NBER WORKING PAPER SERIES

INFORMATION-OPTIONAL POLICIES AND 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 October 2024

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, or the National Institutes of Health. This paper supersedes material previously included in "The Transparency Gap" and "Anticipated Discrimination and Grade Disclosure". The views expressed herein are those of the authors and do not necessarily reflect the 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

Information-Optional Policies and 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 October 2024
JEL No. D82,J16,J71

ABSTRACT

We analyze data from two universities that allowed students to replace a letter grade with "credit" on their transcript. At both schools, we observe a significant and substantial gender concealment gap: women are less likely than men to conceal grades, particularly grades that would harm their GPA. This gender concealment gap produces differential GPA gains from the policy with men benefiting nearly 50% more than women. Additional complementary data, including surveys and experiments with students and employers, suggest why women may conceal less: women may expect observers to have more negative inferences about their concealed grades.

Christine L. Exley
Department of Economics
University of Michigan
Lorch Hall 365B
611 Tappan Ave
Ann Arbor, MI 48109
clexley@gmail.com

Raymond Fisman
Department of Economics
Boston University
270 Bay State Road, 304A
Boston, MA 02215
and NBER
rfisman@bu.edu

Judd B. Kessler
The Wharton School
University of Pennsylvania
320 Vance Hall
Philadelphia, PA 19104
and NBER
judd.kessler@wharton.upenn.edu

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

Xiaomeng Li
Department of Economics
611 Tappan St.
Ann Arbor
lixiaom@umich.edu

Corinne Low
The Wharton School
University of Pennsylvania
318 Vance Hall
Philadelphia, PA 19104
and NBER
corlow@wharton.upenn.edu

Xiaoyue Shan
1 Arts Link
AS2, #05-36
Singapore
x.shan@nus.edu.sg

Mattie Toma
Warwick Business School
University of Warwick
Scarman Road
Coventry CV4 7AL
Mattie.Toma@wbs.ac.uk

Basit Zafar
Department of Economics
University of Michigan
611 Tappan Street
Ann Arbor, Michigan 48109
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 found is influenced by the amount of information available to employers.[2]

In this paper, we present evidence on how choices made by prospective workers—in response to information-optional policies—contribute to differences in information available to employers. We document a *gender concealment gap*: women are less likely than men to conceal relatively poor academic performance when given the opportunity to do so.

Our main sources of data are two large-scale natural experiments at highly selective universities, Boston University and a flagship public school in the Midwest, which both introduced grade-optional policies in response to the Covid-19 pandemic. Students at these schools had the option of choosing—for each of their classes—whether to conceal a passing letter grade by choosing to have it officially recorded as "Credit."[3] When students concealed a passing letter grade in this way, the grade appeared as "Credit" on their transcript and no longer impacted their grade point average (GPA).

A key component of the policies at both universities is that students observed the assigned letter grade they received for each class *before* deciding whether to conceal it. We study responses to this policy for the one semester it was in place at Boston University (BU) and for the two semesters it was in place at the large public Midwestern university (hereafter "Midwestern Flagship"). Both schools provided us with information on assigned letter grades and whether each grade was converted to Credit, as well as additional infor-

[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).

[3] These policies emerged unexpectedly in 2020, in recognition of the challenges faced by many students when Covid brought unprecedented disruption to their educations as a result of mid-semester campus shutdowns and the shift to online instruction.

mation on students. Thus, one strength of our setting is that, unlike in many situations in which students may avoid taking a test or submitting a test score altogether, we observe the performance outcome of students even when they decide to conceal their performance.

We find that students hardly ever conceal grades that would increase their GPA or leave it unchanged. In contrast, 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 observe a large and robust gender gap in concealment. 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. Within each school, we find that the concealment gap is present across a range of student-level traits, types of 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); 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 their difficulty, demand for seats, and their gender balance. The robustness of the concealment gap across these various dimensions and across the two schools—with different student populations and different implementations of the policy—suggests its potentially wide relevance to other settings.

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.[4] 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 qualification for awards, internships, majors, or jobs; further career and educational decisions; and others' perceptions about students' abilities.

We then consider the possible causes of the gender concealment gap. Features of the programs and our rich administrative data allow us to rule out many of the usual explanations for gender gaps that arise in other settings. For example, the gender concealment gap cannot be driven by gender differences in confidence or risk aversion about course performance or eventual course grades, since students learn their course grade before deciding whether or not

[4]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.

to conceal it. The gap is unlikely to be driven by strategic concerns about how a course grade will compare to eventual collegiate GPA, since the gap arises for students in all academic years, including seniors, and arises even when grades are substantially below a student's current GPA. The gap is unlikely to be driven by gender differences in understanding how concealment affects one's GPA, since it appears across various ranges of GPA that include relatively high and low performers. The gap is unlikely to be driven by inattention, since we see a gender concealment gap in the number of courses concealed even among students who are clearly aware of the policy because they conceal at least one grade.

Having ruled out a number of possible stories for the gender concealment gap, we then turn to investigate one possible explanation that could still contribute to the concealment gap. This possible explanation relates to students' beliefs about the inferences that others will draw about them as a result of concealing performance information. If female students believe that employers and admissions committees will draw an overly negative inference about their course performance if a grade is concealed, they may choose to reveal the grade, even if doing so will pull down their GPA. If men believe that a concealed grade will not be viewed that negatively—that they will receive the "benefit of the doubt" about their performance—they may choose to conceal the harmful grade, helping to boost their GPA.

To look for evidence of this explanation, we run a student belief study and an employer study described in Section 3.7. Related to the nascent but growing body of work on anticipated discrimination based on gender (Alston, 2019; Dustan, Koutout and Leo, 2022; Gagnon, Bosmans and Riedl, 2022; Koutout, 2022; Ruebeck, 2023), we find that students at the Midwestern Flagship indeed believe that employers and admissions committees will make worse inferences about women who conceal a grade than about men who do so. Data from an additional sample of adult subjects reveals that such beliefs are also held by the population at large.

While student beliefs about observer inferences could drive the gender concealment gap regardless of whether those beliefs are accurate, we also investigate whether employers indeed respond differentially when men and women conceal information about their performance. To do so, we run a contemporaneous incentivized resume rating experiment—following Kessler, Low and Sullivan (2019)—involving actual employers hiring college seniors in the fall of 2020. In a similar paradigm, which involves concealing GPA from a resume rather than a specific grade from a transcript, we find that employers make worse inferences about female applicants than male applicants when information is concealed. This experiment highlights that students have good reason to believe that employers may respond differentially by gender when performance information is unavailable.

To summarize, the main contribution of this paper is to document a robust gender gap in

grade concealment at two large, selective universities. We find that women are substantially less likely to conceal grades below their GPA when given access to an information-optional policy, a relationship that is observed across a range of student traits and course-level characteristics. This gap meaningfully lowers GPAs for women relative to men.

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. For instance, women effectively conceal their performance more often 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), and applying for challenging work less often (Coffman, Collis and Kulkarni, Forthcoming).[5] One potential reason for this difference could be that a willingness to reveal performance information could be reversed when considering clearer "failures" (e.g., while women may be more reluctant to speak up about potential successes, women may be more willing to speak up about clearer failures such as grades that fall below their GPA). In addition, as is shown in prior gender research, the degree of ambiguity in a context is known to influence gender gaps. 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 in which women are choosing whether to negotiate, compete, speak-up, apply for a job, or provide other information such as other performance or salary information in applications and negotiations.[6]

Our results also speak to the unintended consequences that information-optional policies may have. In our context, because men are more likely to conceal harmful grades, the policy shifts GPA distributions in favor of men, which might exacerbate existing 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.[7]

[5]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.

[6]For instance, consistent with the importance of ambiguity in these decisions, Agan, Cowgill and Gee (2020) show 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).

[7]Agan, Cowgill and Gee (2020) provide 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,

This discussion is growing in importance as many universities have considered widening information-optional policies (e.g., test-optional admissions, broader pass-fail policies) as part of their efforts to increase equity.[8] Our findings also suggest that it will be valuable to further explore the presence and implications of concealment gaps in other settings involving, for example, decisions on what is included on a resume or job application, and whether to report negative information to a supervisor.

2 The concealment gap

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 and any grade from A to C- at the Midwestern Flagship[9]) by converting it into a grade of "Credit" (at BU) or "Pass" (at the Midwestern Flagship).[10] If a student chooses to

Hirshman and Willén (2022) find 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.

[8]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) find 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).

[9]At the Midwestern Flagship, there are two particular 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.

[10]Our analyses exclude the 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."

conceal a letter grade in a course, the grade of "Credit" 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 still 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.[11]

The policy was implemented in the Spring semester of 2020 at BU and the Fall semester of 2020 at the Midwestern Flagship. In each case, students were informed about the policy unexpectedly during these semesters. Prior to the Spring semester of 2021, the Midwestern Flagship announced that the policy would continue that semester. We always analyze the data separately for BU and the Midwestern Flagship, and we also note that our results are robust to only examining the Fall 2020 semester at the Midwestern Flagship or only examining the Spring 2021 semester at the Midwestern Flagship, confirming that the concealment gap persists when the policy is a surprise *and* when it is known in advance.[12]

2.1 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 decision 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

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 then 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.

[11] Grading policies where students do not know their letter grade before deciding whether or not to conceal it have also been explored, and the results from these analyses suggest that the timing of the concealment decision could have important implications for student behavior (Kolb et al., 2023; Trost and Wooten, 2023).

[12] That the concealment gap arises both when the policy is and is not a surprise shows that the concealment gap, in addition to arising when the policy cannot affect class selection, is also robust to any potential impact it has on class selection (although we observe no meaningful differences by gender in class selection at the Midwestern Flagship in the Spring 2021 semester). To see the robustness of the concealment gap when it is a surprise and not, see Appendix Table A.12. Recall also that the policy is a surprise at BU.

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.

2.2 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]

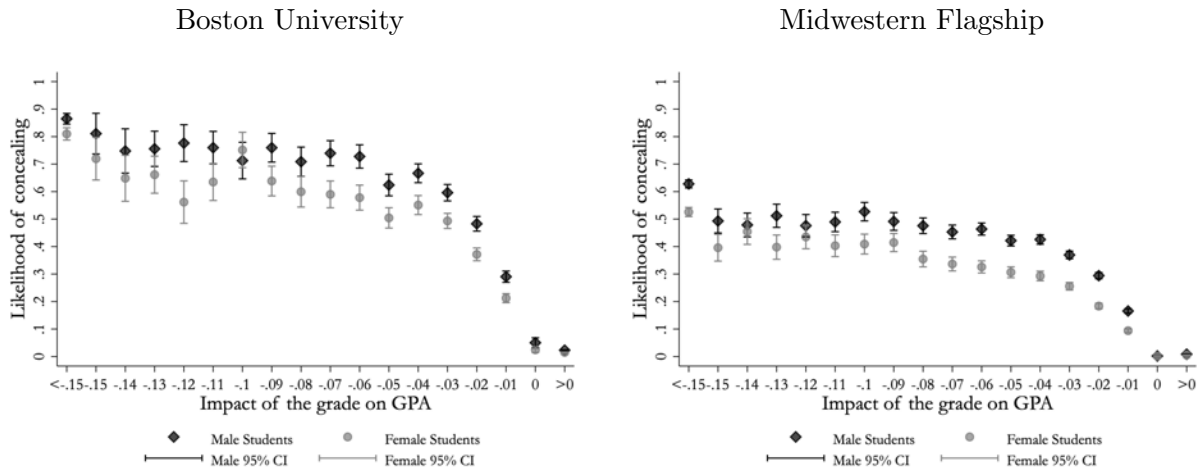
We begin by examining how students make their concealment decisions as a function of GPA impact, and whether those decisions differ by gender. Figure 1 shows the rates at which male and female students conceal their letter grades at Boston University (in the left panel) and at the Midwestern Flagship (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 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 is little evidence for any gender differential.

By contrast, evidence for gender differences in concealment rates is substantial for harmful grades, which are much more likely to be concealed. That is, 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.

[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.

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 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 those at BU. While there are many reasons as to why the concealment rates may differ across these institutions—e.g., relating to differences in the student populations or how the grade-optional policies were announced—one reason that could contribute to

the Midwestern Flagship's lower concealment rate relates to the higher cumulative GPAs of their students (see Appendix Table A.1). Despite the level difference in concealment rates across schools, however, the *gender gap* in concealment—the focus of this paper—is remarkably robust both across schools and across the various impacts that grades may have on GPA at each school. As discussed later, the gender gap in concealment is also robust to a range of features, including 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×ΔGPA 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 program year. Allowing these fixed effects to vary by program 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 are robust to including GPA fixed effects but not allowing them to vary by program 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).[14] 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 when focusing on 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 also evident from the concealment rates observed in Figure 1, pooling across all grades masks substantial and important heterogeneity. While students rarely conceal neutral or helpful grades and hence the potential to observe a concealment gap is minimal—and indeed only a small concealment gap arises in these cases (see Column 3 and 4)—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.

[14]Building off of 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 in Section 2.4.

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. 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.

2.3 Impact of the concealment gap

At both schools, students use the grade-optional policies to conceal grades below their GPA. These policies make it possible for students to end up with substantially higher GPAs than they would have earned if they had to keep letter grades for all of their classes. However, students also frequently choose *not* to conceal harmful grade and sometimes—albeit rarely—conceal helpful grades. In this section, we investigate whether the grade-optional policies significantly increase students' GPAs and whether, given the gender concealment gap, there are significant gender differences in how these policies affect students' GPAs as a result.

Table 2 reports the GPA increase that students receive due to the grade-optional policies 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.1 for a graphical representation of this shift in GPA distributions).

In summary, as one may expect given that women are less likely to take advantage of the grade-optional policy in a way that boosts their GPAs (i.e., women are less likely to conceal harmful grades), these policies increase the GPAs of men relative to the GPAs of women.

Table 2: 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, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

2.4 The robustness of the concealment gap

We conduct a series of analyses to show that the gender concealment gap of harmful grades is remarkably robust. These analyses also generate findings that help us better understand potential drivers of the concealment gap, as discussed in Section 3.

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 broad 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 3: 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***	-0.270***	-0.146***
	(0.010)	(0.021)	(0.023)
Constant	0.616***	1.111***	1.746***
	(0.007)	(0.017)	(0.017)
Observations	10,576	10,163	5,838
Conceal mean	0.552	0.960	1.671
Panel B: Midwestern Flagship			
	(1)	(2)	(3)
Dep. variable:	Ever concealed	# concealed	# concealed | >0
Female	-0.113***	-0.255***	-0.186***
	(0.006)	(0.011)	(0.015)
Constant	0.449***	0.750***	1.649***
	(0.004)	(0.009)	(0.011)
Observations	37,574	37,241	14,823
Conceal mean	0.395	0.626	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. 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.

The concealment gap is also 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 3, 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 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.

3 What could drive the gender concealment gap?

In Section 2 we documented a large and robust gender concealment gap and showed that it has important implications on the relative GPA distributions of men and women. We now discuss potential drivers of the concealment gap.

In Sections 3.1–3.6, we highlight explanations that are unlikely to be driving the concealment gap. We then turn, in Section 3.7, to results from additional data collection efforts that provide evidence in support of a remaining channel, namely that the concealment gap may arise from gender differences in beliefs about how concealed grades will be perceived by others, such as employers or admissions committees.

3.1 Could gender differences in confidence or risk aversion about grades drive the concealment gap?

Suppose students had to make their concealment decisions *before* learning their grade. 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 actually 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 these authors note—be reflective of gender differences in confidence or risk aversion concerns about what grade they will receive.

In our setting, however, we re-emphasize that 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, mitigating the relevance of gender differences in confidence or risk aversion about the grade itself for concealment decisions. This may contribute to why we instead observe that women are *less* likely to conceal their grades.

3.2 Could gender differences in concerns about future grades drive the concealment gap?

One may wonder whether gender differences in beliefs about future grades contribute to the concealment gap, because, for example, women 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 this possibility—and more generally to address any concerns related to future GPAs driving the concealment gap—we examine whether the concealment gap persists across first-year, sophomore, junior, and senior students. Appendix Table A.8 shows that the concealment gap is present and statistically significant in all 8 out of the resulting 8 cases. That is, the concealment gap persists even among seniors in college, when uncertainty about future GPAs is mostly or entirely resolved.

In addition, recalling Figure 1, we see that the concealment gap arises for grades well below 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 over future GPA).

3.3 Could differential awareness or inattention drive the concealment gap?

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.[15] But features of the setting and results discussed above argue against this possibility.

In terms of the setting, the grading policies were prominently announced and advertised at each institution. In addition, 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 3).[16]

[15] 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, it would push against observing the gender concealment gap that we document in our main results.

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

3.4 Could gender differences in strategic behavior about grades drive the concealment gap?

To maximize their GPAs, students need to recognize the benefits of concealing harmful grades. One may speculate that students with high GPAs are more motivated by this goal and perhaps better able to achieve it. One may then wonder if the concealment gap persists even among students with high GPAs. 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.

Both men and women also appear quite sophisticated in terms of when they conceal a grade. In addition to our main results showing that students rarely conceal a helpful grade and frequently conceal harmful grades (recall Figure 1), we observe a similar pattern when we look at concealment rates by grades. Appendix Table A.3 presents our results across letter grades that could be changed to a grade of "Credit" if concealed.[17] 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. The only time we observe evidence inconsistent with the gender concealment gap is for the grade of D at BU. At BU, a D is the worst letter grade that could earn a grade of "Credit." This is, arguably, the only case where it is unambiguous that a student should choose a grade of "Credit" (since doing so at least weakly increases GPA as well as inferences about performance in the course, a topic we return to in Section 3.7). This evidence thus supports women, even women who get a low grade of D in a class, being at least as "strategic" as men when making concealment decisions.[18]

Finally, we also observe the gender concealment gap among students with different types of training, including in quantitative fields, as proxied by their majors. We classify majors into six different categories (the particular majors differ across the universities, but both sets can be classified within this broader framework). Appendix Table A.9 reveals significant evidence for the concealment gap at both universities in Health and Medicine majors, STEM majors, and among students who are undecided. At the Midwestern Flagship, with the much larger sample size of students, evidence for the concealment gap is observed for all major

[17] 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 10.

[18] A similar logic does not apply to Ds at the Midwestern Flagship, where neither Ds nor Fs could be recorded as Credit; both Ds and Fs were defaulted to "No Record Covid" as detailed in Footnote 10.

categories.

3.5 Could gender differences in preferences 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).[19] Could such gender differences—or gender differences in preferences for transparency per se—drive the gender concealment gap?

We first note that lying is conceptually distinct from concealing grades in the context of the grade-optional policies. In our settings, given how the policies were advertised and implemented, choosing to conceal grades is neither a lie by omission nor by commission.

In addition, we do not see a concealment gap for relatively poor letter grades such as "C-" and we see a reverse concealment gap for grades of "D" at BU—grades for which a pure preference for transparency argument would apply.[20]

3.6 Could particular features of courses drive the concealment gap?

One could speculate—for any number of reasons—that the concealment gap is driven by students in particular courses. For example, the gender composition of a class could affect the perceived social norms or gender norms in the class, which could lead to differential concealment rates. Alternatively, the difficulty of a course could affect whether students are more or less disappointed in a particular letter grade, which could also differ by gender.[21] What these explanations have in common is that the concealment gap should only arise—or be substantially larger—in courses with certain features.

Appendix Tables A.4–A.6 and A.10 show that the concealment gap is robustly stable across courses that vary in terms of the gender composition of the course, the number of students in the course, the demand for the course, the difficulty level of the course (i.e., the average and standard deviation of grades in a course), and whether the course falls within or outside a student's major. This robustness suggests that any complete explanation for the concealment gap must account for it arising across a wide variety of courses with different features.

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

[20] 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 does not change the information content of the grade but might appear more palatable on a transcript.

[21] For prior literature on how gender differences can depend on group composition, see 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).

3.7 Could beliefs about observer inferences drive the concealment gap?

When a student decides to conceal a harmful grade, they secure a higher GPA and their grade is recorded as "Credit" on their transcript. If women are more concerned about the inferences made by observers (e.g., employers, reference letter writers, graduate school admissions committees, etc.) when they conceal a grade, this could drive the gender concealment gap.

To investigate whether there is evidence of women expecting more negative inferences than men when grade information is concealed, we collected additional data on beliefs about observer inferences in our *student belief study* and additional data on observer inferences in our *employer study*. These studies are detailed in the following two subsections.

3.7.1 The student belief study

We recruited 407 rising juniors and seniors at the Midwestern Flagship to participate in the *student belief study*.[22] In addition to receiving a $10 Amazon gift card for completion, students may receive a $20 bonus payment as part of the study.

In the study, 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. In two questions, students are asked to make predictions about others' inferences when both students received a grade of "credit" or "pass" in the same relevant course. In one question, they are asked to make predictions about the employers' inferences 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 another question, they are asked to instead make predictions about graduate school admission committee inferences when both students apply to exactly the same graduate school program by selecting from three similar options (see Appendix Figure C.7). Beliefs about graduate school admission committees is of interest since many college graduates pursue graduate school and graduate school applications often request applicants to submit full transcripts.[23]

[22] 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.

[23] 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. While we are interested in beliefs about graduate school admissions committees, we note that we observe a large gender concealment

As shown in Figure 2, students believe employers and graduate school admissions committees will hold more negative inferences about a woman than a man with a grade of credit. While nearly half the students report that employers and graduate admissions committees will generate equal inferences about men and women (39–47%), the remainder are 2–2.5 times more likely to believe that men will be viewed more favorably than women (36–40%) than the opposite (17–21%).

Additional questions in the *student belief study* confirm that students also expect that women will be viewed less favorably in a different context, specifically, when information about their GPA is not provided on their resume (see Panel C of Figure 2, based on Appendix Figure C.5). We extend our investigation to this context for two reasons. First, students frequently make decisions about whether to provide their GPA on their resumes, and we hope that future work examines the drivers of such decisions and that our findings are informative to that future work. Second, a contemporaneous experiment that we ran, detailed in the next section, allows us to incentivize students' beliefs about employers' inferences when GPA information is not provided on a resume. As shown in Appendix Figure C.10, in response to a question with incentives for predicting employer inferences, students predict that employers will believe that a male student has a GPA of 3.34 but that a female student only has a GPA of 3.22 ($p < 0.01$).

While we are particularly interested in the beliefs of our student subjects at the Midwestern Flagship (since we documented the gender concealment gap among students there), we were also interested in whether these beliefs were held more broadly. We thus recruited 399 adults from Prolific to participate in the same study. As shown in Appendix C.4, we replicate the findings from the student sample in this broader subject pool, suggesting the potential relevance of these beliefs beyond our setting.[24]

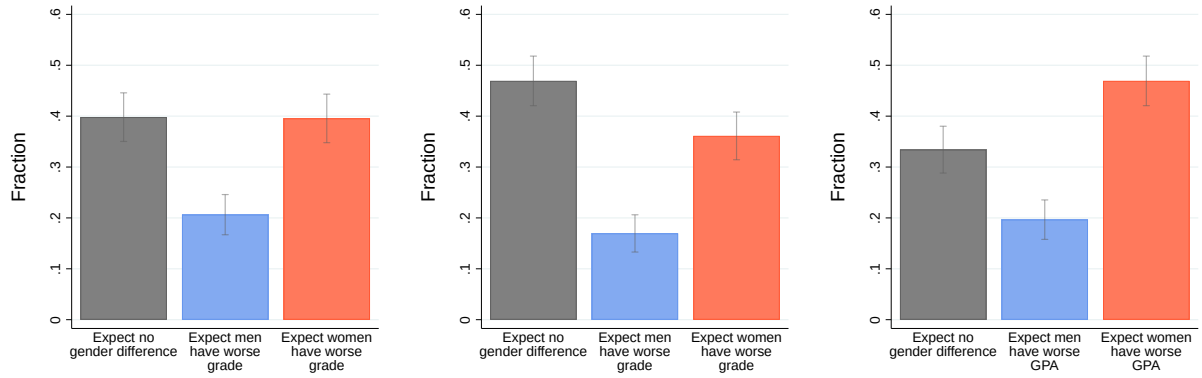
In summary, these studies show that students and the general public expect men—relative to women—to be evaluated more favorably by employers and admissions committees when specific information about their performance is unavailable. This belief is consistent with the concealment gap that we observe. Women may anticipate that employers and admissions committees will make worse inferences about them if they conceal their performance, so they avoid doing so. Men, in contrast, may expect relatively better inferences when they conceal such performance information.[25]

gap even when considering courses outside of a student's major (see Appendix Table A.10).

[24]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 ran this study in June 2024.

[25]These beliefs are also consistent with prior evidence that women are often viewed more negatively than men in response to the same performance information (Coffman, Exley and Niederle, 2021), particularly when

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.

3.7.2 The employer study

While a difference in beliefs about observer inferences—as shown in the prior section—would be sufficient to drive the gender concealment gap, one might wonder if employers indeed display the types of differential inference that students may worry about.

In order to investigate this, we analyze data from an Incentivized Resume Rating (IRR) experiment with employers, where GPA information was randomly omitted from resumes. 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, who 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

performance was poor (Sarsons, 2017). Our results are also consistent with—and contribute to—a nascent body of work on expected discrimination. That literature shows that women anticipate discrimination based on gender (Alston, 2019; Dustan, Koutout and Leo, 2022; Ruebeck, 2023), which may in turn impact incentives to apply for or exert effort at a job (Gagnon, Bosmans and Riedl, 2022; Koutout, 2022).

resume ratings.[26]

GPA on the resumes is randomly varied between 3.00 and 4.00, or omitted for 10% of resumes (approximating the real fraction of students who choose not to show GPA on their resumes). This study, in turn, provided the actual statistics with which to incentivize the student belief study.

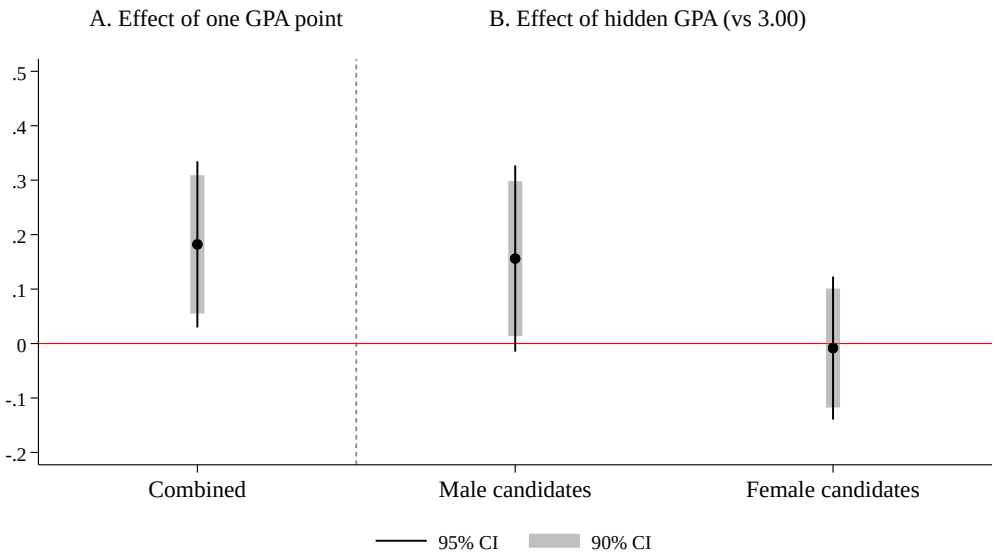
Figure 3 reports the main results of the study, using the binary outcome measure of whether an employer would invite a candidate for an interview. First, Panel A demonstrates substantial employment returns to GPA: candidates with a one-point higher GPA are 18 percentage points more likely to receive an interview. This return to GPA suggests that 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 men relative to women.[27]

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. Thus, consistent with student beliefs documented in Section 3.7, employers respond more negatively to concealed information from female candidates. Appendix Table D.3 reports the regressions these coefficients are drawn from, as well as showing robustness to LASSO controls, and that estimates become more precise when we exclude two hiring managers with less than one year of experience evaluating candidates.

[26] Additional details about implementation can be found in Appendix D. 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), shown in Appendix 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*).

[27] As shown in Appendix 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 (Bertrand and Mullainathan, 2004; Kessler, Low and Sullivan, 2019; Kessler, Low and Shan, 2023).

Figure 3: 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 Appendix 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 Appendix Table D.3, Columns 2–5).

4 Conclusion

In this paper, we document a robust gender gap in grade concealment at two large, selective universities. Women are substantially less likely to conceal grades below their GPA when given access to an information-optional policy, a relationship that is observed across a range of student traits and course-level characteristics. This gap has adverse consequences for gender equity. It results in lower GPA gains for women relative to men. In the IRR experiment, we find—perhaps unsurprisingly—that a higher GPA leads to a higher probability of a candidate being interviewed, suggesting that the concealment gap may have labor market consequences favoring men relative to women.

Considering the various possible causes of the gender concealment gap, our data rule out a number of possible explanations and, along with findings from the *student belief study*, lend support to the notion that female students may be more worried about the potential negative inferences of concealing grades than men are, and women may therefore choose to conceal grades less often. Additional findings from the IRR study suggest that women may be correct that they will face more particularly negative inferences when they hide

information about their performance.

Our findings open up several avenues for future work related to when the concealment gap arises and the potential implications of gender concealment gaps. On *when* gender gaps in concealment are likely to arise, it could be that women are *more reluctant* to reveal their *potential successes*—consistent with prior work on gender differences—but *less reluctant* to reveal their *known failures*. Future work may examine the presence and implications of concealment gaps in the choice to report negative information in other settings, such as in the workplace.

On the potential implications, building on our finding that grade-optional policies introduced gender equity concerns, future work may investigate whether information-optional policies more broadly—including undergraduate admission policies that allow students to selectively report information like SAT scores and job applications that allow individuals to selectively report information (e.g., via resumes)—could lead to differences in realized labor market outcomes or differences in graduate school admissions. Future work should also carefully consider the tradeoffs between optional information policies versus mandatory information policies. 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 (e.g., groups that are viewed more favorably in the absence of information). Institutions should carefully consider these equity implications when discussing information-optional policies, and those institutions with extant information-optional policies may wish to consider ways to mitigate the impact of the gender concealment gap we document.[28]

[28]Indeed, 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.

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).

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, 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. Forthcoming. "Whether to Apply?" *Management Science*.

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.

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*.

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. 2022. "Lowering the Playing Field: Discrimination through Sequential Spillover Effects." *Working Paper*.

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.

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."

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.

Sarsons, Heather. 2017. "Interpreting Signals in the Labor Market: Evidence from Medical Referrals." *Working Paper*.

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*.

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
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
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

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. 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. 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.

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

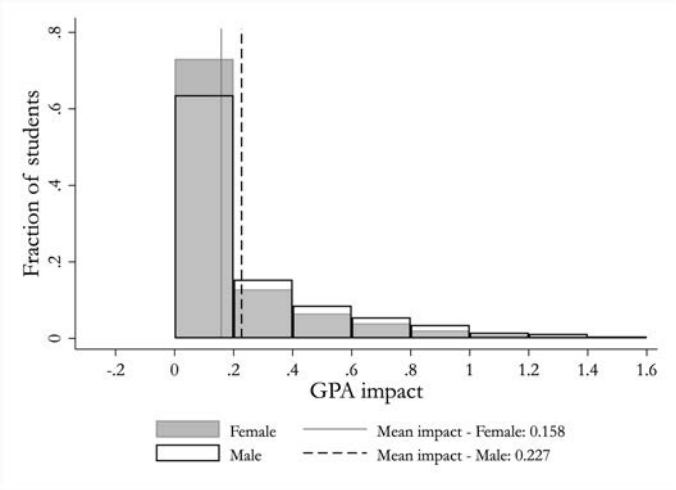
	(1)	(2)	(3)
	Fall 2020	Spring 2021	Spring 2021|concealed before
Female	-0.097***	-0.074***	-0.056***
	(0.007)	(0.005)	(0.010)
Observations	30,543	40,135	15,179
Conceal mean	0.351	0.312	0.495
Year × ΔGPA FEs	Yes	Yes	Yes
p-value		0.001	

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.

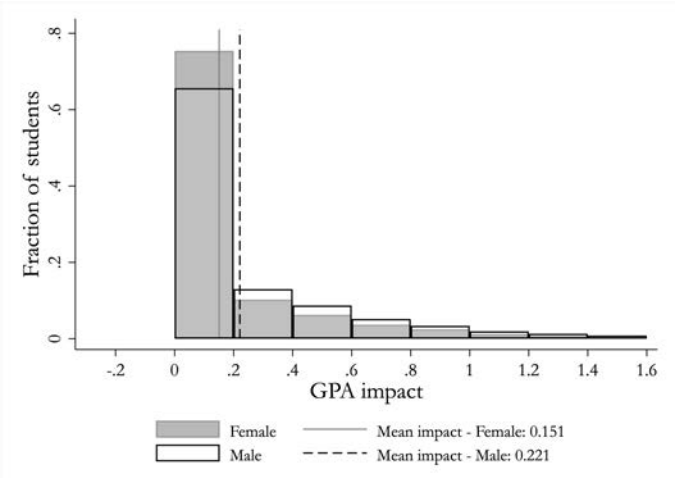
B Additional figures

Figure B.1: 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 Section 3.7.2 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 varies 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, a point we return to in our Conclusion).

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

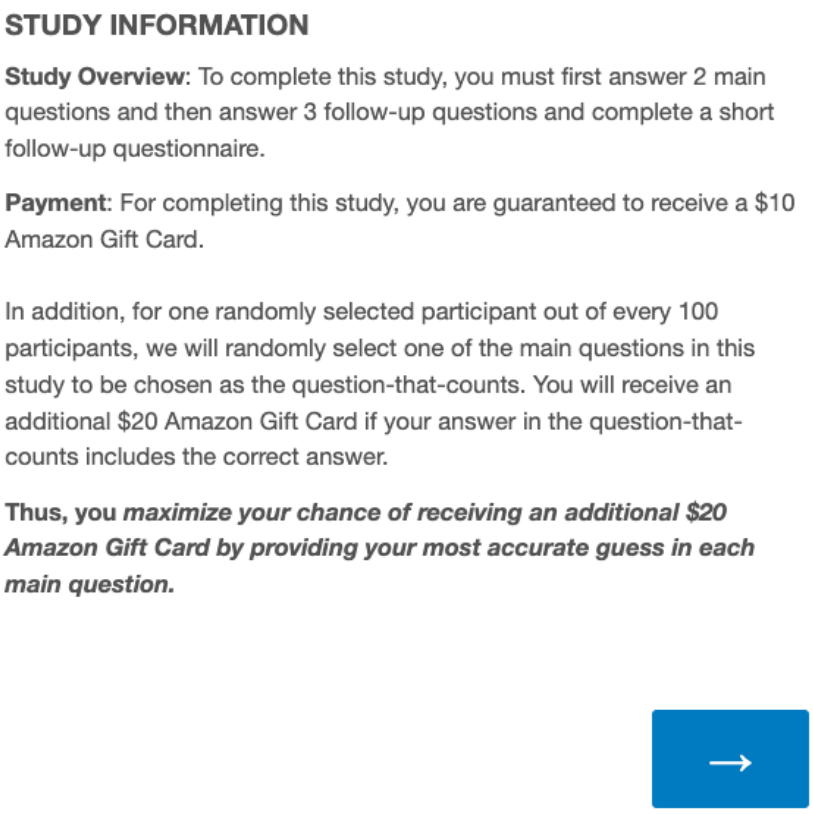
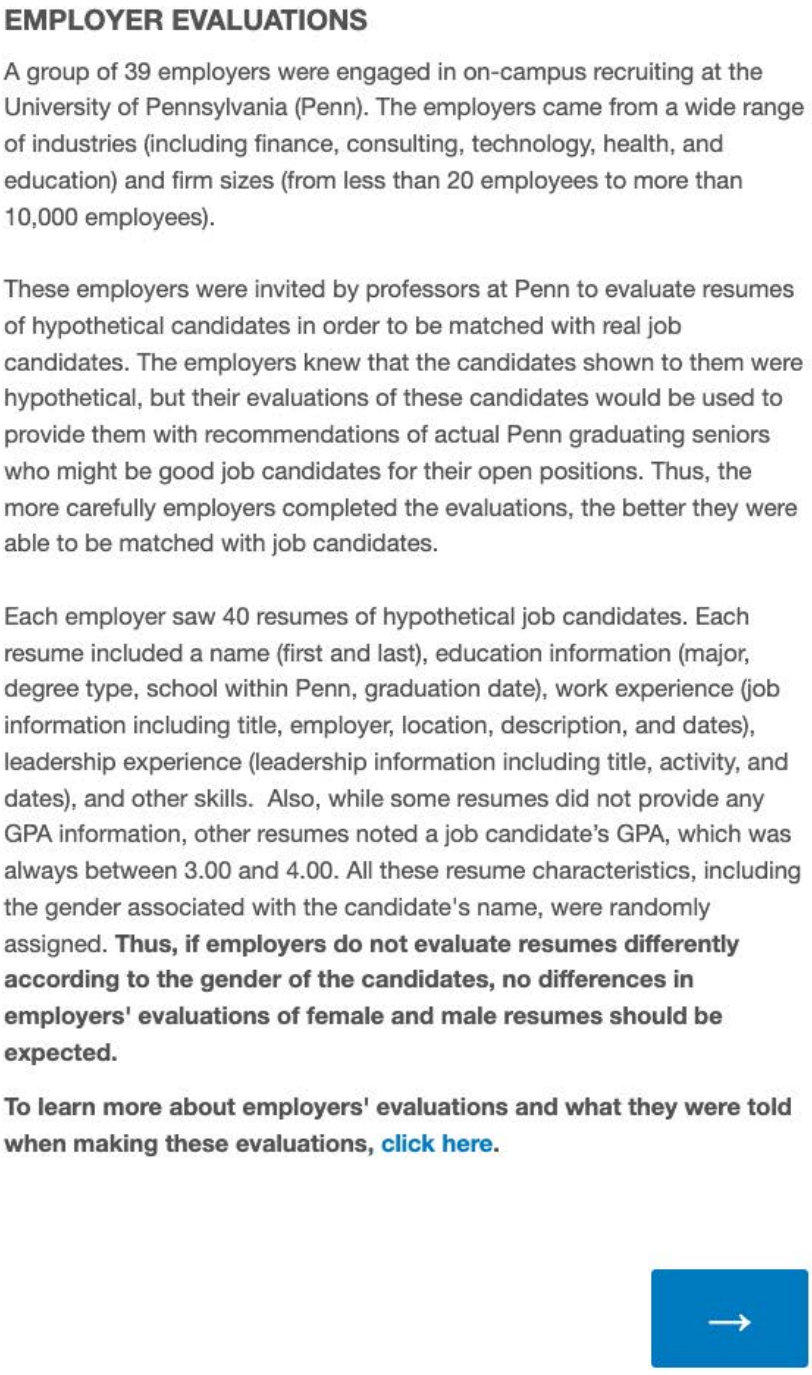


Figure C.2: Introduction about employer evaluations



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

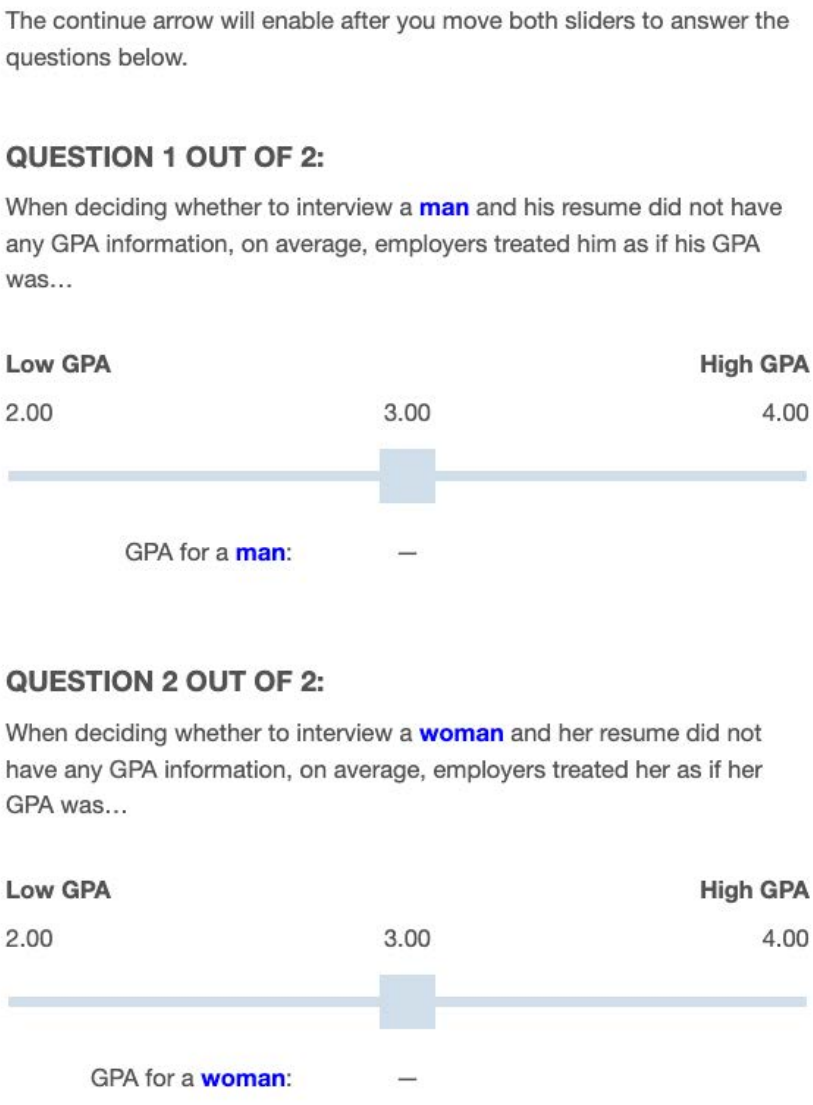


Figure C.4: Instruction for additional questions

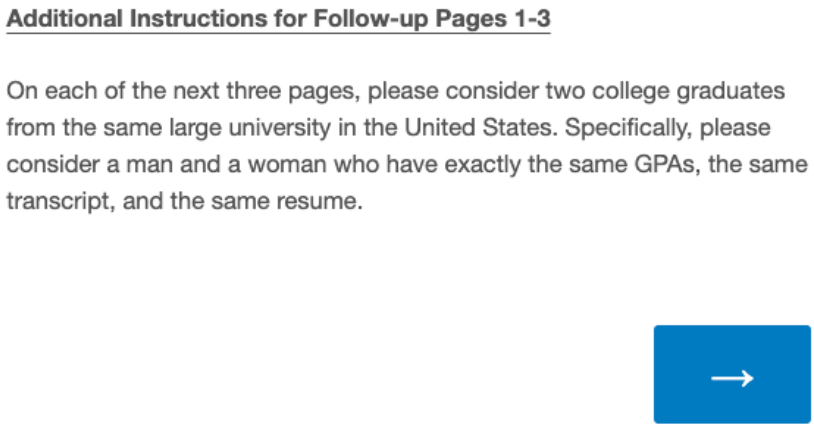


Figure C.5: Job application with no GPA

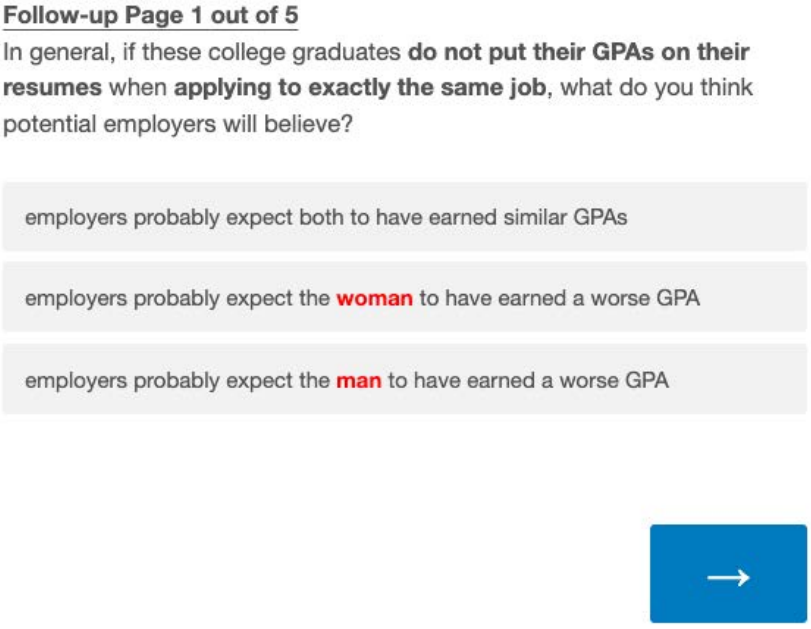


Figure C.6: Job application with a grade of "Credit"

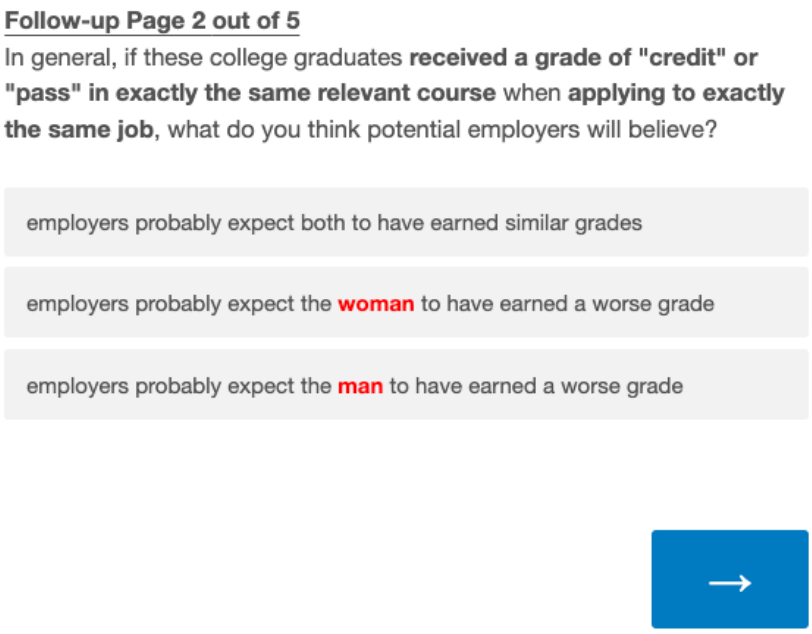


Figure C.7: Graduate school application with a grade of "Credit"

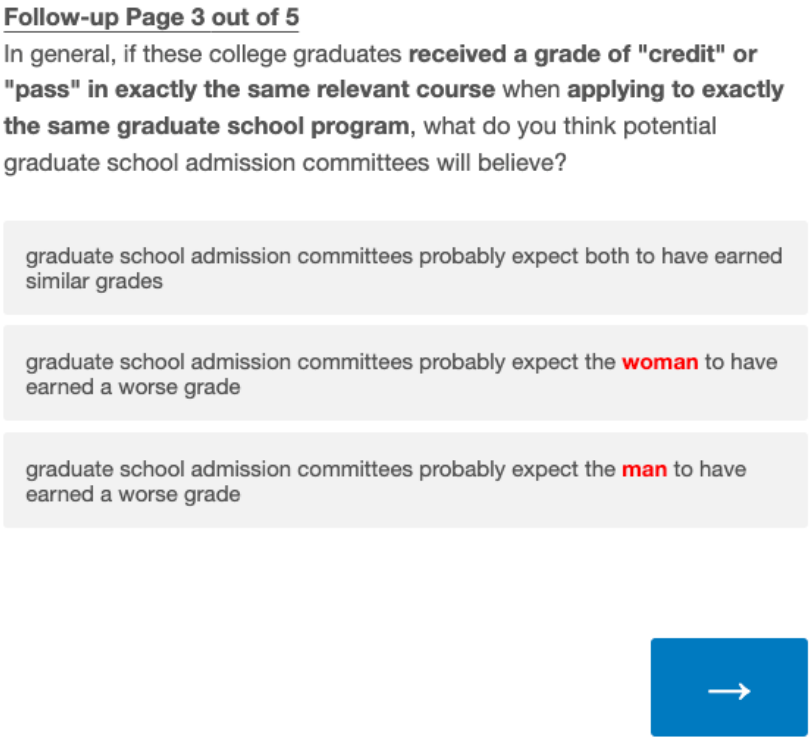


Figure C.8: Open question about grade concealment

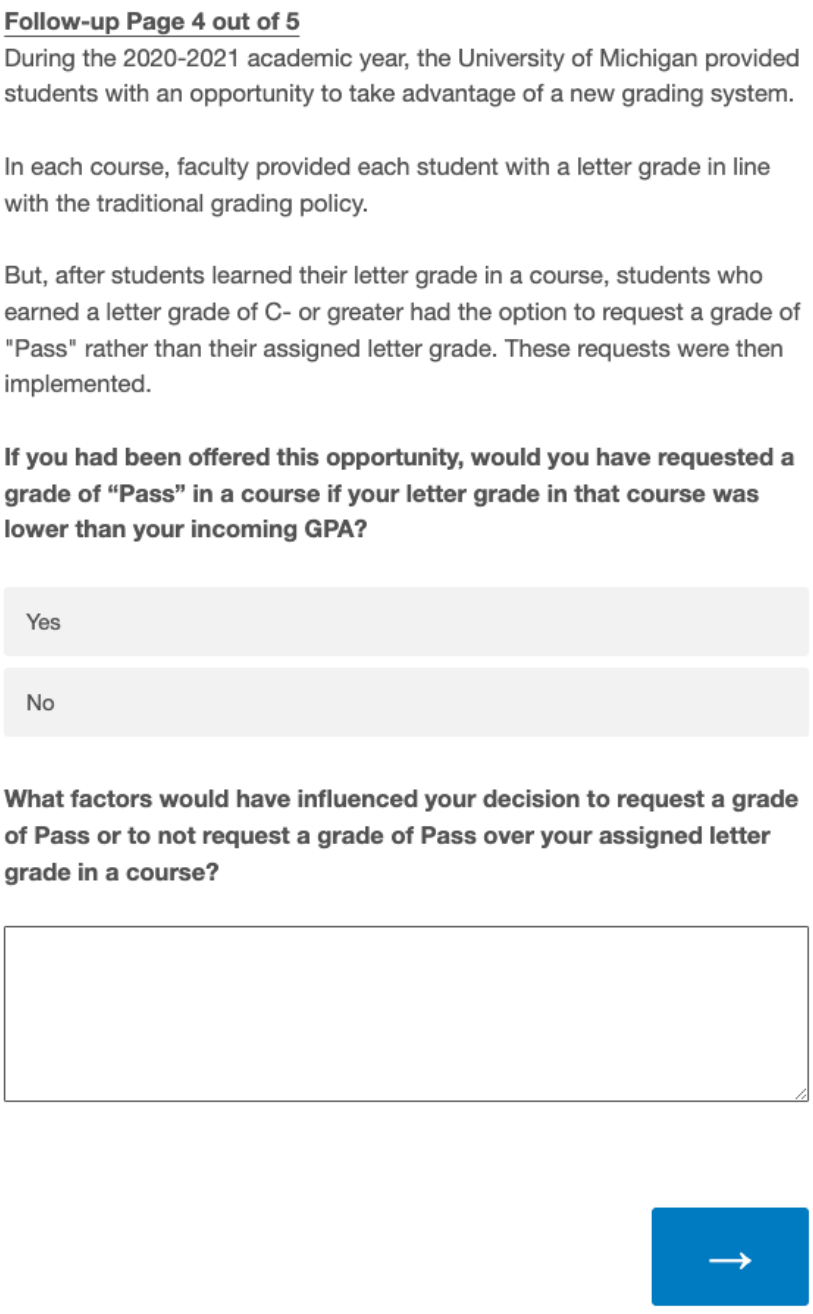


Figure C.9: Short followup survey

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?

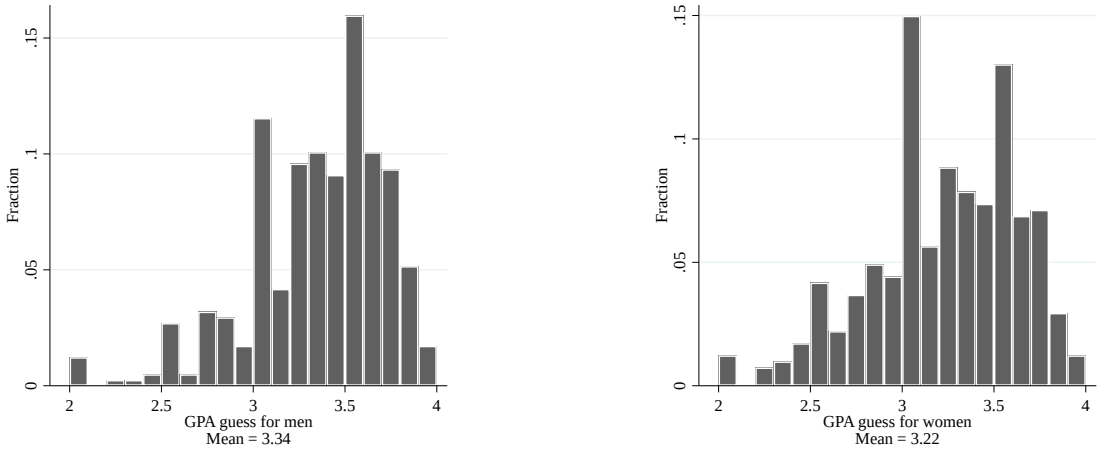
C.3 Additional results from the student belief study

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

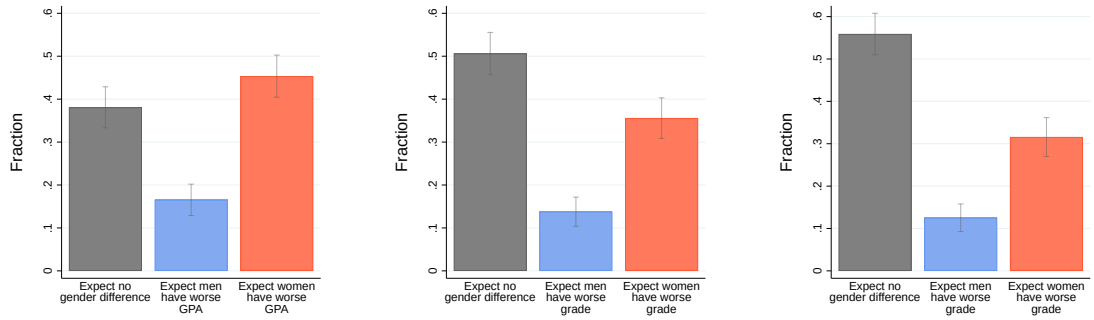
This figure plots the distribution of students' incentivized predictions of employers' inferences about GPA, when this information is missing from an applicant's resume. Panel (a) shows predictions about men, while Panel (b) shows predictions about women. All answers were provided on a slider form that 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.

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]

Appendix Figures C.11 and C.12 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 Anticipated Discrimination Study, anticipated discrimination is statistically significant ($p < 0.01$).

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



(a) Expected GPA, employers

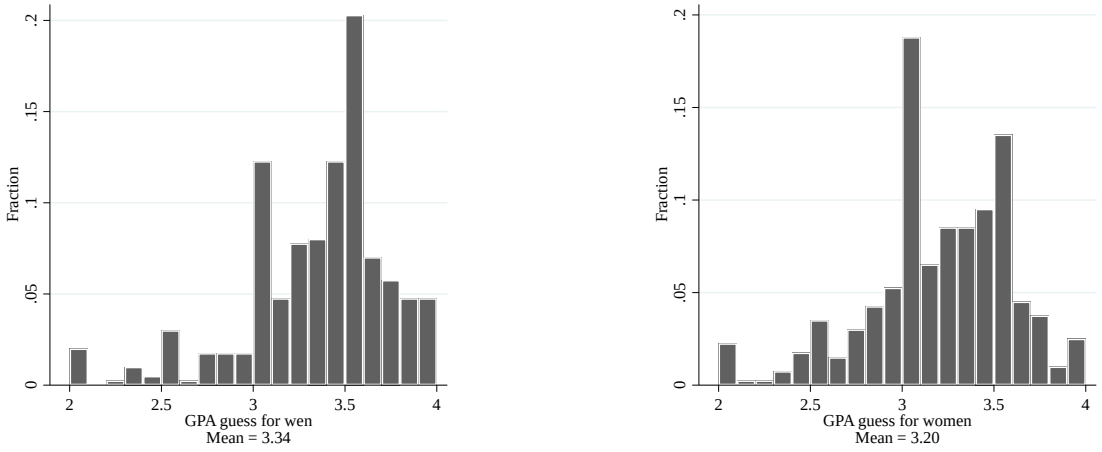
(b) Expected grade, employers

(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.

Figure C.12: 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

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. Table D.1 summarizes the employers who participates in the IRR experiment during the 2020/21 academic year.

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.

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.

We utilize the randomly assigned resume names to be able to examine the interaction of this 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 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 Additional results from the employer study

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 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.[31]

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).[32]

[31] 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).

[32] 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

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 Appendix 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.

justify undesirable behavior (Exley and Kessler, Accepted). Kessler, Low and Shan (2022) 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.

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

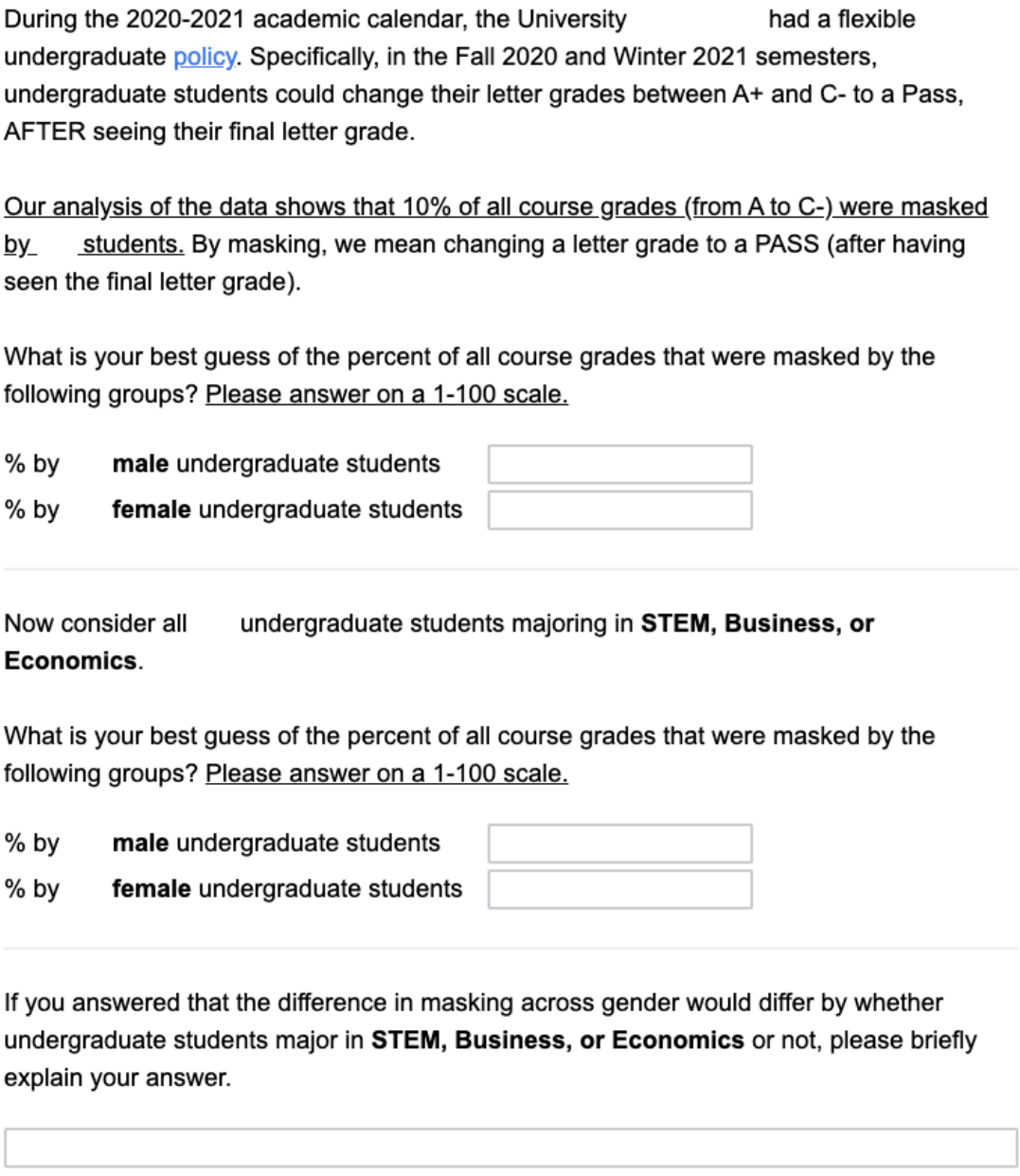


Figure E.2: Questions on perceived gender differences in grade concealment, conditional on same GPA and grade