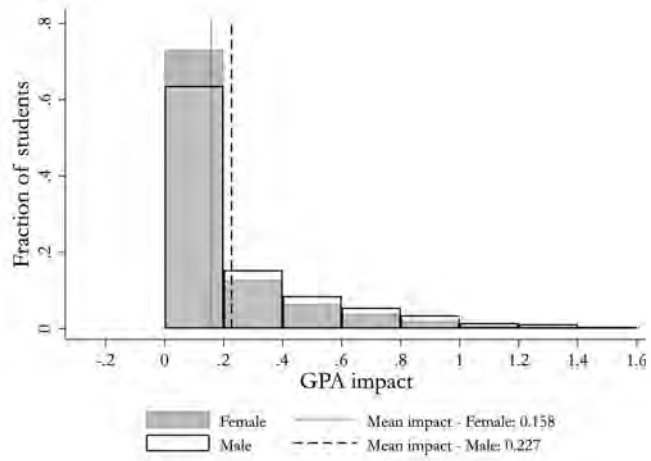
Figure B.1: Distribution of Grades Received Pre- and Post-Policy, by Gender



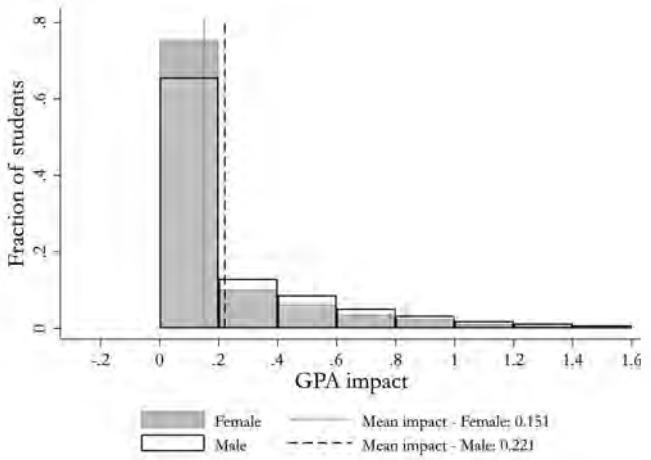
This figure displays the distribution of grades received by students at a Midwestern flagship university in Fall 2020 (pre-policy) and Fall 2021 (post-policy), disaggregated by gender. The grades reflect raw, unconcealed values prior to the implementation of any grade concealing policies. Letter grades are converted to the 4.0 GPA scale: 4.0 corresponds to an A or A+, 3.7 to an A-, and so on. Grades of C or below are grouped into a single category labeled "<=2". The average pre-policy grade was 3.58 for women and 3.46 for men. After the policy, the average grade was 3.74 for women and 3.64 for men.

Figure B.2: Distributional impacts of the grade-optional policies on GPA

Boston University



Midwestern Flagship



This figure plots the distribution of changes in students' GPA from the grade-optional policy, subtracting each student's GPA assuming no optional-information policy (i.e. assuming students revealed all of their letter grades) from their effective GPA after making their grade-concealment decisions. The figure compares the impact among male students (white bars with black outline; average impact shown by the dashed black line) and female students (gray bars with no outline; the average impact shown by the gray solid line). The distribution is truncated at the bottom and top 1%. The impact is calculated over one term for BU and two terms for the Midwestern Flagship, given the timing of the policy at each institution. This figure includes every grade obtained for each term covered by the policy at either institution, which involved 65,090 grades at BU from the decisions made by 15,690 students and 264,131 grades at the Midwestern Flagship from the decisions made by 37,574 students.

C The student belief study

C.1 Overview of the student beliefs study

In the main part of the study, students answer two incentivized prediction questions. These questions ask students to guess the GPA that an employer would infer that a male or female applicant has when deciding whether to interview them. In particular, students are asked: "When deciding whether to interview a [man/woman] and [his/her] resume did not have any GPA information, on average, employers treated [him/her] as if [his/her] GPA was..."

These questions are incentivized based on evaluations of job candidates in a complementary experiment among actual employers engaged in on-campus recruiting at the University of Pennsylvania (see Appendix D for more details). Students in our survey learn about the employers' evaluation decisions in detail. First, they are told about the employers' incentives: although employers are evaluating hypothetical resumes, employers' evaluations are incentivized because they will be used to provide recommendations of actual University of Pennsylvania graduating seniors who might be good candidates for their open positions. Second, they are provided with information on the resumes that randomly vary in terms of: (i) the candidates' name, education information, work experience, leadership experience, and other skills; (ii) whether the GPA information is shown on the resume; and (iii) the GPA value, which ranges from 3.0 to 4.0 when it is shown on the resume. Third, to incentivize the students' answers to the prediction questions, students know that their chance of receiving the $20 bonus is higher if they answer each prediction question correctly.[29]

The prediction questions reveal whether students expect employers to treat women as if they have a lower GPA when their GPA information is not shown. To shed light on the generalizability of these findings beyond beliefs about employers' assessments of GPA (the questions that we can incentivize with actual employer data), we also ask three unincentivized questions. Students are asked to consider a man and a woman from the same large university in the United States who have exactly the same GPA, the same transcript, and the same resume. They are then asked to make predictions about whether women will be assessed more negatively, men will be assessed more negatively, or men and women will be assessed similarly when:

1. Employers are asked to make predictions about their GPAs when they are applying to exactly the same job;

[29] In particular, participants are accurately told that, for one randomly selected participant out of every 100 participants, we will randomly select one prediction question and they will receive the $20 bonus if their guess in that prediction question contains the right answer.

2. Employers are asked to make predictions about their grades when they both took the same class for "credit" and are applying for the same job;

3. Graduate school admission committees are asked to make predictions about their grades when they both took the same class for "credit" and are applying for the same graduate school program.

While Questions 2-3 connect more closely to the concealment gap we observe in the transcript data because they speak to whether students expect employers to treat women as if they have worse grades when information on their individual courses is not provided, Question 1 (as well as the incentivized prediction questions) about employers' beliefs about GPAs when information on the GPA is not provided speaks to the potentially broader range of settings in which expected discrimination may arise. These latter settings may be particularly interesting for future work to investigate, including the potential relevance of observing other concealment gaps (e.g., concealment gaps may align with expected discrimination and hence help to inform information-optional policies).

C.2 Full experimental instructions for the student belief study

After consenting to participate in the study, each participant is informed about the $10 study completion fee and the opportunity to earn additional payments. Figure C.1 shows the overview provided to the participants. Then, participants are shown the introduction in Figure C.2 about a complementary experiment among actual employers engaged in on-campus recruiting at the University of Pennsylvania, specifically regarding how employers make evaluation decisions.

Next, participants proceed to the main part of the experiment, where they answer two incentivized prediction questions about employer evaluations, as shown in Figure C.3.

Additionally, we ask three hypothetical questions where participants predict whether employers or graduate school admissions committees will infer that a man or woman has worse (or equal) performance based on missing information, as illustrated in Figures C.4 through C.7.

Finally, an open-response question is asked about whether and why students would choose to (not) conceal a grade (Figure C.8). After completing these questions, participants complete a short follow-up survey that collects demographic information, partly shown in Figure C.9.

Figure C.1: Study overview

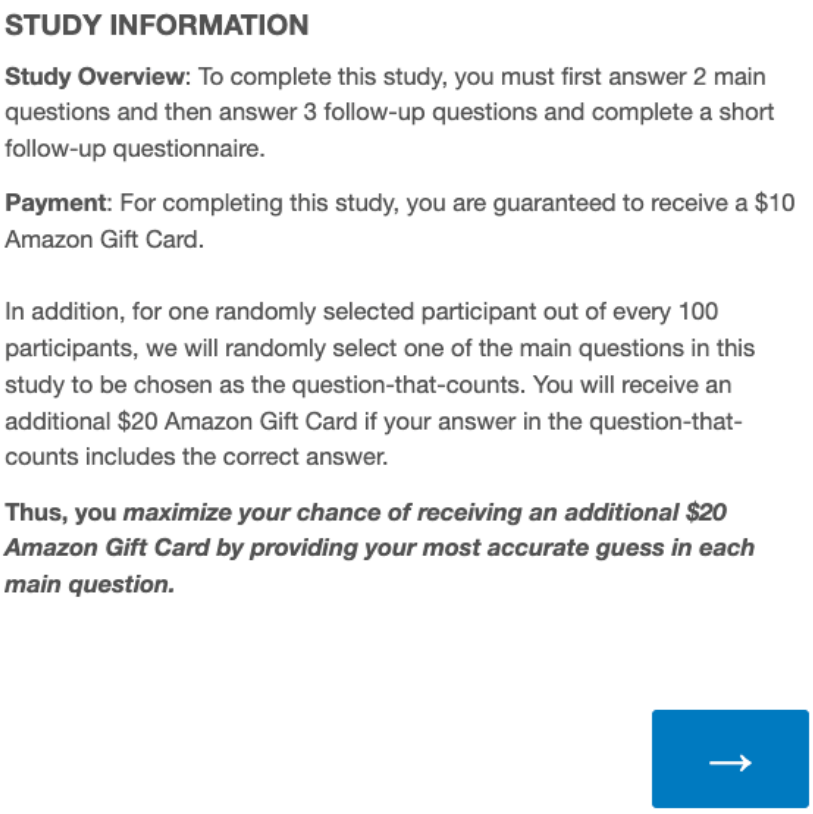
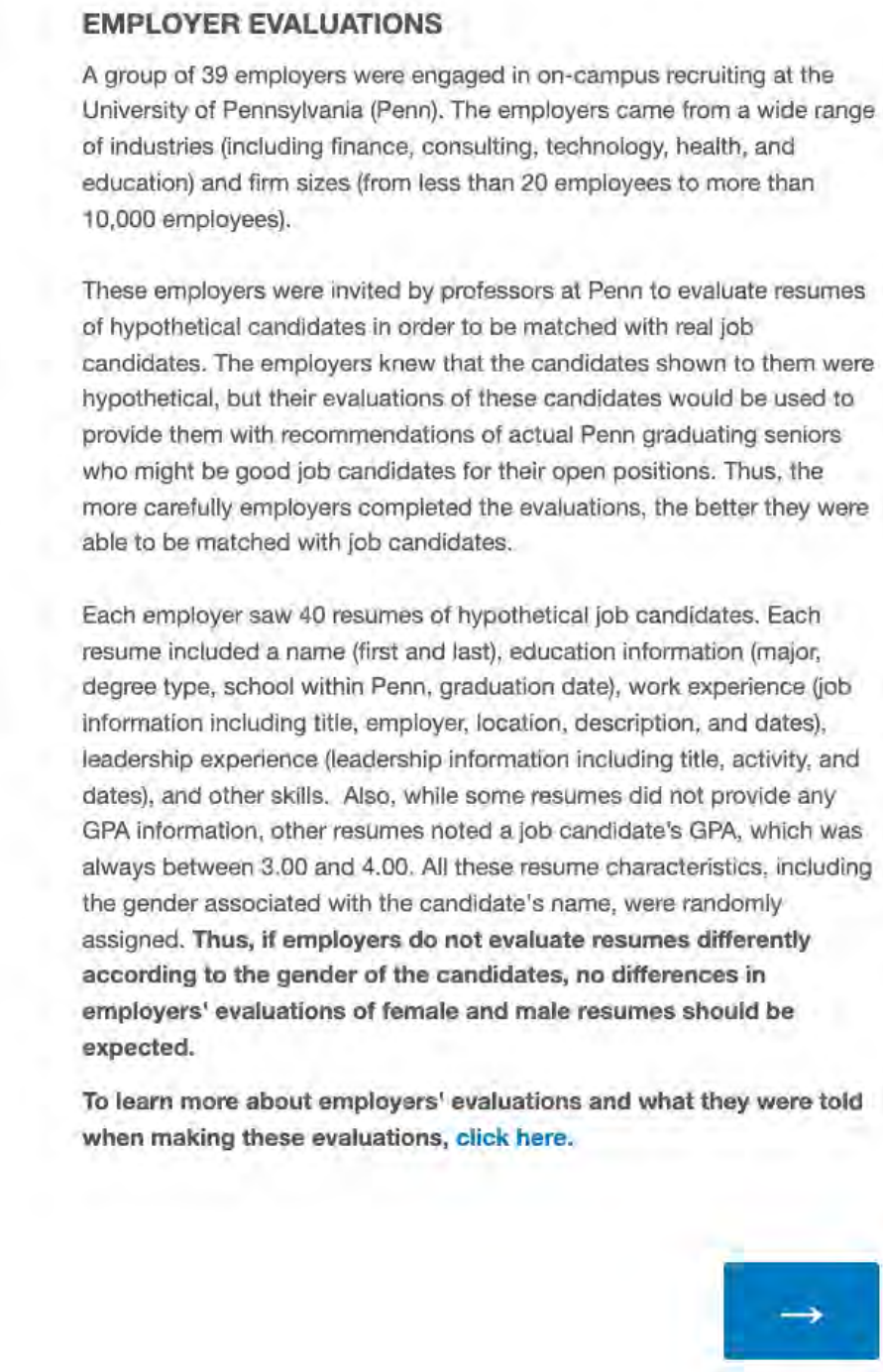


Figure C.2: Introduction about employer evaluations



EMPLOYER EVALUATIONS

A group of 39 employers were engaged in on-campus recruiting at the University of Pennsylvania (Penn). The employers came from a wide range of industries (including finance, consulting, technology, health, and education) and firm sizes (from less than 20 employees to more than 10,000 employees).

These employers were invited by professors at Penn to evaluate resumes of hypothetical candidates in order to be matched with real job candidates. The employers knew that the candidates shown to them were hypothetical, but their evaluations of these candidates would be used to provide them with recommendations of actual Penn graduating seniors who might be good job candidates for their open positions. Thus, the more carefully employers completed the evaluations, the better they were able to be matched with job candidates.

Each employer saw 40 resumes of hypothetical job candidates. Each resume included a name (first and last), education information (major, degree type, school within Penn, graduation date), work experience (job information including title, employer, location, description, and dates), leadership experience (leadership information including title, activity, and dates), and other skills. Also, while some resumes did not provide any GPA information, other resumes noted a job candidate's GPA, which was always between 3.00 and 4.00. All these resume characteristics, including the gender associated with the candidate's name, were randomly assigned. **Thus, if employers do not evaluate resumes differently according to the gender of the candidates, no differences in employers' evaluations of female and male resumes should be expected.**

To learn more about employers' evaluations and what they were told when making these evaluations, click here.

→

Figure C.3: Study main questions

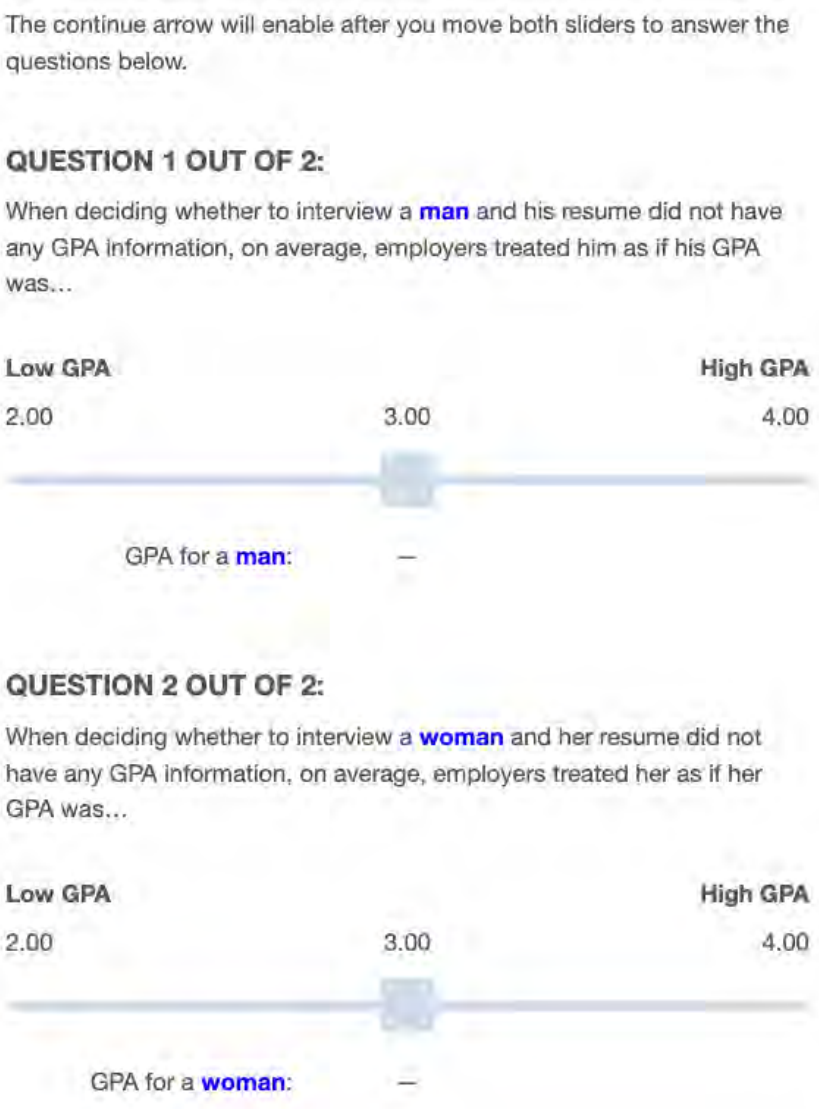


Figure C.4: Instruction for additional questions

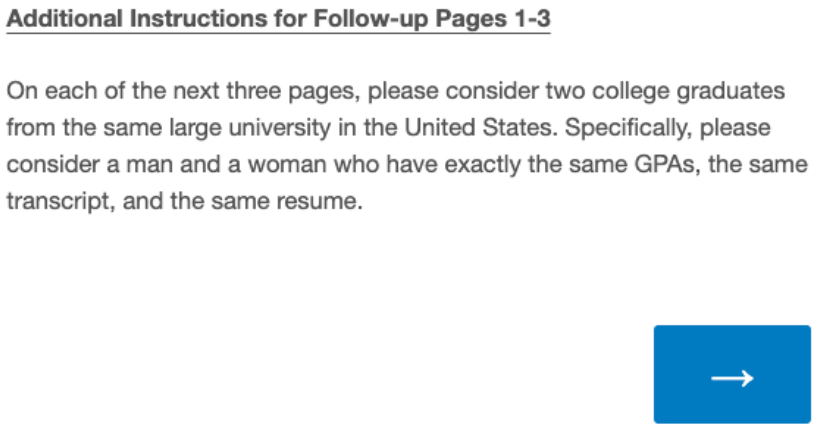


Figure C.5: Job application with no GPA

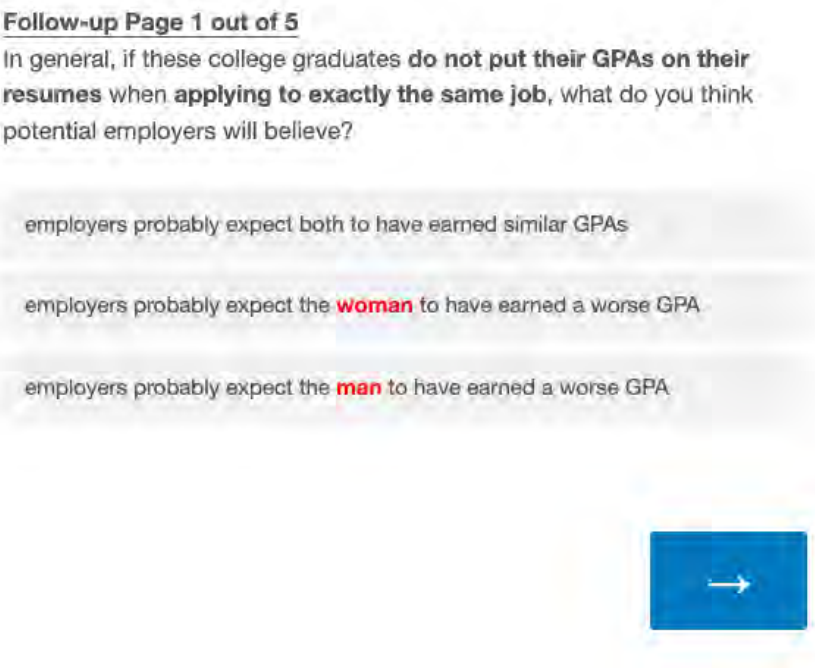


Figure C.6: Job application with a grade of "Credit"

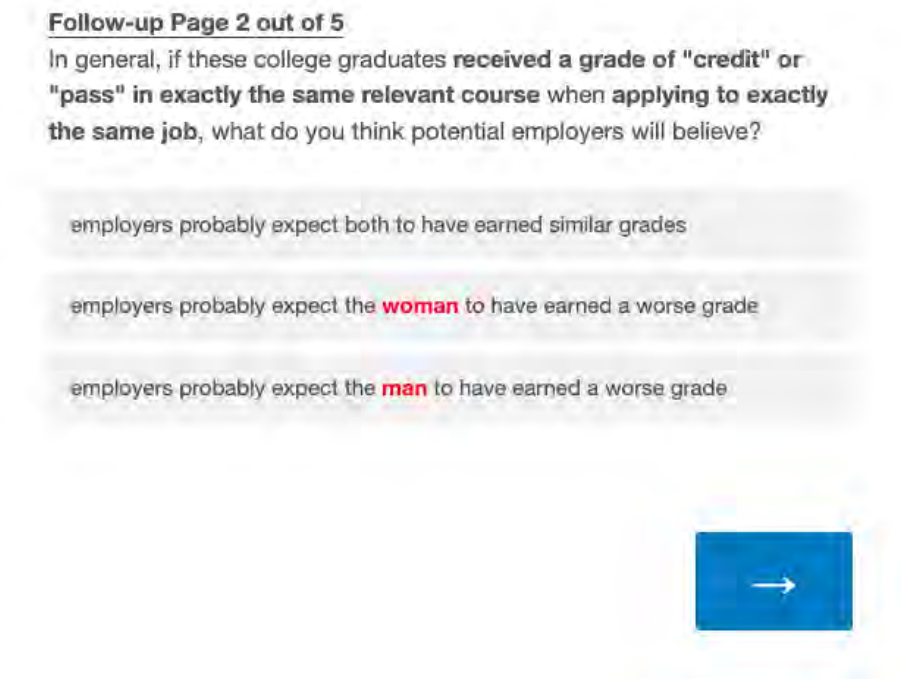


Figure C.7: Graduate school application with a grade of "Credit"

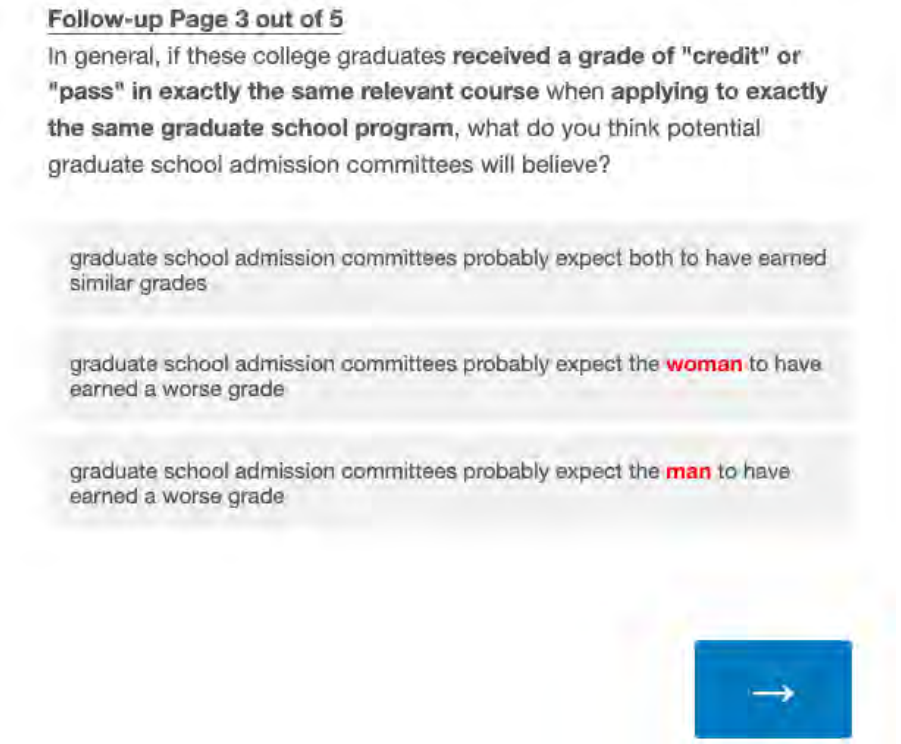


Figure C.8: Open question about grade concealment

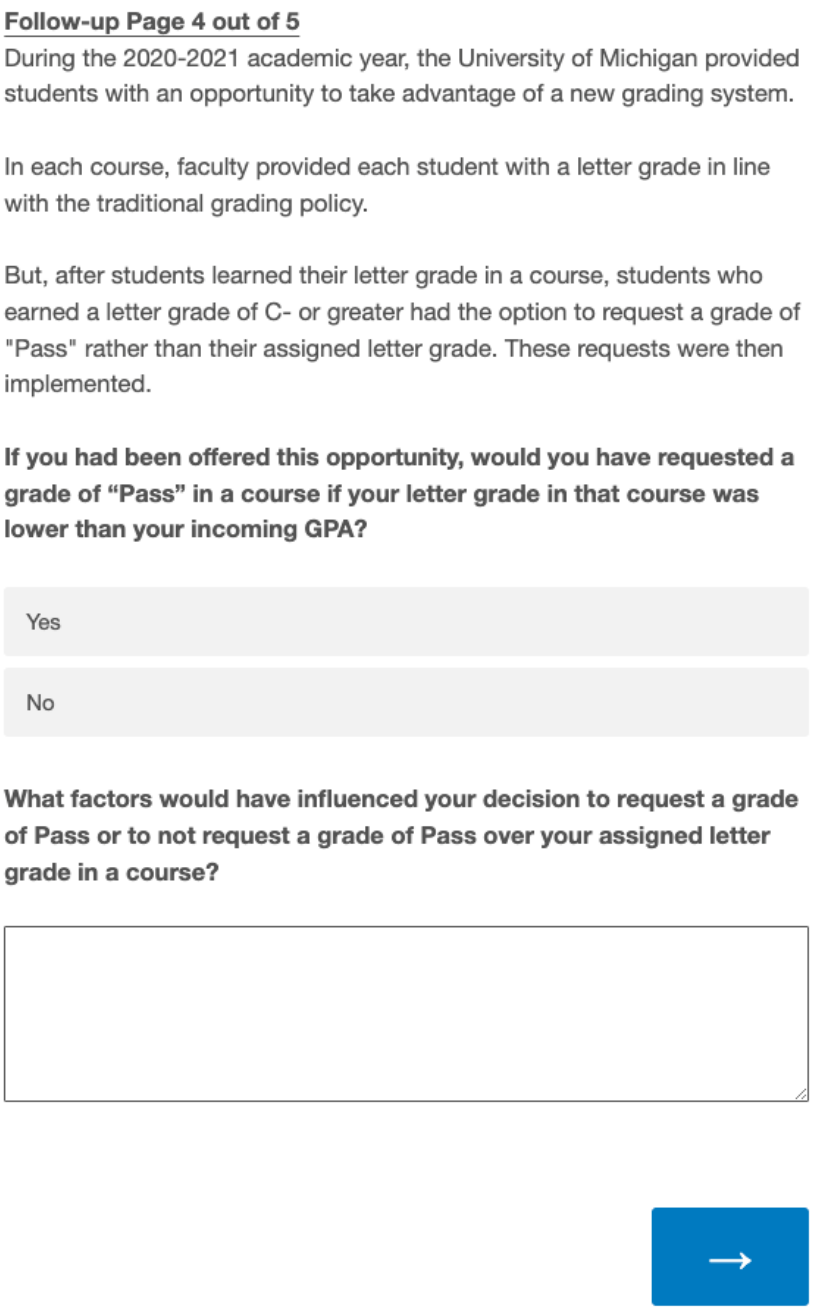
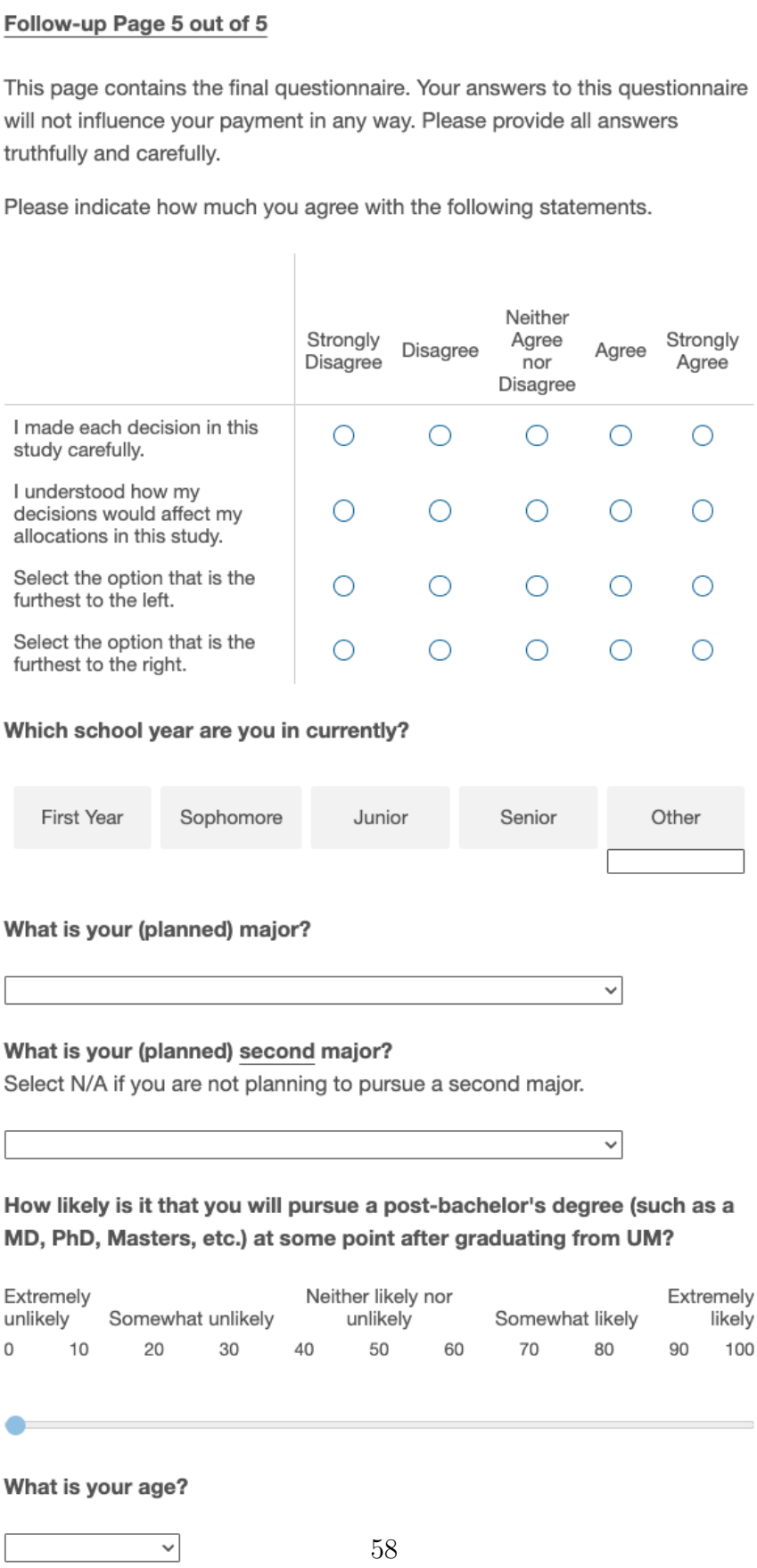


Figure C.9: Short followup survey



Follow-up Page 5 out of 5

This page contains the final questionnaire. Your answers to this questionnaire will not influence your payment in any way. Please provide all answers truthfully and carefully.

Please indicate how much you agree with the following statements.

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I made each decision in this study carefully.	○	○	○	○	○
I understood how my decisions would affect my allocations in this study.	○	○	○	○	○
Select the option that is the furthest to the left.	○	○	○	○	○
Select the option that is the furthest to the right.	○	○	○	○	○

Which school year are you in currently?

First Year | Sophomore | Junior | Senior | Other

What is your (planned) major?

What is your (planned) second major?

Select N/A if you are not planning to pursue a second major.

How likely is it that you will pursue a post-bachelor's degree (such as a MD, PhD, Masters, etc.) at some point after graduating from UM?

Extremely unlikely — Somewhat unlikely — Neither likely nor unlikely — Somewhat likely — Extremely likely

0 10 20 30 40 50 60 70 80 90 100

What is your age?

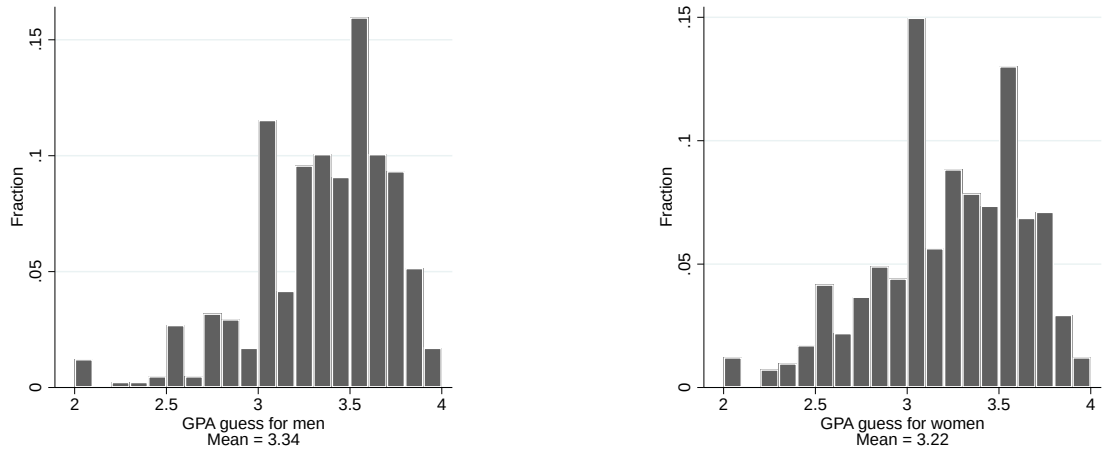
C.3 Additional results from the student belief study

Appendix Figure C.10 shows that students expect employers to make worse inferences about the performance of women compared to men. On average, when GPA information is not provided, they expect employers to infer that a male student has a GPA of 3.34 but that a female student only has a GPA of 3.22 ($p < 0.01$).

These beliefs persist when restricting to the employer inferences that students themselves may expect to face given their own gender. Specifically, when GPA information is not provided, men expect that employers will infer that male students have a GPA of 3.33 while women expect that employers will infer that female students only have a GPA of 3.21 ($p < 0.01$).

These beliefs also persist when exploiting within-subject level data. While 56% of students indicate that they expect employers to make worse inferences about women than men, only 28% expect the opposite (and 16% expect the inferences about men and women to be the same).

Figure C.10: Expected GPA inferences by employers



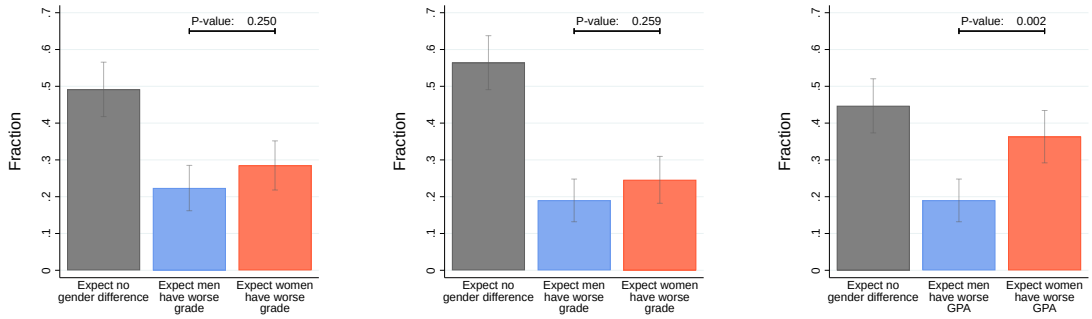
(a) Expected GPA for men

(b) Expected GPA for women

This figure plots the distribution of students' incentivized predictions of employers' inferences about GPA, when this information is missing from an applicant's resume. Panel (a) shows predictions about men, while Panel (b) shows predictions about women. All answers were provided on a slider form that allowed students to select a 0.07 point GPA range on a slider from a GPA of 2.00 to 4.00. This figure plots the midpoint of this range. The sample includes 407 predictions for each gender.

Finally, Appendix Figures C.11 and C.12 show that both male and female students believe employers and graduate school admissions committees will hold more negative inferences about a woman than a man with a grade of credit or a concealed GPA, although these relative expectations are stronger for women.

Figure C.11: Male Students: Expectations of relative inferences by gender, for employers and admissions committees



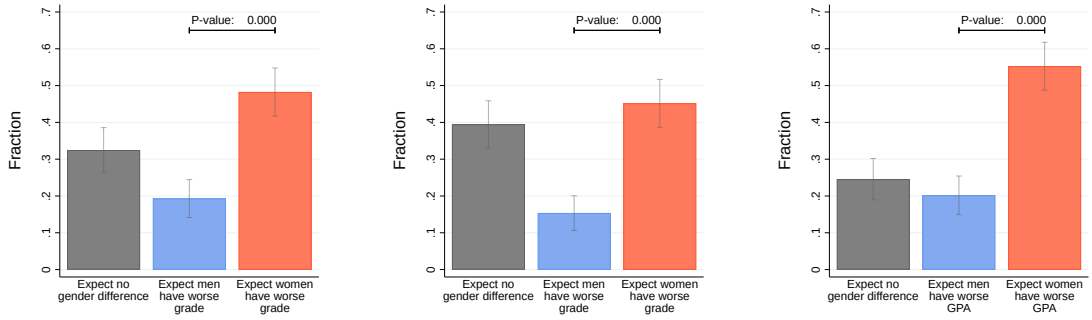
(a) Expected grade, employers

(b) Expected grade, admissions committees

(c) Expected GPA, employers

This figure plots the percent of male students who expected the same inference to be made about men and women, worse inferences for men, and worse inferences for women, respectively, for three different scenarios, as in Figure 2. The sample includes 179 predictions for each scenario.

Figure C.12: Female Students: Expectations of relative inferences by gender, for employers and admissions committees



(a) Expected grade, employers

(b) Expected grade, admissions committees

(c) Expected GPA, employers

This figure plots the percent of female students who expected the same inference to be made about men and women, worse inferences for men, and worse inferences for women, respectively, for three different scenarios, as in Figure 2. The sample includes 228 predictions for each scenario.

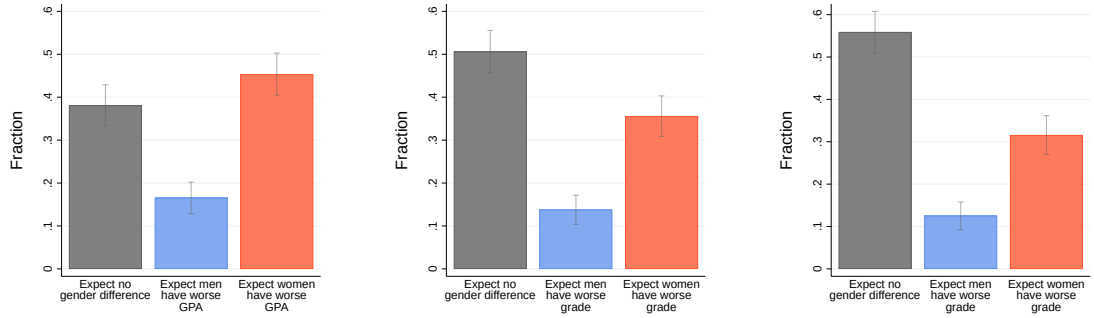
C.4 The general public belief study

This general public belief study is very similar to the student belief study, except the incentives are adjusted to appropriate levels for the Prolific platform.[30]

To be eligible for our study, participants needed to have completed at least 100 prior submissions on Prolific with an approval rating of 95% or greater and to have chosen the United States as their residence. Also, since we recruited a gender-balanced sample, participants must have selected either Male or Female for their sex on the Prolific platform. We recruited 399 participants for this study in June 2024.

Appendix Figures C.13 and C.16 show that the general public also expects employers and admissions committees to make relatively worse inferences about women. When considering each of the same set of measures as in the Student Study, this difference is statistically significant ($p < 0.01$).

Figure C.13: Expectations of relative inferences by gender, for employers and admissions committees



(a) Expected GPA, employers

(b) Expected grade, employers

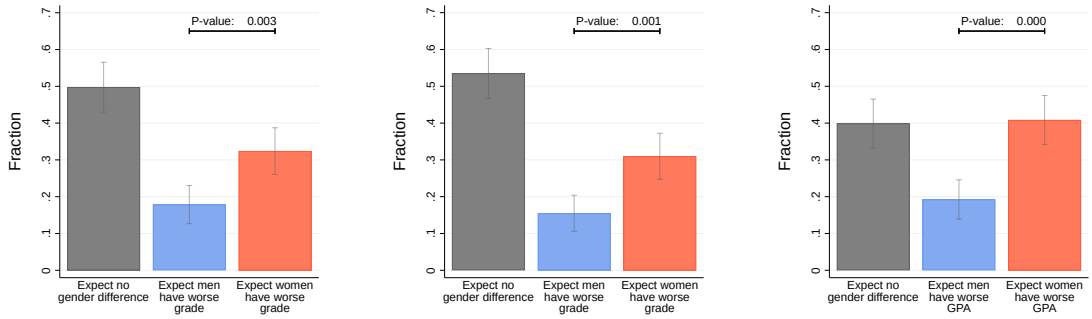
(c) Expected grade, admissions committees

This figure plots the percent of Prolific participants who expected the same inference to be made about men and women, worse inferences for men, and worse inferences for women, respectively, for three different scenarios: Panel (a): employers assessing a man and woman applying to the same job who have missing GPA information on their resumes; Panel (b): employers assessing a man and woman applying to the same job who received a grade of "credit" or "pass" in the same relevant course (and therefore have missing grade information); and Panel (c): graduate school admissions committees assessing a man and woman applying to the same graduate program who received a grade of "credit" or "pass" in the same relevant course (and therefore have missing grade information). Gray bars reflect 95% confidence intervals. The sample includes 407 predictions for each scenario.

[30]Participants are informed that they will receive a guaranteed payment of $2 within 24 hours for completing the study. Additionally, one of their predictions in the study will be randomly selected as the "question-that-counts." Participants earn a bonus payment of $1 if their answer to the "question-that-counts" is correct. There are also other small differences, such as cutting out the multiple choice question on whether they would have requested a grade of "Pass" before the open response about factors that influence this decision, and eliciting different demographic information given the change in the population.

Appendix Figures C.14 and C.15 show that both men and women believe employers and graduate school admissions committees will hold more negative inferences about a woman than a man with a grade of credit or a concealed GPA.

Figure C.14: Male participants: Expectations of relative inferences by gender, for employers and admissions committees



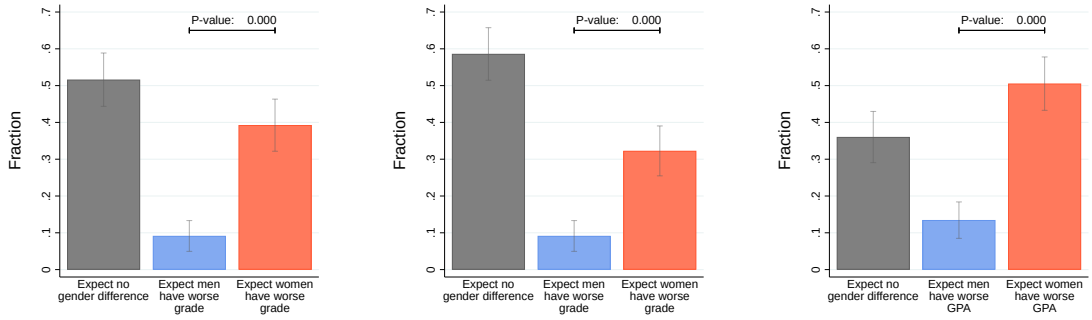
(a) Expected grade, employers

(b) Expected grade, admissions committees

(c) Expected GPA, employers

This figure plots the percent of male students who expected the same inference to be made about men and women, worse inferences for men, and worse inferences for women, respectively, for three different scenarios, as in Figure 2. The sample includes 198 predictions for each scenario.

Figure C.15: Female participants: Expectations of relative inferences by gender, for employers and admissions committees



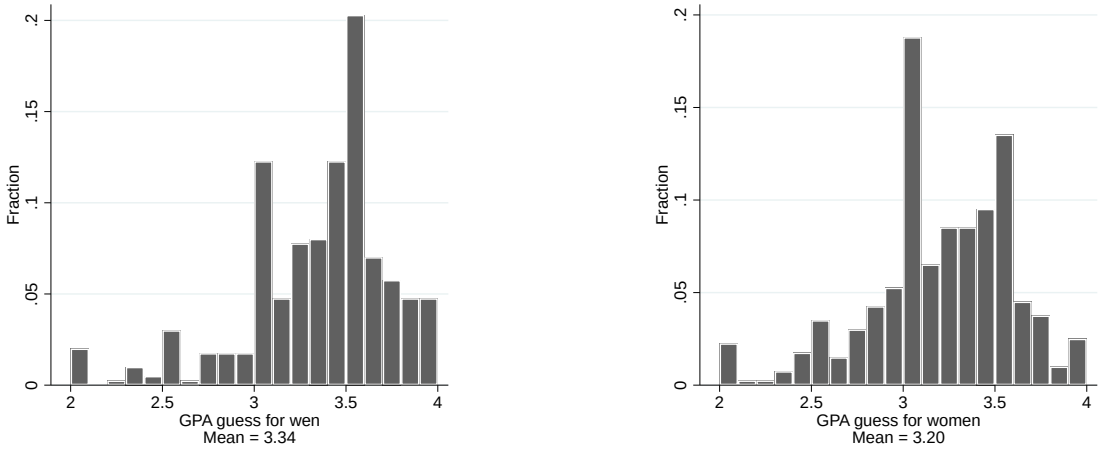
(a) Expected grade, employers

(b) Expected grade, admissions committees

(c) Expected GPA, employers

This figure plots the percent of female students who expected the same inference to be made about men and women, worse inferences for men, and worse inferences for women, respectively, for three different scenarios, as in Figure 2. The sample includes 186 predictions for each scenario.

Figure C.16: When a result did not reveal GPA, men/women were expected to be treated as if they had the following GPAs:



(a) Expected GPA for men

(b) Expected GPA for women

This figure plots the distribution of Prolific participants' incentivized predictions of employers' inferences about GPA, when this information is missing from an applicant's resume. Panel A shows predictions about men, while Panel B shows predictions about women. All answers were provided on a slider form that allowed students to select a 0.07 point GPA range on a slider from a GPA of 2.00 to 4.00. This figure plots the midpoint of this range. The sample includes 399 predictions for each gender.

D The employer study

We recruited 39 actual employers hiring at the University of Pennsylvania during the 2020–2021 academic year, around when the grade-optional policies we study were implemented. The employers each rated 40 resumes, yielding 1,560 resume evaluations. Following the method first introduced by Kessler, Low and Sullivan (2019), employers rate hypothetical candidates with randomly assigned characteristics including GPA and name (indicative of gender), incentivized by being matched with 10 real Penn students via machine learning, based on their resume ratings.

Given the labor market in 2020–2021, there was less on-campus recruiting activity than during the original IRR experiment discussed in Kessler, Low and Sullivan (2019). Nevertheless, the 39 employers came from a wide range of industries (including finance, consulting, technology, health, and education) and firm sizes (from less than 20 employees to more than 10,000 employees)—as summarized in Table D.1. While participants were told their responses would be used for research, they were not recruited using language about research, and their primary motivation, and thus incentive, was to receive the 10 real student matches. This helps to alleviate concerns about experimenter demand effects raised by Agan, Cowgill and Gee (2023*b*).

D.1 Design overview of the employer study

Similar to the implementation in Kessler, Low and Sullivan (2019), all resume characteristics such as name, major, GPA, as well as work and leadership experiences were independently and randomly assigned. Table D.2 lists all resume components and how they are randomized in the tool. In this new iteration, the GPA of some students was randomly omitted, allowing us to analyze employer response to missing information. GPA is randomized in two steps. First, each resume had a 10% chance of GPA being omitted (set roughly equal to the fraction of student participants in Candidate Match at Penn who chose not to show GPA on their resumes). Second, when GPA was revealed, it was randomly drawn from a uniform distribution between 3.00 and 4.00.

We utilize the randomly assigned resume names to examine the interaction of performance information and gender. First names were chosen to be highly indicative of gender: 50% of the names were indicative of a female candidate (e.g., Claire, Emma, Michelle), and 50% were indicative of a male candidate (e.g., Adam, Luke, Scott), where indicative means statistically very likely for the specified gender and unlikely to be the other gender.

Employers were asked to provide a binary measure of whether they would interview the candidate, which we use as our outcome variable. We also asked how interested they were in

Table D.1: Characteristics of participating employers

Recruiter Characteristics	
Identify as Female	66.7%
Identify as White	60.5%
Has a Graduate Degree	35.9%
Has at least 2-Year Tenure in Organization	60.5%
Firm Industry	
Consulting	10.3%
Finance	23.1%
Education, Non-Profit, or Government	12.8%
Technology	17.9%
Health	7.7%
Others	28.2%
Firm Size (in Employees)	
1–19	17.1%
20–49	22.0%
50–99	17.1%
100–249	9.8%
250–999	9.8%
1,000–9,999	12.2%
10,000 or more	12.2%
Position Location (Multiple Responses Allowed)	
East Coast	87.2%
West Coast	30.8%
Midwest	17.9%
South	12.8%
International	10.3%

This table shows descriptive statistics about the employers who participated in the incentivized resume rating experiment in the 2020–2021 academic year.

the candidate, ignoring likelihood of acceptance, and how likely they thought the candidate was to accept their offer, and can use these measures as robustness checks. We use the interview probability as our main outcome variable since it is easiest to interpret.

D.2 Results from the employer study

Figure D.1 reports the two main results of the study, using the binary outcome measure of whether an employer would invite a candidate for an interview.

First, Panel A demonstrates substantial employment returns to GPA: candidates with a one-point higher GPA are 18 percentage points more likely to receive an interview. These higher returns to GPA suggest a relative GPA shift in favor of men because women are less likely to use the grade concealment policy is likely to benefit the employment outcomes of

Table D.2: Randomization of resume components

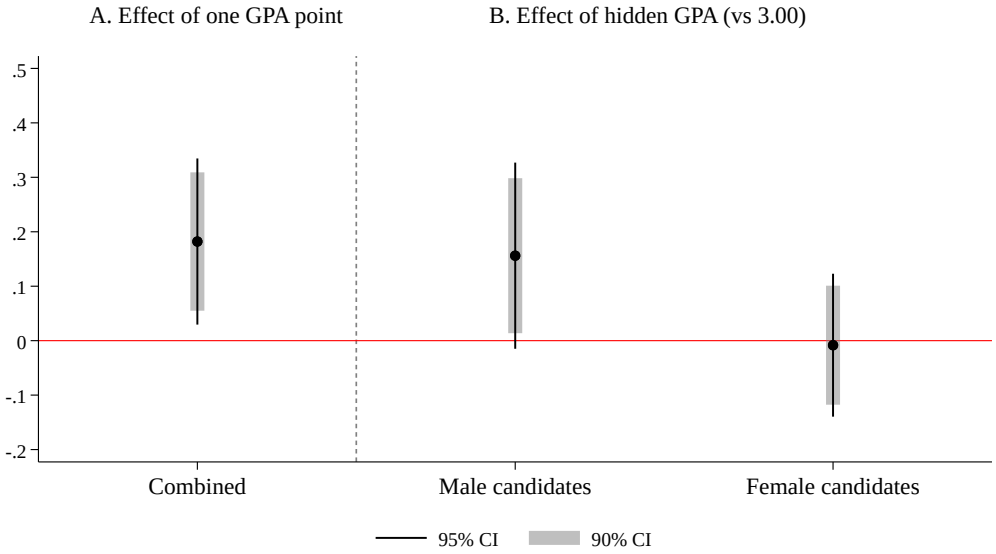
Resume Component	Description
Personal Information	
First & last name	50% Male, 50% Female; drawn from list of 50 possible names given selected race and gender
Education Information	
GPA display condition	90% displayed, 10% missing
Revealed GPA	Drawn from $Unif[3.00, 4.00]$ to second decimal place
Major	Drawn from a list of majors at Penn
Degree type	BA, BS fixed to randomly drawn major
School within university	Fixed to randomly drawn major
Graduation date	Fixed to upcoming spring (i.e., May 2021)
Work Experience	
First job	Drawn from curated list of top internships and regular internships
Title and employer	Fixed to randomly drawn job
Location	Fixed to randomly drawn job
Description	Bullet points fixed to randomly drawn job
Dates	Summer after candidate's sophomore year (i.e., 2019)
Second job	Left blank or drawn from curated list of regular internships and work-for-money jobs
Title and employer	Fixed to randomly drawn job
Location	Fixed to randomly drawn job
Description	Bullet points fixed to randomly drawn job
Dates	Summer after candidate's junior year (i.e., 2020)
SAT Scores	
Display condition	39% displayed, 61% missing
Math Score	50% drawn from $Unif[700, 790]$, 50% equal to 800
Reading Score	50% drawn from $Unif[680, 760]$, 50% drawn from $Unif[770, 800]$
Writing Score	50% drawn from $Unif[690, 780]$, 25% equal to 790, 25% equal to 800
Leadership Experience	
First & second leadership	Drawn from curated list
Title and activity	Fixed to randomly drawn leadership
Location	Fixed to Philadelphia, PA
Description	Bullet points fixed to randomly drawn leadership
Dates	Start and end years randomized within college career, with more recent experience coming first
Language Skills	
Display condition	75% displayed, 25% missing
Displayed skills	Drawn from six combinations between English and one foreign language (Mandarin, Spanish, French)

This table lists resume components in the order that they appear on hypothetical resumes.

men relative to women.[31]

Second, we demonstrate the impact of concealing GPA, compared to listing the lowest GPA in our study, a 3.00. Panel B shows that male candidates with hidden GPA instead of a GPA of 3.00 are 16 percentage points more likely to receive an interview—an effect that is marginally statistically significant. By contrast, female candidates with hidden GPA have the same interview probability as those listing a 3.00, suggesting no benefit of GPA concealment for women.

Figure D.1: Employer response to GPA and hidden GPA



Panel A shows the effect of GPA on candidates' chance of receiving an interview, estimated with a linear probability model using the 1,401 resume observations that list GPA between 3.00 and 4.00 (see Table D.3, Column 1). Panel B shows the estimated effects of hidden GPA, compared to listing a 3.00 on resume, separately for male and female candidates. The p-value of the gender difference is 0.083 in the specification shown, or 0.047 when we exclude two hiring managers with less than one year experience evaluating candidates. The estimates are derived from linear probability models that include all 1,560 observations. Hidden GPA is replaced with a value of 3.00 in the regressions, so that we can simultaneously estimate the effect of GPA levels and the effect of concealed GPA (see Table D.3, Columns 2–5).

Given that men and women differentially conceal information, one might also wonder whether this response to concealed GPA is rational inference on the part of employers. If fewer women conceal, employers might expect them to be concealing worse grades, and therefore rationally penalize them more. We can use our IRR data to calculate the "equivalent GPA" for resumes without GPA information; that is, the GPA level of candidates rated

[31] As shown in Table D.3, Columns 2–5, female candidates receive a lower return to GPA than male candidates, which may further enlarge the differential benefits of the concealment policy for men. This result is consistent with a body of evidence showing that marginalized groups often receive lower returns to quality, conversely to what one would expect as a result of statistical discrimination (**?**Kessler, Low and Sullivan, 2019; Kessler, Low and Shan, 2023).

equally likely to receive an interview as those with concealed GPAs. In an unraveling model (Grossman, 1981; Milgrom, 1981), those who conceal would be expected by employers to have lower GPAs than all those who reveal, so fewer and fewer individuals would end up concealing. For women, the equivalent GPA for resumes without GPA information is 2.91, below the distribution of revealed GPAs in the IRR experiment. However, for men, the equivalent GPA for resumes without GPA information is 3.51—at the midpoint of revealed GPAs.[32]

That women are treated worse than men when their resumes lack GPA information, and that men are actually treated as though their GPAs fall in the middle of the distribution, appears hard to reconcile with rational behavior by employers. However, such a result is consistent with a large body of evidence in which lack of information or ambiguity can allow a greater influence of subjective beliefs or provide "moral cover" for undesirable behaviors, such as favoritism toward an advantaged group (e.g., see Dana, Weber and Kuang (2007) and more recently Chan (2022) in the context of discrimination).[33]

[32]We emphasize that the type of performance information and context certainly affect whether the disclosure of performance information benefits men and women relatively more (e.g., Agan, Cowgill and Gee (2023*a*) finds that men benefit relatively more from salary disclosures).

[33]For evidence on how individuals may more generally use ambiguity, uncertainty, or subjectivity to justify undesirable behavior, see also Snyder et al. (1979); Kunda (1990); Haisley and Weber (2010); Di Tella et al. (2015); Exley (2016). Individuals may even act as if they are confused when processing information to justify undesirable behavior (Exley and Kessler, Accepted). Kessler, Low and Shan (2023) shows evidence that individuals may look for internal justification for bias, by doing so only when the preferred group has high quality, as favoring a low quality candidate would make the discrimination obvious.

Table D.3: Gender gap in the impact of GPA information on interview likelihood

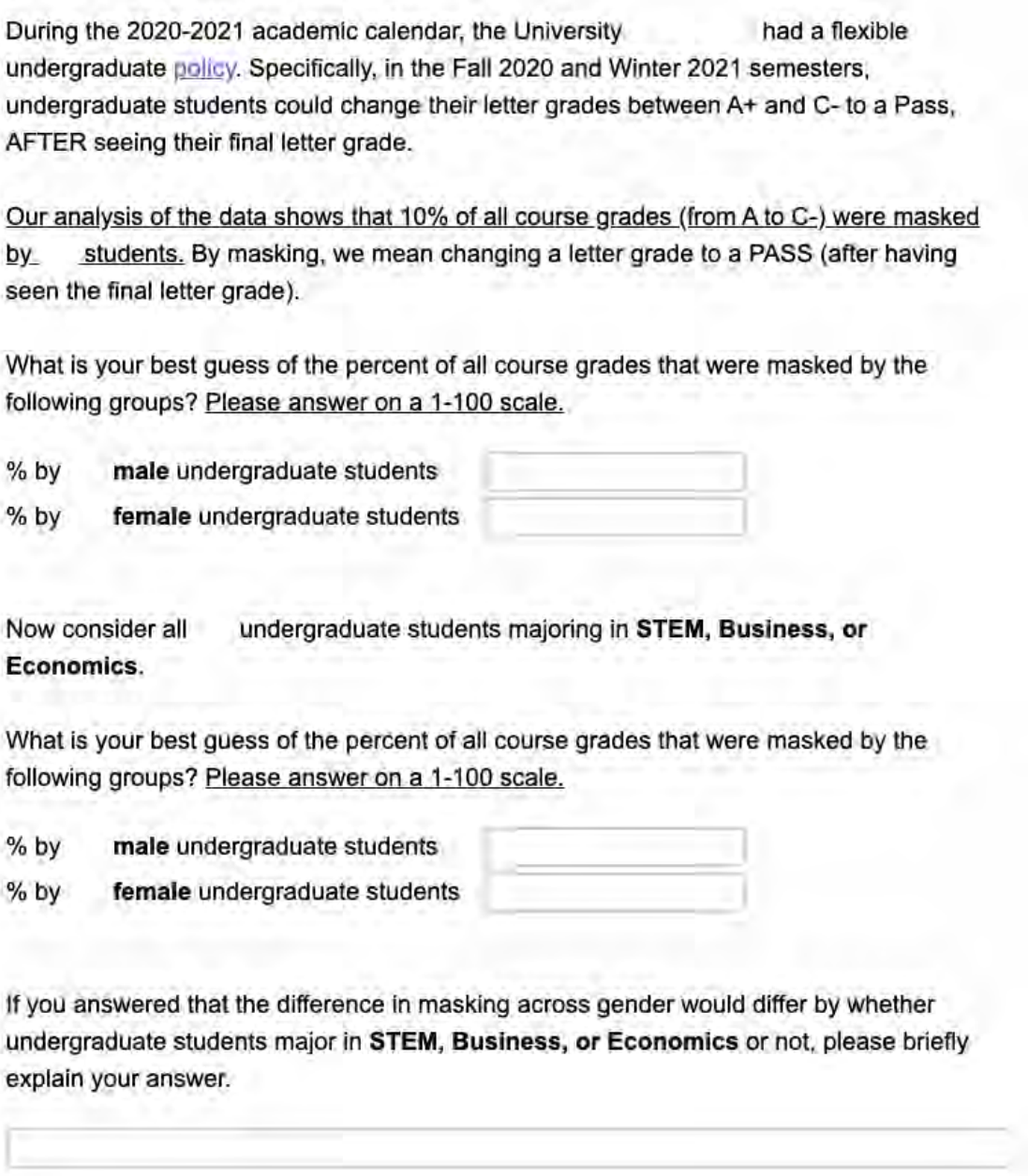
	(1)	(2)	(3)	(4)	(5)
	Dependent variable: Employer would interview candidate				
	All	All	Experience≥1	All	Experience≥1
GPA	0.182**				
	(0.075)				
GPA × Male		0.280***	0.333***	0.286***	0.340***
		(0.094)	(0.092)	(0.094)	(0.093)
GPA × Female		0.088	0.116	0.086	0.117
		(0.082)	(0.082)	(0.080)	(0.080)
GPA Concealed × Male		0.156*	0.186**	0.159*	0.190**
		(0.084)	(0.085)	(0.086)	(0.087)
GPA Concealed × Female		-0.008	-0.004	-0.003	0.004
		(0.065)	(0.065)	(0.063)	(0.063)
Female	-0.002	0.668**	0.748**	0.694**	0.766**
	(0.029)	(0.311)	(0.317)	(0.295)	(0.306)
Resume work exp. controls	Yes	Yes	Yes	Yes	Yes
Major category fixed effects	Yes	Yes	Yes	Yes	Yes
Lasso controls	No	No	No	Yes	Yes
Observations	1,401	1,560	1,480	1,560	1,480
p-value: gender diff. in *GPA*		0.017	0.003	0.008	0.001
p-value: gender diff. in *GPA Concealed*		0.083	0.047	0.077	0.045

This table estimates the effects of the GPA level and concealing GPA on the interview likelihood. The outcome variable is a binary indicator for a candidate receiving an interview. Column 1 focuses on the 1,401 resumes that explicitly reveal a GPA between 3.00 and 4.00; Columns 2–5 also include resumes with no GPA information. When included, resumes without GPA are replaced to have a GPA of 3.00, so that the coefficients *GPA Concealed* (interacted with gender indicators) estimate the difference in interview likelihood between male/female candidates with hidden GPA and male/female candidates with a GPA of 3.00. Columns 3 and 5 focus on employers who have had at least one year of working experience at their institution. The last two rows show the p-values derived from tests of the gender differences in the effects of GPA level and GPA concealment. All regressions use a linear probability model and control for *Top Internship* (having an internship at a prestigious company like Google and McKinsey), *Work-for-Money Job* (having a paid summer job such as a waiter or a cashier), *Second Job* (having a second regular internship), and indicators for major categories. Columns 4 and 5 employ double-lasso to select additional control variables from the rich set of resume characteristics: dummies for employer subject, a student's major, first and second student leadership experience, resume review order, SAT writing, reading, and math scores, whether SAT scores are missing, and white name. See Table D.2 for details of various resume components. We cluster standard errors at the employer level. ***, **, * denote that estimates are statistically significant at the 1, 5, and 10% levels, respectively.

E The expert survey

At the beginning of the survey, each participant is informed that their participation is completely voluntary. Figures E.1 and E.2 illustrate the main questions concerning perceived gender differences in grade concealment. Finally, a brief demographic survey about gender and primary field of expertise was administered at the end.

Figure E.1: Questions on perceived gender differences in grade concealment


During the 2020-2021 academic calendar, the University had a flexible undergraduate policy. Specifically, in the Fall 2020 and Winter 2021 semesters, undergraduate students could change their letter grades between A+ and C- to a Pass, AFTER seeing their final letter grade.

Our analysis of the data shows that 10% of all course grades (from A to C-) were masked by students. By masking, we mean changing a letter grade to a PASS (after having seen the final letter grade).

What is your best guess of the percent of all course grades that were masked by the following groups? Please answer on a 1-100 scale.

% by **male** undergraduate students

% by **female** undergraduate students

Now consider all undergraduate students majoring in **STEM, Business, or Economics**.

What is your best guess of the percent of all course grades that were masked by the following groups? Please answer on a 1-100 scale.

% by **male** undergraduate students

% by **female** undergraduate students

If you answered that the difference in masking across gender would differ by whether undergraduate students major in **STEM, Business, or Economics** or not, please briefly explain your answer.

Figure E.2: Questions on perceived gender differences in grade concealment, conditional on same GPA and grade