AI-Adapted Grading I2D Proposal

1) Research Question and Project Overview

This project investigates how grading systems should adapt in an era in which artificial intelligence (AI) tools are widely accessible to students. The central research question is: Holding AI policy and the grading of individual assessments constant, how does shifting grade weight from homework to in-class assessment affect student learning, anxiety, and perceptions of AI?

We propose a randomized, multi-institution field experiment comparing two grading regimes: (1) traditional grading with AI-permissive policies and (2) an AI-adapted grading system that shifts weight away from homework toward in class assessment. In the AI-adapted system, the weight on work completed outside of class is cut in half and the freed weight is shifted to in-class assessments other than the final exam. Our hypothesis is that AI-adapted grading will lead to greater learning outcomes and reduced student stress.

This project is important because advances in AI have fundamentally altered the cost-benefit analysis of completing assignments outside the classroom, undermining traditional homework-based assessment. At the same time, restrictive AI policies may increase student anxiety, reduce learning opportunities, and lead to a bimodal grade distribution, where one group of students cheats by using AI and the other group does not. Identifying grading systems that are both robust to AI misuse and supportive of student learning is therefore critical for the future of economics education and higher education broadly.

2) Background

The emergence of AI tools like large language models (LLMs) introduces new challenges for assessment. AI can successfully complete most forms of homework, including complex problem sets, raising concerns about academic integrity and the productivity of take-home assignments. At the same time, AI may serve as a valuable learning aid when used appropriately, creating tension between restrictive and permissive policies.

Bastani et al. (2024) and Stromberg, Lei, and Wu (2026) found that students with generative AI access outperformed peers initially but performed worse on a closed-book exam once access was removed. Wecks et al. (2024) report similar exam-performance penalties among university students using generative AI, with effects largest for high-ability students. Forero and Herrera-Suárez (2023) document similar declines in a college physics course. These findings align with a related principle in cognitive science: offloading effortful retrieval undermines long-term retention (Grinschgl, Papenmeier, and Meyerhoff 2021). When students short-circuit the

productive struggle of working through a problem, they forgo the very mechanism by which homework was thought to produce learning.

Restrictive AI policies carry their own costs. Because detection of AI-generated work is unreliable, restrictive regimes risk producing bimodal grade distributions in which compliant students are systematically disadvantaged. With cheating in general, McCabe, Treviño, and Butterfield (2001), reviewing a decade of survey research, find that a perception of widespread peer cheating is a strong predictor of student dishonesty. Zhao et al.'s (2022) meta-analysis of 38 studies finds a similar result.

Despite increasing discussion of AI in education, there is little causal evidence on how grading systems should adapt. Instructors can (1) ignore generative AI and continue with the status quo, or (2) redesign courses to be robust to AI by shifting toward in-class assessment. This project will contribute to the literature by providing experimental evidence comparing these approaches across multiple institutions and instructors.

3) Methodology

The study population consists of undergraduate students enrolled in economics courses at U.S. liberal arts colleges and universities. Participating faculty will be recruited through EENE.

Randomization occurs at the term level. Each participating instructor teaches the same course in two of three academic terms (Spring 2027, Fall 2027, and Spring 2028), once under each of the two treatment arms, with the starting arm assigned at random by the research team. Instructors may pair Spring 2027 with Fall 2027 or Spring 2028, or Fall 2027 with Spring 2028, depending on their teaching schedules. Instructors who do not yet know their Fall 2027 or Spring 2028 teaching assignment can still participate if they are confident that they will teach the same class again in one of the two terms.

Because each instructor serves as their own comparison group, the design controls for differences across instructors, courses, and institutions. Final exam grades are used as the measure of learning. The same final exam is administered in both terms and kept protected (not returned to students, not posted, and not used outside the study), so that exam scores are directly comparable across the two terms.

Where an assessment is applicable, courses may additionally administer the econ-assessments.org assessment as a secondary measure of learning.

We aim to recruit approximately 45 instructors. With an average class size of 60 students and about a quarter of instructors teaching two sections per term, this yields roughly 3,400 students per term and approximately 6,750 student observations across each instructor's two study terms (see Power Analysis). To be eligible, a course must be taught by the same instructor in both

terms, determine final grades as a weighted average of graded components with fixed percentage weights, include a proctored final exam worth at least 20% of the course grade, include at least one other in-class or proctored assessment, and place between 20% and 30% of the course grade on work completed outside of class in the control term. If an instructor teaches multiple sections in a term, all sections run under the same arm.

Control/Treatment Arms

1) Traditional Grading, AI-Permissive (Control)

   Standard weighting (e.g., homework, exams, participation). AI use is allowed without restriction on work completed outside of class. Work completed outside a proctored setting (problem sets, online homework, take-home quizzes, papers, projects) carries between 20% and 30% of the course grade (the "control weight") and is graded as the instructor chooses (for correctness, for completion, or a mix), with each assignment graded the same way in both terms. In-class and proctored assessments follow the instructor's normal rules, which typically prohibit AI.

2) AI-Adapted Grading (Treatment)

   Homework has half the weight of the control group. Increased emphasis on in-class or proctored assessments. AI use is permitted but made to be less directly advantageous to a student's final grade absent improved learning. Specifically, the weight on every outside-of-class component is cut to half of the control weight, a reduction of 10 to 15 percentage points. The freed weight is reallocated to in-class and proctored assessments other than the final exam (midterms and quizzes), roughly in proportion to their control weights, rounded so that every component's weight is a multiple of 5. The weight on the final exam is the same in both arms. Assignments, due dates, the way each outside-of-class assignment is graded (for correctness or for completion, as chosen by the instructor), the number and format of in-class assessments, and the AI policy are identical to the control term; only the weights change. For example, a control term with homework at 30%, midterms at 20%, and a final exam at 50% becomes, in the treatment term, homework at 15%, midterms at 35%, and a final exam at 50%.

Standardization

Instructors will receive a detailed implementation guide covering both treatment arms (see Appendix). Common templates for syllabi language (AI policies, grading weight ranges) will be provided. Before the first study term, and before learning their starting arm, each instructor fixes the course's basic grading structure, assessments, and both the control and treatment weights. Instructors may adjust their course design (weights, number of exams, assignments) up to that point, but the design is then locked for both terms except for the prescribed arm difference. Only after the design is locked does the research team randomly assign and reveal the arm in which the instructor starts, so that knowledge of the starting arm cannot influence how the course is

configured. All participating courses determine final grades as a weighted average with component weights in multiples of 5.

Ensuring Treatment Fidelity

Fidelity is supported in five ways. (1) Instructors fix the course's grading structure, assessments, and both sets of weights (control and treatment) on the weight worksheet before they learn their starting arm. (2) The same final exam is used in both terms and kept protected: it is not returned or posted, students may review it only under supervision, and its questions are not used in review materials or other courses while the study runs. (3) The AI policy, assignments, due dates, the way each outside-of-class assignment is graded, in-class assessments, the final exam weight, participation weight, late-work and extra-credit policies, and curving rules are held constant across the two terms, and template syllabus language is provided for every block. (4) Instructors keep a deviation log of anything that departs from the locked design and submit it with their end-of-term data, even if it is empty, along with the syllabus they used. (5) Students are not told which arm their course is in, and the study is described to them only with IRB-approved language.

Outcomes to Measure

1) Final exam scores - the primary outcome. Each instructor switches arms between their two study terms and administers the same final exam, so otherwise identical classes are compared. Instructors report each student's raw, uncurved percentage score on the shared final exam.

2) Performance on standardized assessments (e.g., Cornell-style economics tests) - for courses where an applicable assessment exists

3) Student perceptions (fairness, stress, clarity, AI anxiety), collected via surveys**.** Surveys are administered by Cornell University in the first two weeks and the last two weeks of each term. Students receive a personal link by email; instructors provide class rosters with student email addresses in the first week of the term and make brief in-class announcements when each survey opens. Instructors do not see individual responses.

4) Course completion and grade distributions**.** After final grades are submitted, instructors send the research team a file with one row per enrolled student, identified by student name and institutional email address, containing the final exam score, the final numeric and letter course grade, and completion status (completed, withdrew, or did not take the final). Identified student data are shared through Cornell Secure File Transfer and stored by Cornell, so that exam scores and grades can be linked to survey responses.

Empirical Strategy

We estimate regressions of the following form:

$$Outcome_{ijt} = \alpha + \beta_1 Treated_{jt} + X_i\gamma + \tau_t + \rho_j + e_{ijt}$$

In this equation, the dependent variable $Outcome_{ijt}$ is the test score (or other outcome) of student i in course-instructor pair j in time period (term) t. The variable $Treated_{jt}$ is a binary variable indicating whether instructor-course pair j was treated in term t (equal 1 if treated, and 0, otherwise). X_i is a vector of student demographic characteristics. Then, we control for academic term with term fixed effects τ_t (Spring 2027, Fall 2027, or Spring 2028). ρ_j are instructor-course fixed effects. The error term is denoted with e. Standard errors are clustered at instructor level (students who have had the same instructor are not independent).

This approach controls for time-invariant, unobserved characteristics of the instructors, and compares outcomes of students in the treatment and control groups taught by the same instructor.

Power Analysis

Under plausible assumptions regarding persistence in student performance across terms, a sample of approximately 45 instructors is expected to have 80% power to detect effects of roughly 0.20-0.33 standard deviations. In raw score terms, this corresponds to approximately 3 points when exam-score standard deviations are between 10 and 15 points. If underlying exam-score variation is larger (e.g., a standard deviation of 20 points), a 3-point effect represents only 0.15 standard deviations, and substantially more instructors would be required to achieve conventional levels of statistical power. Power will be greatest when term-to-term performance is highly correlated within instructors and more limited when outcomes exhibit substantial variation across terms. Under moderate to strong within-instructor correlations, approximately 45 instructors should provide adequate power to detect a policy-relevant increase of 3 exam-score points when exam-score standard deviations are no greater than about 15 points.

Raw Exam SD	3-point Effect	Approximate Minimum Instructors Needed
10	0.30 SD	20-45
12	0.25 SD	25-65
15	0.20 SD	40-100
20	0.15 SD	70-175

So, a sample of 45 instructors is sufficient for a 3-point effect if exam-score standard deviations are around 15 points or lower and within-instructor persistence is moderate to high. If standard deviations are closer to 20 points, 45 instructors would likely be underpowered for detecting a 3-point effect.

Technical Challenges and Mitigation

1) Treatment Noncompliance - Mitigation: Clear protocols**,** including a single implementation guide covering both arms, course designs and weights locked before the starting arm is revealed, template syllabus language, a fixed weight worksheet, and a per-term deviation log (see Ensuring Treatment Fidelity).
2) Instructor Heterogeneity - Mitigation: Randomization, inclusion of fixed effects, randomizing the order of treatment and control for the same instructor across two terms.
3) Measurement of AI Use - Mitigation: Focus on outcomes and survey responses rather than attempting to observe AI usage
4) Student Sorting Across Courses - Mitigation: randomize across terms so students are not aware of the specific policy until the beginning of the course. All sections taught by an instructor in a given term run under the same arm, so students cannot select a section by grading policy.
5) Carryover Effects - Instructors may teach slightly differently in their second study term after experiencing the first arm. Mitigation: the order of the arms is randomized, so carryover is balanced across arms in expectation, course designs are locked before the first term, term fixed effects are included, and we will test whether estimated effects differ by arm order.
6) Spillover Between Cohorts - Students from an instructor's first study term may discuss the grading approach or the final exam with students taking the course in the second term, or post about it online. Mitigation: the final exam is kept protected and is not returned, students are not told which arm their course is in, and suspected exam compromises are reported to the research team. We expect this risk to be minor.

4) Timeline

Phase 1 (Spring-Fall of 2026):

1) Present project at EENE annual conference.
2) Finalize research design
3) Recruit faculty participants
4) Develop standardized materials (implementation guide, syllabus templates, weight worksheet)
5) Apply for IRB approval

Phase 2 (Spring Term of 2027):

1) Implement treatments in first wave of courses (instructors starting in Spring 2027 teach their randomly assigned starting arm)

Phase 3 (Fall Term of 2027):

1) Instructors paired across Spring 2027 and Fall 2027 switch to the other arm and administer the same final exam. Instructors starting in Fall 2027 teach their randomly assigned starting arm.

Phase 4 (Spring Term of 2028):

1) The remaining instructors switch to the other arm and administer the same final exam

Phase 5 (Summer-Fall of 2028):

1) Clean and analyze data
2) Draft research paper
3) Present preliminary results at conferences

Phase 6 (Winter of 2028-29):

1) Submit manuscript to an economics education journal

5)Research Team

Josh Staveley-O'Carroll is an associate professor of economics at Babson College. He has published several papers on classroom experiments and demonstrations in *The Journal of Economic Education*, *Southern Economic Journal*, *Eastern Economic Journal*, and more. He also completed a two-year controlled study on his own microeconomic principles courses resulting in the publication, "Impact of Classroom Demonstrations and Online Discussion on Student Achievement," in 2015 and led the EENE project "The Use of Generative Artificial Intelligence in Economics."

Stefani Milovanska-Farrington is an Associate Professor at Florida Polytechnic University. She is the creator of "Choice: Economics materials for success" on Substack, and an author of a Microeconomics textbook. She has collaborated with FEE and MRU, taught Microeconomics at UTampa (2018 – 2026), and worked as a Director of Assessment at UTampa. She is a CSWEP liaison, IZA@LISER Research Fellow, and an AEF Director. Stefani was awarded APEE's *James Gwartney Best Educational Note Award* in 2026, AEF's *Teaching Fellow Award* in 2025, and UTampa's COB *Researcher of the Year Award* in 2022. She has published more than 30

peer-reviewed academic articles. Stefani also participated in the Spartan Incubator Entrepreneurship program at UTampa in 2025, and is the founder of Love QT (Healthy fun food art for kids).

Daniel Savelle is an Assistant Professor of Economics at St. Mary's College of Maryland. He studies alternative grading in Economics with a focus on Specifications Grading. He recently published "More than just a weighted average: Economic statistics with auto specs grading" and is working on a follow up that adapts the grading scheme to the age of AI. At St. Mary's, Dan is a member of AI Governance Faculty Workgroup and co-chaired the data science minor proposal committee which led to a new data science minor in the Fall of 2025.

Jonathan Ernest is an Assistant Professor of Economics at the Weatherhead School of Management, Case Western Reserve University, where he has been teaching since 2019. He serves as the Academic Representative and Undergraduate Program Coordinator for the CWRU Department of Economics and advises the Weatherhead Economic Society (WES), the CWRU Real Estate Club (REC), and CWRU Quiz bowl. Courses taught include Principles of Microeconomics, Principles of Macroeconomics, Intermediate Microeconomics, Economic Analysis of Business Strategies, and the Economics of Sports. For his efforts, he has received multiple nominations for the WSOM Undergraduate Teaching Award and the Bruce Jackson, MD, Award for Excellence in Undergraduate Mentoring. He has been featured in well over 100 news stories across various media platforms, including television, print, online, radio, podcasts, and documentaries. In 2022, he was named "Coach of the Year" for the Emerging Leaders Program, and in 2023-2025, he was recognized as "MVP" by the CWRU Media Relations team.

Anthony (Tony) Underwood is associate professor of economics at Dickinson College in Carlisle, PA. Professor Underwood earned his Ph.D. in Economics from Colorado State University. He has since published in several academic journals, including the *Journal of Risk and Uncertainty*, *Energy Policy*, *Ecological Economics*, *The Journal of Economic Education*, and *Review of Political Economy*. His research interests include household energy use and emissions, the challenges for climate change mitigation posed by demographic change, and advancements in economic education. Professor Underwood regularly teaches environmental economics, econometrics, and microeconomics. In his advanced econometrics course, students practice the importance of reproducible and transparent methods in research through completion of their own empirical research project, which is also the focus of a 2019 co-authored publication in *The Journal of Economic Education*. He is also a member of the Economic Education Network for Experiments (EENE) Review Committee.

6) References

Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman. "Generative AI can harm learning." *The Wharton School Research Paper* (2024).

Forero, Manuel G., and H. J. Herrera-Suárez. "ChatGPT in the classroom: boon or bane for physics students' academic performance?." *arXiv preprint arXiv:*2312.02422 (2023).

Grinschgl, Sandra, Frank Papenmeier, and Hauke S. Meyerhoff. "Consequences of cognitive offloading: Boosting performance but diminishing memory." *Quarterly Journal of Experimental Psychology 74*, no. 9 (2021): 1477-1496.

Marshall, Emily C., Brandon J. Sheridan, Douglas C. McKee, George Orlov, William Goffe, Allen Bellas, Brooks Depro, et al. "Measuring academic mindset in economics courses." *International Review of Economics Education* 52 (2026): 100349.

McCabe, Donald L., Linda Klebe Treviño, and Kenneth D. Butterfield. "Cheating in academic institutions: A decade of research." *Ethics & Behavior 11*, no. 3 (2001): 219-232.

Staveley-O'Carroll, James, Petar Stankov, Emily Marshall, Brandon Sheridan, Douglas McKee, Zahra Akbari, Chelsea Taylor Dowell, et al. "The use of generative artificial intelligence in economics." *SSRN Working Paper* No. 6174240 (2026).

Stromberg, David, Victor Lei, and Yanhui Wu. "The generative AI learning penalty: Evidence from Chinese secondary education." *SSRN Working Paper* No. 6868618 (2026).

Wecks, Janik Ole, Johannes Voshaar, Benedikt Jost Plate, and Jochen Zimmermann. "Generative AI usage and exam performance." *arXiv preprint arXiv:*2404.19699 (2024).

Zhao, Li, Haiying Mao, Brian J. Compton, Junjie Peng, Genyue Fu, Fang Fang, Gail D. Heyman, and Kang Lee. "Academic dishonesty and its relations to peer cheating and culture: A meta-analysis of the perceived peer cheating effect." *Educational Research Review* 36 (2022): 100455.

7) Appendix

Disclosure of AI use: The authors used generative AI (Claude, Anthropic) in preparing this proposal for tasks such as drafting and editing content, formatting, searching for additional relevant research, and checking citations. The research question, study design, and methodology were developed by the authors, although AI was used to critique some of these decisions. All AI-

assisted material was reviewed and approved by the authors, who take full responsibility for the content.

Student Pre-Survey

Section 1: Boilerplate and Demographics

Section 2: AI Usage (Staveley-O'Carroll et al., 2026)

1) Do you pay for a premium AI program?
    a) Yes
    b) No
2) How many hours per day on average do you use AI outside of class?
    a) Slider with number of hours
3) On what percent of your assignments for other classes do you generally use AI?
    a) Slider with percentage
4) Are you allowed to use AI tools to *help you study* for this class (e.g. reviewing notes, summarizing lectures, creating practice questions, etc.)?
    a) Yes, there are no restrictions on use.
    b) Yes, as long as I explicitly state how I used the technology.
    c) Yes, I am allowed to use it for studying in some cases.
    d) No, this use of AI tools is not allowed in this class.
5) Are you allowed to use AI tools to *complete assignments* in this class (e.g. problem sets, homework, papers, presentations, etc.)?
    a) Yes, there are no restrictions on use.
    b) Yes, as long as I explicitly state how I used the technology.
    c) Yes, I can use Gen AI to complete some assignments.
    d) No, this use of AI tools is not allowed in this class.
6) What ethical concerns do you have about the use of AI tools? [CHECK ALL THAT APPLY]
    a) Environmental (energy usage)
    b) Copyright
    c) Inequality
    d) Bias
    e) Other [Specify]
    f) None

Section 3: Student Stress, Anxiety, Fairness, and Clarity

7) How often do you generally feel anxious about your grades in other classes?
    a) Never

b) Rarely
c) Sometimes
d) Often
e) Always

8) How stressed are you about using AI on assignments in other courses?
a) Never
b) Rarely
c) Sometimes
d) Often
e) Always

9) Grading in other courses is generally fair.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

10) Grading policies in other courses are generally clear.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

Section 4: Relevance, Belonging, and Growth Mindset (RBG) (Marshall et al., 2026)

11) I have a good relationship with other students in my economics classes.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

12) I think that nearly everyone is capable of understanding economics if they work at it.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

13) Even when an economics course is challenging, I can perform quite well.

a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

14) I think about economic events I experience and witness in day-to-day life (e.g., in my own life and decisions, on the news, in internet articles).

a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

15) I think economics is interesting and applicable for people like me.

a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

16) I plan to take another economics course in the future.

a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

Student Post-Survey

Section 1: Boilerplate and Demographics

Section 2: AI Usage (Staveley-O'Carroll et al., 2026)

1) Do you pay for a premium AI program?
    a) Yes
    b) No
2) How many hours per day on average do you use AI outside of class?
    a) Slider with number of hours
3) On what percent of your assignments for this class did you generally use AI?
    a) Slider with percentage
4) Are you allowed to use AI tools to *help you study* for this class (e.g. reviewing notes, summarizing lectures, creating practice questions, etc.)?
    a) Yes, there are no restrictions on use.
    b) Yes, as long as I explicitly state how I used the technology.
    c) Yes, I am allowed to use it for studying in some cases.
    d) No, this use of AI tools is not allowed in this class.
5) Are you allowed to use AI tools to *complete assignments* in this class (e.g. problem sets, homework, papers, presentations, etc.)?
    a) Yes, there are no restrictions on use.
    b) Yes, as long as I explicitly state how I used the technology.
    c) Yes, I can use Gen AI to complete some assignments.
    d) No, this use of AI tools is not allowed in this class.
6) What ethical concerns do you have about the use of AI tools? [CHECK ALL THAT APPLY]
    a) Environmental (energy usage)
    b) Copyright
    c) Inequality
    d) Bias
    e) Other [Specify]
    f) None

Section 3: Student Stress, Anxiety, Fairness, and Clarity

7) How often do you feel anxious about your grade in this course?
    a) Never
    b) Rarely
    c) Sometimes

d) Often
e) Always

8) How stressed were you about using AI on assignments in this course?
   a) Never
   b) Rarely
   c) Sometimes
   d) Often
   e) Always
9) Grading in this course was fair.
   a) Strongly disagree
   b) Somewhat disagree
   c) Neither agree nor disagree
   d) Somewhat agree
   e) Strongly agree
10) Grading policies in this course were clear.
   a) Strongly disagree
   b) Somewhat disagree
   c) Neither agree nor disagree
   d) Somewhat agree
   e) Strongly agree

Section 4: Relevance, Belonging, and Growth Mindset (RBG) (Marshall et al., 2026)

11) I have a good relationship with other students in my economics classes.
   a) Strongly disagree
   b) Somewhat disagree
   c) Neither agree nor disagree
   d) Somewhat agree
   e) Strongly agree
12) I think that nearly everyone is capable of understanding economics if they work at it.
   a) Strongly disagree
   b) Somewhat disagree
   c) Neither agree nor disagree
   d) Somewhat agree
   e) Strongly agree
13) Even when an economics course is challenging, I can perform quite well.
   a) Strongly disagree

b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

14) I think about economic events I experience and witness in day-to-day life (e.g., in my own life and decisions, on the news, in internet articles).
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

15) I think economics is interesting and applicable for people like me.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

16) I plan to take another economics course in the future.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

Section 5: Out-of-Class Assignments

17) Approximately how many hours per week did you spend working on the out-of-class assignments in this course?
a) Slider with number of hours

18) Out-of-class assignments in this course were very difficult.
a) Strongly disagree
b) Somewhat disagree
c) Neither agree nor disagree
d) Somewhat agree
e) Strongly agree

The faculty implementation guide, which serves as the study protocol, is reproduced starting on the next page as it will be distributed to participating instructors.

Faculty Implementation Guide

Grading in the Age of AI: A Multi-Institution Field Experiment

Economic Education Network for Experiments (EENE) · Spring 2027, Fall 2027, and Spring 2028

One guide for both study arms

Contacts

Study team (design, grading questions, data submission): [NAME], [EMAIL]

Survey administration (Cornell University): [NAME], [EMAIL]

IRB protocol: [NUMBER] · Items in [square brackets] are placeholders to be confirmed by the study team.

1. How to use this guide

Every faculty participant receives this same document. You will teach the same course in two of three academic terms (Spring 2027, Fall 2027, and Spring 2028), one term under each of the two study arms. The guide calls these your first and second study terms. This guide tells you how to set up the course once, what to hold constant, what to change between terms, and what to send to the study team and to Cornell.

Read the whole guide before your first study term. You will come back to Sections 5 and 6 (the arm-specific instructions) and Section 11 (the term checklist) most often.

Quick start

1. Read Sections 2–4 to understand the design and what you are committing to.
2. Complete Step 1 (lock your course design) in Section 4, including the grade-weight worksheet in Appendix A. It gives you your Control weights and your Treatment weights side by side. Send the worksheet to the study team.
3. Complete Step 2 (record your assigned starting arm) in Section 4. You receive your starting arm only after your worksheet is in.
4. Finalize your final exam and store it securely. The identical exam is given in both terms.
5. Prepare two versions of your syllabus using the template language in Appendix B.
6. Use the checklist in Section 11 each term.

2. The study in brief

Research question. How do grading structures that shift weight from homework toward in-class assessment affect student learning, stress, and perceptions of AI? The comparison is with traditional grading, with AI use outside of class permitted under both.

Two arms. The only thing that differs between the arms is how much outside-of-class work counts for.

	Control (traditional grading, AI-permissive)	Treatment (AI-adapted grading)
Grading of outside-of-class work	Your choice for each assignment (correctness or completion)	Same as Control for every assignment (identical)
Weight of outside-of-class work	Between 20% and 30% of the course grade (the Control weight)	Half of the Control weight
In-class / proctored assessments	Your normal weight	Normal weight plus the freed weight, spread roughly in proportion across assessments other than the final exam
Final exam weight	At least 20% of the course grade	Same weight as Control (identical)
AI policy	AI permitted on outside-of-class work	AI permitted on outside-of-class work (identical)
Everything else	Identical	Identical

Design. Each instructor teaches the same course in two academic terms, once under each arm, in a randomly assigned order. You serve as your own comparison group. That is what lets the study separate the effect of the grading structure from differences in instructor, institution, and course. It works only if everything else about the course stays the same across the two terms, which is why the guide is strict about what may and may not change.

Outcomes. Final exam scores (the same exam in both terms), student surveys at the start and end of each term (administered by Cornell University), and final course grades.

Why fidelity matters. Any change between your two terms that is not part of the arm definition cannot be separated from the treatment effect in the analysis. That includes a new assignment, a reweighted midterm, a revised exam, or a different AI policy. Small drifts across many instructors add up. If something must change, record it in the deviation log (Appendix C) and tell the study team.

3. Eligibility and your commitments

Participation requires all of the following.

- You teach the same course (same course number and level) in two of the three study terms: Spring 2027 and Fall 2027, Spring 2027 and Spring 2028, or Fall 2027 and Spring 2028.
- The final course grade is a weighted average of the graded components listed in the syllabus, with a fixed percentage weight for each. This is the standard grading method. Courses that use another scheme, such as specifications grading, standards-based grading, or a points total without fixed category weights, must convert to a weighted average before their first study term.
- The course has a proctored final exam (in class, or in your institution's proctored testing setting) worth at least 20% of the course grade, with the same weight in both terms.

- The course has at least one graded component completed outside of class, and outside-of-class work totals between 20% and 30% of the course grade in the Control term.
- The course has at least one in-class / proctored assessment other than the final exam (for example, a midterm or in-class quizzes). It receives the freed weight in the Treatment term.
- You will lock your course design and both sets of weights before you learn your starting arm (Section 4.2).
- You will administer the identical final exam in both terms and keep it protected (Section 7).
- You will hold everything other than the prescribed arm differences constant across the two terms (Section 4.2).
- In the first week of each term you will send your class roster with student email addresses to Cornell. You will also make brief in-class announcements when each survey opens (Section 9).
- After each term you will send the study team each student's final exam score paired with their final course grade, with student identifiers, by secure file transfer (Section 10). If you teach more than one section in the same term, all sections run under the same arm.

4. One-time setup before your first study term

4.1 Key terms

Term	Meaning in this study
Weighted average	The standard grading method. Each graded component has a fixed percentage weight, and the final course grade is the weighted average of the component scores. Every participating course must use it.
Outside-of-class work	Any graded work students complete outside a proctored setting. This includes problem sets, online or publisher-platform homework, take-home quizzes, papers, projects, discussion posts, and similar work.
In-class / proctored assessment	Any graded work completed under supervision, such as midterms, the final exam, in-class quizzes, in-class writing, and proctored online tests.
Participation / other	Attendance, participation, and similar components that are neither outside-of-class work nor a proctored assessment. These are held at the same weight in both arms.
Control weight	The share of the course grade coming from outside-of-class work in the Control term. It must be between 20% and 30%.
Freed weight	Half of the Control weight (10 to 15 percentage points). In the Treatment term, this amount moves from outside-of-class work to in-class / proctored assessments other than the final exam, spread roughly in proportion to their Control weights.

4.2 Step 1: Lock your course design

Before you lock your design, you may change anything about the course, including the number of exams, the weights, the assignments, the textbook, and the participation policy, within the eligibility limits in Section 3. You lock the design by completing Appendix A and sending it to the study team before your first study term and before you learn your starting arm. **Once locked, the course stays fixed** through the end of your second study term, except for the prescribed arm difference.

Whatever you decide, the final course grade must be a weighted average of the components you list, with a fixed percentage weight for each. Choose weights that are multiples of 5. For outside-of-class components, use multiples of 10 so that half of each weight is still a multiple of 5. Outside-of-class work must total between 20% and 30% in the Control column, and the final exam must be at least 20% with the same weight in both columns.

Design both terms at the same time. The lists below show what is held constant and what varies. Appendix A walks you through the weights.

Held constant across both terms

- The weighted-average grading method
- Course content, textbook, topic sequence, and pacing
- Number, timing, format, and length of in-class / proctored assessments
- Number, timing, content, and due dates of outside-of-class assignments
- How each outside-of-class assignment is graded, including rubrics and partial-credit rules. You may grade for correctness, for completion, or a mix across assignments, as long as each assignment is graded the same way in both terms.
- The final exam itself, which is identical and protected (Section 7), and its weight
- AI policy, which permits AI on outside-of-class work and keeps your normal rules in proctored settings (Section 8)
- Participation / attendance weight, late-work policy, drop-lowest rules, extra credit (preferably none), grading scale, curving policy
- Modality, TA support, office hours, and, where feasible, meeting times and section size

Varies by arm (and nothing else)

- **Weight** of outside-of-class work, which is the Control weight in Control and half of it in Treatment
- **Weight** of in-class / proctored assessments other than the final exam. In Treatment, the freed weight is added roughly in proportion to each component's Control weight and rounded so that every component is a multiple of 5 (Appendix A).
- **Component weights listed in the syllabus** (Appendix B)

Do these now, before your first study term:

1. Complete Appendix A. Enter your Control weights and derive your Treatment weights. The freed weight is spread across in-class components other than the final exam roughly in proportion to their

Control weights, rounded so that every component is a multiple of 5. Fix this split once and use it whenever you are in the Treatment term. Send the completed worksheet to the study team. You receive your starting arm only after it is in.

2. Write and finalize the final exam. Store it securely. Do not use any of its questions in practice materials, review sessions, or other courses.
3. Prepare both syllabi from Appendix B. You do not need to send them to the study team in advance. You will send the syllabus you actually used with your end-of-term data (Section 10).
4. Set up your LMS or publisher platform for your first-term arm once you receive it (Sections 5 and 6 explain the grading configuration).

4.3 Step 2: Record your assigned starting arm

After the study team receives your completed worksheet, it assigns your starting arm at random and sends it to you by email. The random order is what makes the comparison valid, and locking the design first means that knowing your starting arm cannot influence how the course is set up. Do not change the worksheet after you receive your assignment.

Step 2 record

First study term (_____): ☐ Control ☐ Treatment

Second study term (_____): the other arm.

Course: _____ Sections per term: ____ Assignment received on: _____

5. Control arm instructions

Control arm at a glance (traditional grading, AI-permissive)

- Outside-of-class work carries your Control weight (20% to 30%).
- In-class / proctored assessments carry their normal weight.
- AI is permitted on outside-of-class work (same policy as Treatment).

5.1 Outside-of-class work

- Use the Control weights from Appendix A. The total weight on outside-of-class work is your Control weight.
- Assign the same assignments, with the same due dates and the same late policy, as in your Treatment term.

5.2 In-class and proctored assessments

- Give the same number of in-class assessments, in the same format and at the same points in the term, as in the Treatment term. Only the weights differ between the terms.
- Administer the study final exam under proctored conditions and keep it protected (Section 7).

5.3 Syllabus and communication

- Use the grading block with your Control weights and the shared AI-policy, final-exam, and study-notice language from Appendix B.
- Do not describe the Treatment arm, the study hypotheses, or which arm your course is in (Section 8.4).

5.4 LMS setup tips

- Enter the Control weights in your gradebook categories before the first day of class.
- Do not change any other gradebook or platform settings between terms.

6. Treatment arm instructions

Treatment arm at a glance (AI-adapted grading)

- Outside-of-class work carries half the Control weight.
- The freed weight moves to in-class / proctored assessments other than the final exam, spread roughly in proportion to their Control weights and rounded to multiples of 5 (Appendix A). The final exam keeps its Control weight.
- Same assignments, same due dates, same grading of each assignment, same in-class assessments as Control. Only the weights change.
- AI is permitted on outside-of-class work (same policy as Control).

6.1 Weights

- Each outside-of-class component is worth half its Control weight. Participation and similar components keep their Control weight.
- The freed weight is added to in-class / proctored components other than the final exam roughly in proportion to their Control weights. The final exam keeps its Control weight. Round the result so that every component is a multiple of 5. All of the freed weight must be reallocated, and the Treatment column must total 100%.
- If you have not completed Appendix A, do it now. Do not improvise the split during the term.

6.2 In-class and proctored assessments

- Same number, format, timing, and length as the Control term. Only the weights differ.
- Administer the study final exam under proctored conditions and keep it protected (Section 7).

6.3 Syllabus and communication

- Use the grading block with your Treatment weights and the shared AI-policy, final-exam, and study-notice language from Appendix B.
- Students may notice that homework counts less than in-class work. Describe it as the course's grading policy for the term. Do not describe the Control arm, the study hypotheses, or which arm the course is in (Section 8.4).

6.4 LMS and publisher-platform setup tips

- Do not change any other gradebook or platform settings from the Control term.
- Enter the Treatment weights in your gradebook categories before the first day of class.

7. The final exam (both terms)

Final exam scores are the study's primary measure of learning. The comparison between your two terms is only as good as the exam is identical and protected.

7.1 Same exam

- Give the identical exam in both of your study terms. That means the same questions, order, point values, time allowed, permitted materials (calculator, formula sheet), format (paper or online), and proctoring.
- Finalize the exam before your first study term and store it securely.
- Do not edit the exam between terms. If you discover an error in your first study term that must be fixed by the second, note it in the deviation log.
- While this study runs, do not administer this exam, or any of its questions, in any course or term outside the study (including summer sections or a colleague's section).

7.2 Keeping the exam protected

- Do not return the exam, post it, or distribute the answer key. Collect every page, including scratch paper, before students leave.
- Do not permit photos or copies. If the exam is online, disable answer review and post-exam access to questions. Use your normal lockdown or proctoring tools.
- Do not use exam questions in practice exams, review sessions, homework, or study guides.
- Students may review their graded exam only under your supervision, for example in office hours, and may not take notes, photos, or copies. Handle grade disputes the same way.
- If you believe the exam has been compromised (for example, a copy is circulating), tell the study team immediately.

7.3 Administration and scoring

- Proctor the exam yourself or through your institution's testing center, following the same procedure both terms. Handle accommodations through your normal process.
- Record each student's raw, uncurved percentage score. This is the number you report to the study team (Section 10).
- If you normally curve, apply the curve when computing course grades and use the same rule in both terms. Still report the raw exam score.

8. Rules that apply in both terms

8.1 AI policy

- Students may use generative AI tools on all outside-of-class work in both terms. Do not restrict, penalize, or attempt to detect AI use on that work, and do not run AI-detection software.
- In-class and proctored assessments follow your normal rules (typically no devices and no AI). Apply those rules identically in both terms.
- Use the same AI-policy syllabus text (Appendix B) word for word in both terms.
- Suspected misconduct on a proctored assessment is handled under your institution's academic-integrity policy, the same way in both terms.

8.2 Everything else stays the same

Section 4.2 has the full lists. In short, these must not change between your two study terms.

- The weighted-average grading method.
- Assignments, due dates, how each outside-of-class assignment is graded, late policy, drop-lowest rules, participation weight, grading scale, curving policy, extra credit.
- Number, timing, and format of in-class assessments, and the final exam and its weight.
- Textbook, content, pacing, modality, TA support, office hours.
- Any incentive attached to the surveys (Section 9).

8.3 Multiple sections

- If you teach more than one section of the course in a term, every section runs under the same arm with the same weights, assignments, and exam.
- Report section counts to the study team, and submit data for all sections (Section 10).

8.4 Talking with students about the study

- Use only the IRB-approved language provided by the study team when describing the study to students. Appendix B includes a placeholder for it (to be replaced with approved text).
- Do not tell students which arm the course is in, describe the other arm, or discuss the study hypotheses. If a student asks why the course is weighted the way it is, answer in terms of the course's grading policy rather than the study design.
- Consent for the surveys is handled by Cornell within the survey itself. You do not need to collect consent forms.

8.5 Deviations

Things happen. A weather closure cancels an exam, your institution mandates a policy change, or a platform discontinues a feature. If anything about the course changes from the locked design, record it in the deviation log (Appendix C) and email the study team as soon as practical. Deviations you report can be accounted for in the analysis. Deviations we do not know about cannot.

9. Student surveys (administered by Cornell)

Students complete two short surveys each term, an opening survey in the first two weeks and a closing survey in the last two weeks. Cornell University administers both. Students receive an email with a personal link. You do not host, distribute, or see the surveys, and you will not see individual responses.

Your role is limited to three things. Send the roster, make announcements, and keep the procedure identical in both terms.

When	Cornell	You
Before the term	Sends you the roster template, the submission address, and the deadline.	Watch for the email. Prepare to export your roster with email addresses from your LMS or registrar system.
Week 1	—	Send your roster with every enrolled student's email address to Cornell by [DATE], using the format Cornell specifies ([FORMAT / ADDRESS]). If enrollment changes materially after the add/drop deadline, send an updated roster.
Weeks 1–2	Emails each student a link to the opening survey and notifies you that it is open.	Make a brief in-class announcement the day you are notified (script below). Repeat it once before the survey closes.
Last two weeks	Emails each student a link to the closing survey and notifies you that it is open.	Same two announcements. Avoid scheduling the announcement on the day of the final exam.

Repeat the identical procedure in your second study term, with the same steps, the same announcement wording, and the same incentive (if any).

Suggested in-class announcement

"You should have received, or will receive today, an email from Cornell University with a link to a short survey for a study on grading in economics courses that this course is part of. It takes about [X] minutes and closes on [DATE]. Participation is voluntary, and I will not see anyone's individual responses. I'd appreciate it if you would complete it."

- Do not attach course credit or any other incentive to survey completion unless the study team instructs you to. Whatever you do must be identical in both study terms.
- Do not add your own questions or run a competing course survey during the survey windows.
- Please respond promptly to any roster questions from Cornell. A missing email address means a missing survey.

10. What you send the study team after each term

After final grades are submitted, send the study team one spreadsheet (.xlsx or .csv) per term. It should contain each student's final exam score paired with their final course grade. Send it through Cornell Secure File Transfer by [DATE] (e.g., within two weeks of the grade deadline). The study team will send upload instructions ([LINK]).

The file must identify each student. Include each student's name and institutional email address so that exam scores and grades can be matched to survey responses. Because the file contains identified student records, send it only through Cornell Secure File Transfer. Do not send it by email.

Column	Contents
student_email	The student's institutional email address, exactly as it appears on the roster you sent to Cornell.
student_name	The student's full name as it appears on your roster.
course	Course number or title.
term	Spring 2027, Fall 2027, or Spring 2028.
arm	Control or Treatment.
section	Section identifier if you teach more than one (A, B, …). Otherwise leave blank.
final_exam_score	Raw, uncurved percentage score on the study final exam (0–100).
final_course_pct	Final numeric course grade as a percentage, after any curve you normally apply. This is the weighted average used to assign the letter grade.
final_letter_grade	The letter grade recorded with the registrar.
status	Completed / Withdrew / Did not take final. Include a row for every student who was enrolled after the add/drop deadline, even if they withdrew, so that completion rates can be measured.

- Include all sections you teach. If you have more than one section, one file with the section column filled in is fine.
- Also send your deviation log (Appendix C), even if it is empty, and a copy of the syllabus you actually used that term.

11. Term checklist

Print this page and keep it with your course materials. Run the full list in your first study term and again in your second.

When	Task
Before your first study term (once)	☐ Complete Step 1 (lock your course design), Section 4. ☐ Confirm the course grade is a weighted average with fixed percentage weights, all multiples of 5, with outside-of-class work at 20% to 30% in Control and the final exam at 20% or more. ☐ Complete the weight worksheet in Appendix A (Control column and Treatment column) and send it to the study team. ☐ After the worksheet is in, record your assigned starting arm (Step 2, Section 4). ☐ Finalize the final exam and store it securely. Remove any of its questions from other course materials. ☐ Prepare both syllabi from Appendix B.
Before each term	☐ Confirm which arm this term is (Step 2 record) and post the matching syllabus. ☐ Enter that arm's weights in the gradebook. ☐ Confirm nothing else has changed from the locked design. If it has, log it.
Week 1	☐ Send the roster with student email addresses to Cornell by [DATE]. ☐ Watch for Cornell's notice that the opening survey is open.
Weeks 1–2	☐ Make the opening-survey announcement in class. Repeat it once before the survey closes.
During the term	☐ Keep the exam protected. Keep exam questions out of review materials. ☐ Record any deviation in Appendix C and email the study team.
Last two weeks	☐ Make the closing-survey announcement in class. Repeat it once before the survey closes. ☐ Administer the study final exam. Collect every page and do not return it.
After grades are submitted	☐ Send the identified exam-score and course-grade file through Cornell Secure File Transfer by [DATE]. ☐ Send the deviation log (even if empty) and the syllabus you used.

When	Task
Between terms	☐ Switch to the other arm. Update the weights in your gradebook and syllabus. Change nothing else. ☐ Re-read Section 5 or 6 for the arm you are about to run.

12. Frequently asked questions

I use total points, specifications grading, or standards-based grading rather than percentage weights. Can I participate?

Only if you convert to a weighted average with fixed category weights before your first study term and keep it for both terms. The Treatment arm works by changing category weights, so the study needs every participating course to grade this way.

Outside-of-class work is currently more than 30% (or less than 20%) of my course grade. Can I participate?

Yes, if you bring the Control weight into the 20% to 30% range before you lock your design. Your Treatment term then puts half of that weight on outside-of-class work and moves the rest (10 to 15 percentage points) to in-class assessments other than the final exam. After you lock, the weights are fixed.

The proportional split gives one of my in-class components a weight like 23%. What do I do?

Round so that every component is a multiple of 5 while keeping the Treatment column at 100%. Appendix A shows an example. Small departures from exact proportionality are fine, but large ones are not.

Can I move some of the freed weight to the final exam?

No. The final exam has the same weight in both terms, so that the stakes of the exam do not change with the arm. The freed weight goes only to other in-class / proctored assessments, such as midterms and in-class quizzes.

I do not yet know what I will teach in Fall 2027. Can I still join?

Yes. You need to teach the same course in two of the three study terms. You can pair Spring 2027 with Spring 2028, or start in Fall 2027 and finish in Spring 2028.

Can I drop the lowest homework score, or offer extra credit?

Only if you do exactly the same thing in both terms. We recommend no extra credit during the study. Dropping the lowest score is fine if the rule is identical both terms.

How do I treat participation or attendance?

As “participation / other.” It keeps the same weight in both arms and is not part of the outside-of-class weight.

What if a student uses AI on a proctored exam?

That is academic misconduct under your institution’s policy, handled the same way in both terms. The AI-permissive policy applies only to outside-of-class work.

Can I curve the final exam?

Report the raw, uncurved score to the study team. You may apply your usual curve when computing course grades, provided the rule is the same in both terms.

A student needs to take the final early or late. Is that allowed?

Yes, under your normal make-up and accommodation procedures, as long as the exam stays protected. Note any unusual arrangement in the deviation log.

I teach two sections of the course. Can one be Control and one Treatment?

No. All sections in a given term run under the same arm. Cross-section comparisons within a term are not part of the design and would let students sort between policies.

Can I give the study exam in a summer section, or share it with a colleague?

No. The exam must not be administered anywhere outside your two study terms, and it should not leave your possession.

Students figured out the exam is reused, or a copy is circulating. What now?

Tell the study team immediately and record it in the deviation log.

My institution requires a change mid-study (for example, a new attendance policy). What do I do?

Comply with your institution, record the change in the deviation log, and email the study team. Wherever possible, apply the change in both terms so it is not confounded with the arm.

Will I see the survey results?

You will not see individual responses. The study team plans to share aggregate findings with all participating faculty once the analysis is complete.

Who do I contact with a question this guide does not answer?

The study team contact on the first page. For anything about the surveys or the roster, contact Cornell (also on the first page).

Appendix A. Grade-weight worksheet

Complete this once, before your first study term and before you learn your starting arm. It produces the weights for both of your terms. The worksheet assumes the final grade is a weighted average of the listed components, which every participating course must use.

1. List every graded component of the course and mark its type. O is outside-of-class work, I is in-class / proctored assessment, and P is participation / other.
2. In the Control column, enter your weights. They must sum to 100. Use multiples of 5, and use multiples of 10 for O rows so that half of each is still a multiple of 5. The O rows must total between 20 and 30, and the final exam must be at least 20.
3. In the Treatment column, enter half the Control weight for each O row and copy the Control weight for each P row and for the final exam. The freed weight is the total of the O reductions.
4. Spread the freed weight across the I rows other than the final exam roughly in proportion to their Control weights. Then adjust so that every component in the Treatment column is a multiple of 5 and the column sums to 100. Fix this split now. It is used whenever you are in the Treatment term. Two worked examples follow the worksheet.

Your worksheet

Course: _____ Completed on: _____ Sent to study team on: _____

Component	**Type (O / I / P)**	**Control weight**	**Treatment weight**
Total	N/A	**100%**	**100%**

Worked example 1

Component	Type	Control weight	Treatment weight
Problem sets	O	30%	15%
Midterm exams (two, equally weighted)	I	20%	35%
Final exam	I	50%	50%
Total		**100%**	**100%**

Freed weight is 15%. The midterms are the only in-class component other than the final exam, so all 15 points move to them (20% becomes 35%, or 17.5% per midterm). The final exam stays at 50%.

Worked example 2 (with a paper and participation)

Component	Type	Control weight	Treatment weight
Online homework	O	20%	10%
Term paper	O	10%	5%
Participation	P	5%	5%
In-class quizzes	I	10%	15%
Midterm	I	20%	30%
Final exam	I	35%	35%
Total		**100%**	**100%**

Freed weight is 15%. In proportion to the Control weights of the in-class components other than the final exam (10, 20), the split is 5 / 10, giving 15 / 30. The final exam stays at 35%. The term paper is still assigned, still due on the same date, and still graded the same way. It simply carries half the weight.

Appendix B. Syllabus template language

Build two syllabi before your first study term. They are identical except for the component weights in the grading block: one lists your Control weights and the other your Treatment weights, both from Appendix A. Every other block below must appear word for word in both.

B.1 Grading block (both terms; only the weights differ)

Grading. Your course grade is a weighted average of the following components. [list each component and its weight for this term, Control or Treatment, from Appendix A]

[Describe how homework and other assignments completed outside of class are graded. Use identical sentences in both syllabi.] In-class assessments and the final exam are graded for correctness.

B.2 AI policy (both terms, identical)

Use of AI tools. You may use generative AI tools (for example, ChatGPT, Claude, Gemini, or Copilot) on homework and other work completed outside of class. You are responsible for the accuracy and quality of everything you submit. AI tools and other electronic aids are not permitted during in-class quizzes, exams, or the final exam unless I explicitly say otherwise. Using them there is academic misconduct under [institution]'s policy.

B.3 Final exam (both terms, identical)

Final exam. The final exam is given [in class / in the testing center] on [date]. Exams are not returned. You may review your graded exam with me during office hours. Copies, photos, and notes are not permitted.

B.4 Study notice (both terms, identical)

Research study. This course is participating in a multi-institution study of grading practices in economics courses. You will receive an email from Cornell University with a link to a short survey near the beginning and end of the term. Participation is voluntary, does not affect your grade, and your individual responses are not shared with your instructor. [Replace this paragraph with the IRB-approved text supplied by the study team before use.]

B.5 What not to put in the syllabus

- Do not mention the two arms, the word "control" or "treatment," the study hypotheses, or that the final exam is reused.
- Do not add AI-detection language, disclosure requirements, or AI-use penalties for outside-of-class work.
- Do not attach credit to survey completion unless instructed by the study team.

Appendix C. Deviation log

Record anything that differs from the locked design in either term. Examples include a cancelled or rescheduled assessment, a changed due date or policy, an exam question corrected, a platform change, or a suspected exam compromise. Send this log with your end-of-term data, even if it is empty.

Date	Term / arm	What changed	Why	Reported on

Course: _____ Instructor: _____