Generative AI's Impact on Student Achievement and Implications for Worker Productivity

Naomi Hausman
Hebrew University

Oren Rigbi
Hebrew University

Sarit Weisburd
Hebrew University

January 26, 2026

AI is changing work AND learning

- Many jobs and tasks are affected by AI
  - At least 10% of tasks for 80% of workers (Eloundou et al. (2024))
  - 2024: 39% of US 18-64 pop using gen AI (Bick et al. (2024))
- Research focus on today's workers: AI increases productivity
  - Occupational exposure/ applicability: coding, coding, coding (Bick et al. (2024); Handa et al. (2025); Tomlinson et al. (2025))
  - Distributional effects:
    - AI tends to boost productivity most for less-skilled workers (Noy and Zhang (2023); Brynjolfsson et al. (2025); Kanazawa et al. (2025); Cui et al. (2025))
    - AI boosts most-skilled entrepreneurs in Kenya (Otis et al. (2024))
- Current workers were originally educated without AI
- Student use of AI is redefining the skills and knowledge of future workers

AI is changing work **AND learning**

- Many questions re **future workers:**
  - Which jobs to train for? What skills are/will be needed?
  - How to educate when AI is available? (~~The dog ate my~~ AI did my homework)
  - How to signal skill to employers?
- Student AI adoption could affect skills and job matches of future workers
  - **AI-HC:** Improve AI-related skills, increase future workplace productivity
  - **Basic HC:** Substitute for essential learning, create knowledge gaps
  - **Grade signal:** Weaken, obfuscate underlying ability, complicate job matching
- **RQ:** How does AI availability affect student performance and signaling?
  - Admin data on courses, grades (top Israeli university)
  - Natural, quasi-random variation in AI availability (time, cross-course)
  - After vs. before ChatGPT, courses more/less amenable to AI use
- Find: increase in average performance, compression of distribution

Motivation: education design and workforce productivity

- Educational programming and assessment: balance objectives
  - Training in technology to be used on-the-job (e.g. AI iteration)
  - Harvest HC-AI complementarities
  - Risk of over-reliance on AI tools to substitute for learning
- Tomorrow's workforce productivity
  - Reduced signal value of grades → alt screening devices, reduced job match quality
  - HC development challenge (AI-specific and basic)
    - AI use is productive (Cui et al. (2025); Babina et al. (2024); Gofman and Jin (2024))
    - AI could spur innovation, knowledge accumulation (Cockburn et al. (2018))
    - But most job tasks still not amenable to AI! (Eloundou et al. (2024))
  - At firm level: human expertise and skilled management complement big data, computing technology (McElheran et al. (2025))
- Productivity heterogeneity → income inequality (Acemoglu (2025); Agrawal et al. (2023))

How is AI used in education?

- Widely available tools (use what's out there): ChatGPT/ Gemini/ Claude/ Copilot, etc.
- Existing tools with restrictions (possible in RCTs...)
  - No copy/paste
  - No solution seeking (can only talk through ideas)
- Specialized tools, designed to improve learning
  - Custom-made chatbots (tailored to purpose, e.g. Hebrew U Med School)
  - Personalized learning platforms (tailored to student)
- AI forward + soft skills (Alpha School: AI-powered private K-12)
- Tools for teachers

Does AI help students? Hard question

- What is "helping?"
  - Better performance on schoolwork?
    - When AI is available or unavailable?
    - When AI is available, performance reflects HC + AI
  - Productivity in education $\neq$ productivity in labor market
    - (producing output vs. developing HC)
  - Better performance on tasks resembling job tasks?
    - But what are job tasks now that AI is available on the job?
    - What will job tasks be in 5, 10, 20 years?
    - Obviously can't rely on teaching to tasks; need basic skills
  - Ultimate goal: learn more / develop HC (basic and AI-specific)
- Measurement challenge:
  - Short run vs. long run and new problems
  - Output vs. skills
  - Students' opinions vs. professors'

Does AI help students? Evidence (mostly) from RCTs

- Two kinds of studies
  1. Randomize access to (general) AI tools when learning
  2. Distinguish btw effects of augmentation vs. automation AI tools
- Unrestricted AI tools that offer solutions: (Bastani et al. (2025); Lehmann et al. (2024))
  - Improve performance with tool (Nie et al. (2025))
  - Hurt performance when tool removed (harm learning) (Darvishi et al. (2024))
- Restricted AI / tutor: improve performance and learning (or break even) (Bastani et al. (2025); Lehmann et al. (2024))
- Students less experienced with AI more prone to replacement (Lehmann et al. (2024))
- We take a complementary approach
  - Novel empirical strategy, real-world: natural, quasi-random variation
  - Large, comprehensive administrative dataset: multiple degrees, many subjects
- Results entirely consistent with RCTs! ▸ Worker Evidence

Empirical Strategy: Main Analysis

- DID: Compare grades after vs. before **ChatGPT release (Nov 2022)** in:
  - **AI-compatible** (treated) courses (essays, at-home projects)
  - **AI-incompatible** courses (labs, in-class exams)
- Course-level treatment facilitates within-student comparison
- Key DID Assumption: AI-C and AI-NC courses would follow **parallel grade trends**
  - Evidence on pre-treatment trends from 4 years of pre-ChatGPT data
- Propensity Score Matching (PSM):
  - Assume selection on observables into treatment
  - Compare outcomes in similar courses that nevertheless got different treatment
- Data: Admin grades, syllabi, surveys of faculty and students ▸ Data

Empirical Strategy: Main Analysis

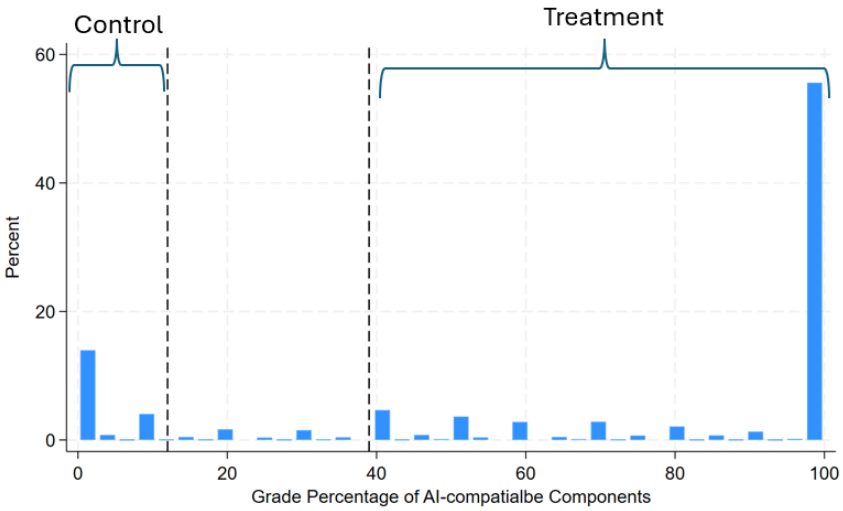
- DID: Compare grades after vs. before **ChatGPT release (Nov 2022)** in:
  - **AI-compatible** (treated) courses (essays, at-home projects)
  - **AI-incompatible** courses (labs, in-class exams)
- Course-level treatment facilitates within-student comparison
- Key DID Assumption: AI-C and AI-NC courses would follow **parallel grade trends**
  - Evidence on pre-treatment trends from 4 years of pre-ChatGPT data
- Propensity Score Matching (PSM):
  - Assume selection on observables into treatment
  - Compare outcomes in similar courses that nevertheless got different treatment
- Data: Admin grades, syllabi, surveys of faculty and students ▸ Data

Defining AI-compatible Courses (Treatment)

AI-Compatible Components	AI-Incompatible Components
Presentations	Lab work
At-home coursework	In-class exams
At-home exams	

- **AI-compatible (Treatment):** $\geq$ 40% of grade determined by AI-compatible components
- **AI-incompatible (Control):** $\leq$ 10% of grade determined by AI-compatible components
- Focus on $\approx$ 37,000 students in $\approx$ 5,700 courses with these definitions

Distribution of AI-compatible Grade Components



DID Estimation: Does AI Availability affect Grades?

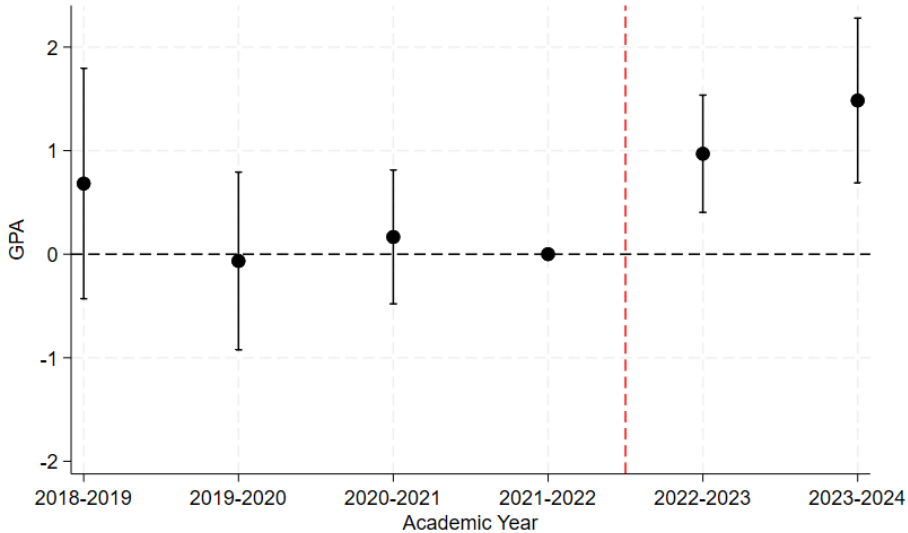
$$g_{ict} = \sum_{\substack{t=2019 \\ t \neq 2022}}^{2024} \delta_t \, \mathrm{D}_t \times \text{AI-Comp}_{ct} + \sum_{\substack{t=2019 \\ t \neq 2022}}^{2024} \gamma_t \, \mathrm{D}_t + \beta \text{AI-Comp}_{ct} + \theta' X_{ct} + \rho_i + \epsilon_{ict}$$

$$\text{AI-Comp}_{ct} = \begin{cases} 1 & \text{if} \geq 40\% \text{ of course grade is based on AI-compatible components} \\ 0 & \text{if} \leq 10\% \text{ of course-grade is based on AI-compatible components} \end{cases}$$

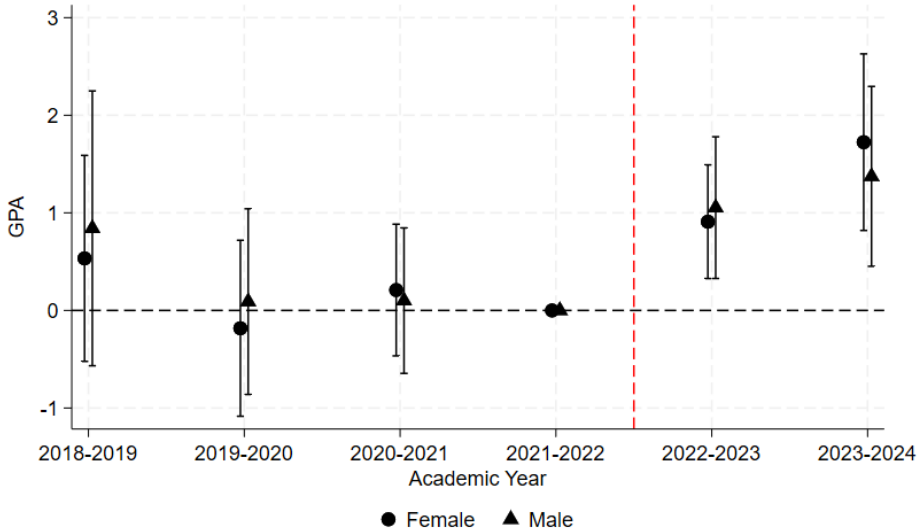
$$D_t = \begin{cases} 1 & \text{if course is completed in year } t \\ 0 & \text{otherwise} \end{cases}$$

X_{ct} – course chars: class size, dept, degree, credits, language, mandatory
ρ_i – student FE
ϵ_{ict} – unobserved time- and course- varying factors, clustered by course

AI Availability Effect on Grade Outcomes (0-100)



Do Men Experience a Greater AI Boost? No!

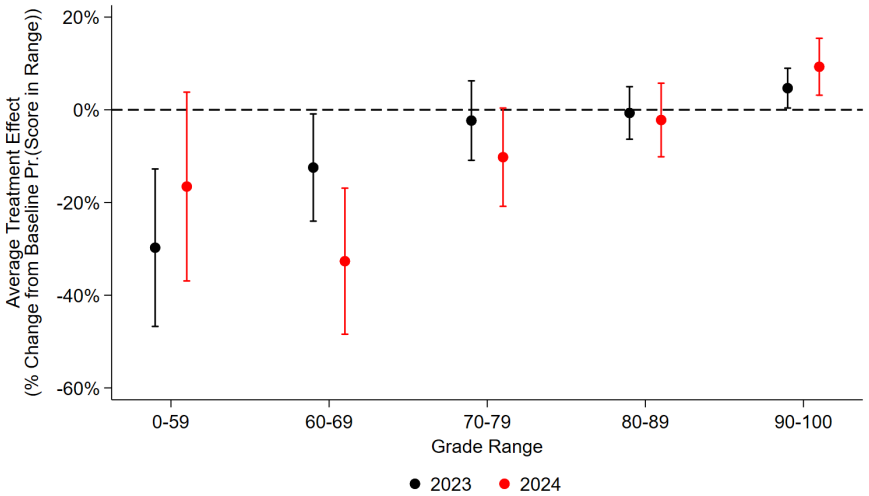


Heterogeneity in AI Effect on Grades

Dimension		No	Yes
Advanced Course	AI Compatible X 2022-2023	0.999	0.843**
		(0.742)	(0.355)
	AI Compatible X 2023-2024	0.935	1.466***
		(0.876)	(0.511)
STEM	AI Compatible X 2022-2023	0.769***	1.332
		(0.287)	(0.920)
	AI Compatible X 2023-2024	1.511***	1.269
		(0.409)	(0.969)
Business	AI Compatible X 2022-2023	0.769**	0.188
(vs. non-Business non-STEM)		(0.298)	(1.101)
	AI Compatible X 2023-2024	1.398***	0.522
		(0.416)	(1.328)
Standardized Grading	AI Compatible X 2022-2023	1.208**	0.525
		(0.512)	(0.532)
	AI Compatible X 2023-2024	1.587**	-1.350
		(0.705)	(0.842)

Effect on Probability of Scoring in Grade Range

30% decline in Pr(fail)



Interpretation

- AI availability (ITT) magnitude: 1-1.3 points (0-100 scale) are meaningful
  - AI availability Can push B+ student into A range
  - Driven by non-STEM, non-Business, non-curving, advanced courses
  - Biggest effects at bottom of distribution. Reduces signal value of grades
  - AND Not all students use AI even when it's available (take-up<100%)
- AI *use* effects (ATT) must be larger: 3.5 points in 2022-23 (B to A-)
  - AI availability an instrument for AI use
  - Estimate with Wald estimator: $LATE = \frac{ITT}{take-up}$ (Angrist et al. (1996))
- Both availability and use effects are policy-relevant parameters
  - ITT informs cost-effectiveness and scaling when AI is made available
  - ATT could help promote AI use and inform targeting to subpopulations
- Does the grade increase reflect more learning???

Human Capital Development in AI Regime

AI Experience Effects

- How does student AI use affect underlying HC development?
  - May replace key learning processes (problem solving, critical thinking, idea synthesis)
  - Practice may increase productive labor market use of AI
- Hard to measure underlying HC accumulation
  - When AI is available, grades reflect HC + AI
  - When AI is not available, grades should reflect HC
  - (Students now need AI-specific AND basic HC)
- Estimate effects of AI experience on subsequent performance with/without AI
- Compare students with vs. without AI exposure in 1st-year AI-compatible intro courses
  - Use students who enter univ before vs. after ChatGPT release (arguably identical)
  - Measure outcomes in 2nd year AI-compatible and AI-incompatible courses

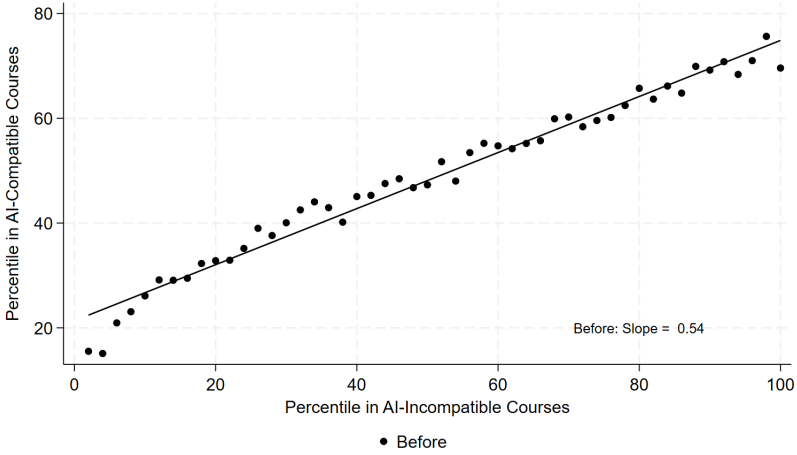
The Tradeoff Between Basic and AI-Specific Human Capital

	AI-Compatible Advanced Course Grade			AI-Incompatible Advanced Course Grade		
	(1)	(2)	(3)	(4)	(5)	(6)
2nd Year*AI Experience	1.194***	0.656***	0.928***	-9.232***	-0.431	-0.507
	(0.267)	(0.245)	(0.229)	(0.353)	(0.382)	(0.353)
Course Char. Controls	✗	✓	✓	✗	✓	✓
Student FE	✗	✗	✓	✗	✗	✓
R^2	0.029	0.227	0.555	0.080	0.246	0.548
N	22,806	22,806	22,806	34,829	34,829	34,829

- AI experience in early coursework boosts performance in AI-supported tasks but may undermine preparation for contexts requiring unaided human capital

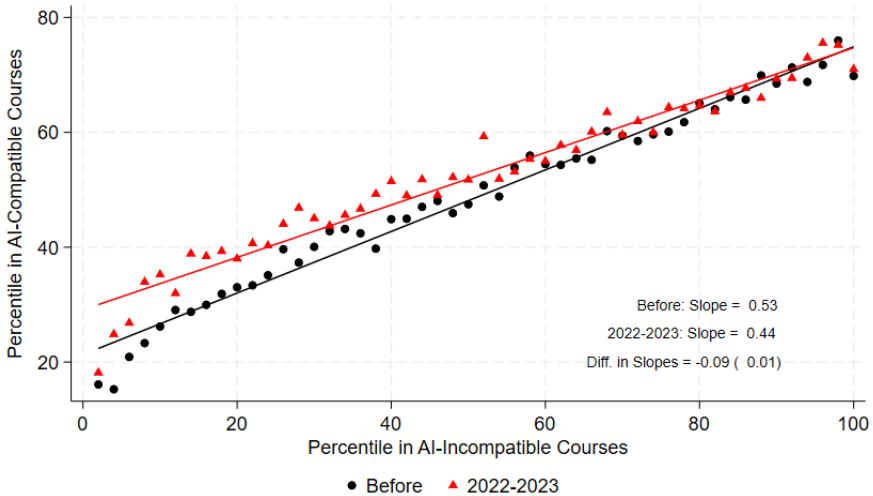
The Human Capital Tradeoff: Supportive Evidence

Rank-Rank Slope Before ChatGPT



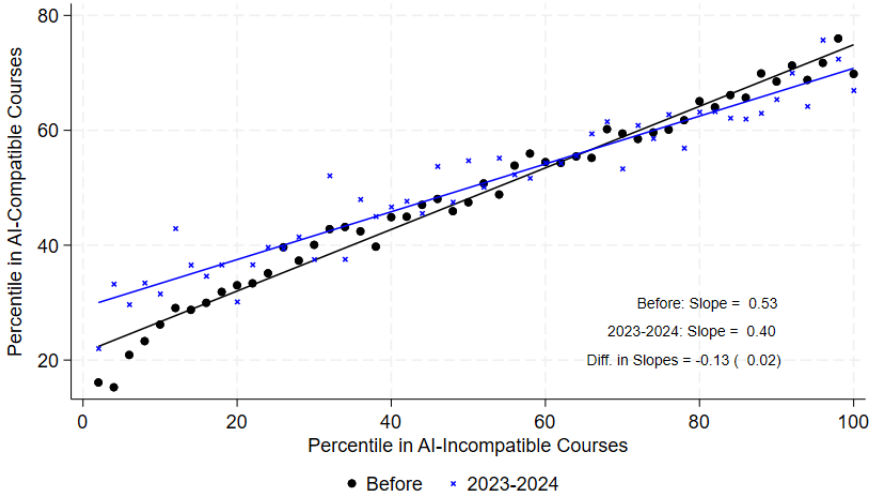
Weakening Link between HC and Grades

The Growth of Student Rank Inconsistencies Post-ChatGPT (2023 vs. 2022)



Weakening Link between HC and Grades

The Growth of Student Rank Inconsistencies Post-ChatGPT (2024 vs. 2022)



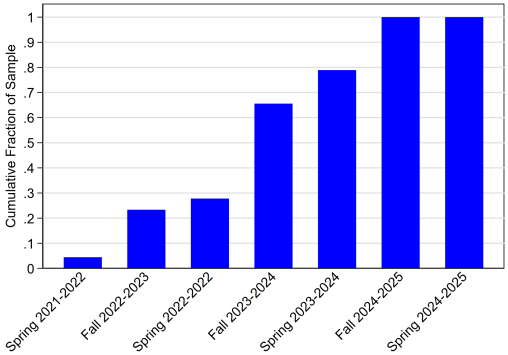
Student and Faculty Adjustment to AI Regime

and Workforce Implications

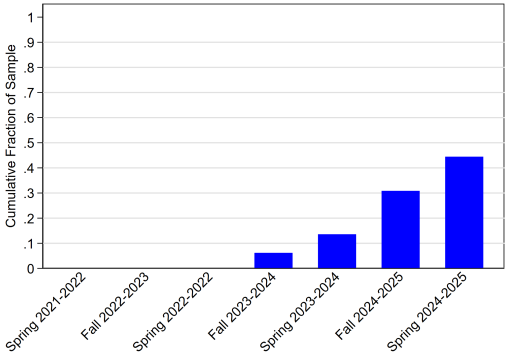
- Surveyed students and faculty on AI adoption for coursework and adjustment of courses
- Bottom line:
  - Students adopt quickly. 100% adoption by Fall 2024-25.
  - Faculty have barely adjusted.

Survey Data: Student AI Use is Rising Sharply

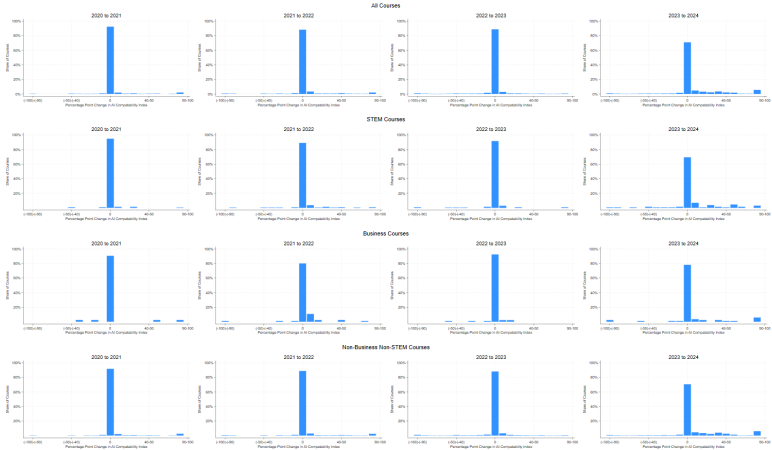
Students: Help with Coursework



Faculty: Course Prep or Evaluation



Composition of Assignments: Year-to-year Changes in AI-Compatibility of Courses by Academic Field

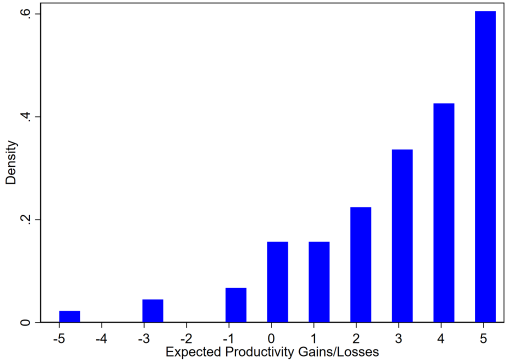


Disconnect in Faculty and Student Beliefs on AI

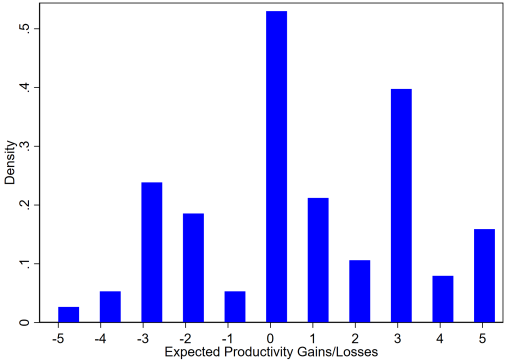
- Subjective beliefs $\neq$ actual productivity effects, but
- Faculty anticipate substantially lower AI productivity gains for students entering labor market than students do
  - 0.61 (faculty) vs. 2.96 (students) on [-5,5] scale
  - Teachers' greater understanding of basic HC gaps?
  - Students' greater understanding of what they can do with AI?

Expected Impact of Student AI Use on Productivity Gains at Work

Student Expectations



Faculty Expectations



Discussion & Ongoing Work

- Student access to AI is likely to shape skill set and productivity of future workforce
- Early causal evidence: AI availability in higher education improves average performance
  - Bifurcated learning effect: improves performance on future tasks with AI, undermines basic HC that drives performance when AI is absent
- Four main findings:
  1. AI availability: 0.6-1.3 point avg increase in grades, driven by bottom of dist
  2. Grade distribution compressed, reducing grades' signal value, potential job match quality and productivity
  3. AI experience has costs and benefits for HC development
  4. Faculty have mustered little response
- Implications for universities: design of coursework, assessments (mixed?), specialized AI tools, incentives for use, incentives for faculty adjustment
- Implications for employers: alt screening tools
- Ongoing: connecting to labor market outcomes in tax data

References I

Acemoglu, D. (2025). The simple macroeconomics of ai. *Economic Policy*, 40(121):13–58.

Agrawal, A., Gans, J. S., and Goldfarb, A. (2023). Do we want less automation? *Science*, 381(6654):155–158.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Babina, T., Fedyk, A., He, A., and Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., and Mariman, R. (2025). Generative ai without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26):e2422633122.

References II

Bick, A., Blandin, A., and Deming, D. J. (2024). The rapid adoption of generative ai. Technical report, National Bureau of Economic Research Working Paper no. 32966.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*, 140(2):889–942.

Cockburn, I. M., Henderson, R., and Stern, S. (2018). *The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis*, pages 115–146. University of Chicago Press.

Cui, Z. K., Demirer, M., Jaffe, S., Musolff, L., Peng, S., and Salz, T. (2025). The effects of generative ai on high skilled work: Evidence from three field experiments with software developers. *Available at SSRN 4945566*.

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., and Siemens, G. (2024). Impact of ai assistance on student agency. *Computers & Education*, 210:104967.

References III

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., and Lakhani, K. R. (2023). Navigating the jagged technological frontier: Field experimental evidence of the effects of ai on knowledge worker productivity and quality. Harvard Business School Technology & Operations Mgt. Unit Working Paper no. 24-013.

Doshi, A. R. and Hauser, O. P. (2024). Generative ai enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28):eadn5290.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2024). Gpts are gpts: Labor market impact potential of llms. *Science*, 384(6702):1306–1308.

Gofman, M. and Jin, Z. (2024). Artificial intelligence, education, and entrepreneurship. *The Journal of Finance*, 79(1):631–667.

References IV

Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., Mueller, J., Hong, J., Ritchie, S., Belonax, T., et al. (2025). Which economic tasks are performed with ai? evidence from millions of claude conversations. arXiv preprint arXiv:2503.04761.

Kanazawa, K., Kawaguchi, D., Shigeoka, H., and Watanabe, Y. (2025). Ai, skill, and productivity: The case of taxi drivers. *Management Science*.

Lehmann, M., Cornelius, P. B., and Sting, F. J. (2024). Ai meets the classroom: When does chatgpt harm learning? Available at SSRN: http://dx.doi.org/10.2139/ssrn.4941259.

McElheran, K., Yang, M.-J., Kroff, Z., and Brynjolfsson, E. (2025). The rise of industrial ai in america: Microfoundations of the productivity j-curve(s). Technical report.

References V

Nie, A., Chandak, Y., Suzara, M., Malik, A., Woodrow, J., Peng, M., Sahami, M., Brunskill, E., and Piech, C. (2025). The gpt surprise: Offering large language model chat in a massive coding class reduced engagement but may increase adopters' exam performances. *Proceedings of the Twelfth ACM Conference on Learning@ Scale*, pages 376–380.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Otis, N., Clarke, R., Delecourt, S., Holtz, D., and Koning, R. (2024). Does generative ai crowd out social learning? evidence from a field experiment with entrepreneurs. Technical report, UC Berkeley Working Paper.

Tomlinson, K., Jaffe, S., Wang, W., Counts, S., and Suri, S. (2025). Working with ai: Measuring the occupational implications of generative ai.

Data Collected

- **Admin data:** student-course level (2018-2019 – 2023-2024)
  - Grades (exam, final) for all BA, MA, and MBA student-course obs
  - 4 yrs before ChatGPT release (November 2022), 2 years after
  - **Students**: primary field, sex, age, and immigration year
  - **Courses**: format (e.g. lecture/lab), mandatory/elective, department, professor
- **Syllabi:** Grade composition (assignments), text descrip
- **Surveys of faculty and students**: demographics, usage rates, beliefs about AI
  - Conducted spring 2025
  - Same university from which admin data originates
  - Assess student take-up, faculty response, possible workforce implications

▸ Empirical Strategy

Mean Treated and Control Course Characteristics

Variable	**Treatment** (AI-compatible)	**Control** (AI-incompatible)
Course Characteristics		
Avg. Grade	90	84
Class Size	36	78
Mandatory Class	44%	63%
Fields		
Humanities	22%	10%
Natural Sciences	15%	22%
Social Sciences	22%	12%
Agriculture	13%	34%
Law	5%	7%
Social Work	9%	6%
Engineering	6%	1%
Students		
Age	27	26
Male	40%	43%
Num. Course-Semester Obs	4,087	2,640
Num. Student-Course-Semester Obs	130,739	184,597

Propensity Score Matching Procedure

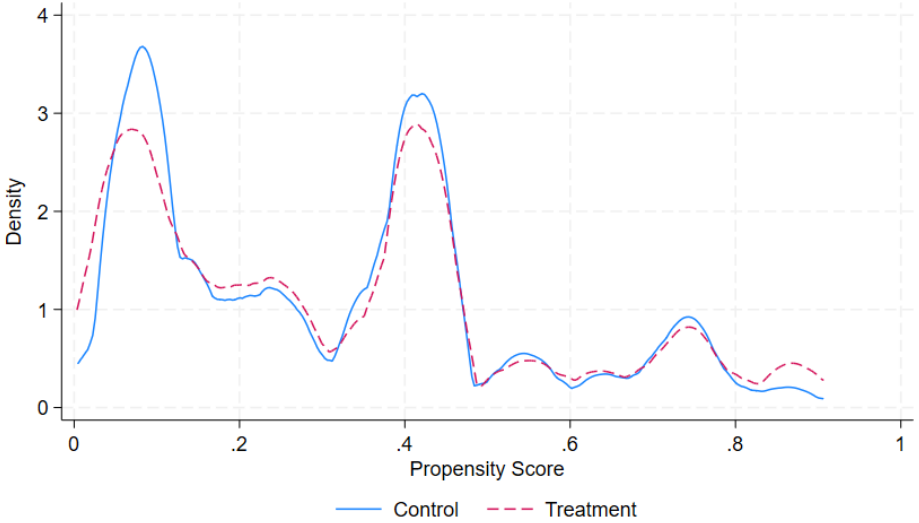
- Compare treated courses to similar-on-observables control courses
- Step 1: Estimate Pr(AI-Compatible) as a function of ex-ante course observables
  - Class size, type (lecture/lab/seminar), mandatory, language, degree, dept
- Step 2: Match AI-C to AI-NC courses with similar PS, NN-1 within faculty
  - Generate "PSM Sample" of matched courses in common support
  - Each matched pair gets unique match group ID
- Step 3: Estimate DID with match group FE and/or PS control
  - Match group ID ensures treated courses compared to own matched control
  - PS control absorbs remaining diffs in Pr(treat) within matched pairs

Propensity Score Matching DID Estimation

$$g_{ict} = \sum_{\substack{t=2019 \\ t \neq 2022}}^{2024} \delta_t \, \mathrm{D}_t \times \text{AI-Comp}_{ct} + \sum_{\substack{t=2019 \\ t \neq 2022}}^{2024} \gamma_t \, \mathrm{D}_t + \beta \text{AI-Comp}_{ct} + \theta' X_{ct} + \rho_i + PS_c + \phi_g + \epsilon_{ict}$$

X_{ct} – course chars: class size, dept, degree, credits, language, mandatory
ρ_i – student FE
PS_c – propensity score control
ϕ_g – match group FE
ϵ_{ict} – unobserved time- and course-varying factors, clustered by course

PSM Sample: Match Quality



Effect of AI Availability on Student Grades

	Baseline DID		PS Matching DID			
Dependent Variable: Grade (0-100)	(1)	(2)	(3)	(4)	(5)	(6)
Before:						
2018-2019*AI-compatible	0.682	0.375	0.362	-0.286	0.135	0.117
	(0.567)	(0.858)	(0.856)	(0.850)	(0.812)	(0.816)
2019-2020*AI-compatible	-0.066	0.530	0.514	0.296	-0.023	-0.048
	(0.438)	(0.634)	(0.636)	(0.557)	(0.563)	(0.565)
2020-2021*AI-compatible	0.166	0.256	0.247	0.077	-0.162	-0.168
	(0.330)	(0.490)	0.491	(0.527)	(0.472)	(0.473)
After:						
2022-2023*AI-compatible	0.970***	0.942**	0.917**	0.809**	0.645*	0.633*
	(0.289)	(0.387)	(0.385)	(0.365)	(0.345)	(0.344)
2023-2024*AI-compatible	1.484***	1.334**	1.303**	0.745	0.566	0.532
	(0.406)	(0.540)	(0.538)	(0.545)	(0.486)	(0.483)
Course Char. Controls	✓	✓	✓	✗	✓	✓
Match Group FE	✗	✗	✗	✓	✓	✓
P-Score Control	✗	✗	✓	✗	✗	✓
R^2	0.477	0.554	0.555	0.565	0.581	0.581
N	500,611	361,591	361,591	361,591	361,591	361,591

Dosage Effect: Continuous Treatment Intensity

	(1) Fixed Intensity	(2) Flexible Intensity
Before:		
2018-2019*AI Compatability Index	0.008	0.012*
	(0.007)	(0.007)
2019-2020*AI Compatability Index	0.005	0.005
	(0.005)	(0.005)
2020-2021*AI Compatability Index	0.001	0.002
	(0.004)	(0.004)
After:		
2022-2023*AI Compatability Index	0.006*	0.007**
	(0.003)	(0.003)
2023-2024*AI Compatability Index	0.003	0.011**
	(0.004)	(0.004)
R^2	0.482	0.483
N	442,992	442,992

Heterogeneity in AI Effect on Grades

Dimension		No	Yes
Male	AI Compatible X 2022-2023	0.909***	1.053***
		(0.297)	(0.370)
	AI Compatible X 2023-2024	1.724***	1.374***
		(0.462)	(0.470)
Young (< Age 26)	AI Compatible X 2022-2023	0.933***	1.132***
		(0.306)	(0.358)
	AI Compatible X 2023-2024	1.287***	2.079***
		(0.398)	(0.537)
Senior Professor	AI Compatible X 2022-2023	0.791**	1.093**
		(0.392)	(0.486)
	AI Compatible X 2023-2024	1.613***	1.105*
		(0.578)	(0.576)

Summary Statistics from Student Survey

Panel A: Statistics from Survey of Students			
Variable	**Mean**	**SE**	**N**
Demographics			
Male	0.52	0.05	91
Jewish	0.86	0.04	91
Age	26.21	0.50	91
New Immigrant	0.29	0.05	91
English Survey Version	0.19	0.04	91
Survey Responses			
Early Adopter (by July 2023)	0.28	0.05	90
Adopter (by July 2024)	0.79	0.04	91
AI Use: Percent of Assignments	60.66%	2.78%	91
Usage Reduces Effort by at least 50%	47.25%	3.27%	91
Beliefs about AI			
AI Replaces Learning (Scale 1-5)	2.63	0.13	91
Importance of Lost Knowledge (Scale 1-5)	2.38	0.11	91
Expected Access to AI for Job Tasks (Scale 1-5)	4.30	0.09	91
Anticipated effect of AI in coursework on students' labor market productivity (Scale -5 to 5)	2.96	0.22	91

AI Availability Effects (ITT) and Use Effects (ATT) on Grades

Dependent Variable: Grade (0-100)	2022-2023 Usage Rate	2022-2023 ITT	2022-2023 ATT	2023-2024 Usage Rate	2023-2024 ITT	2023-2024 ATT
Full Sample	27.7	0.970***	3.489***	78.9	1.484***	1.881***
Gender						
Female	22.7	0.909***	4.004**	72.7	1.724***	2.371***
Male	32.6	1.053***	3.23***	84.8	1.374***	1.62***
Age						
Age>=26	30.2	0.933***	3.089***	83.7	1.287***	1.538***
Age<26	25.5	1.132***	4.439***	74.5	2.079***	2.791***

Results: ITT and ATT Effects

- Early adopters experience large use effects: 3.5 points, B to A-
- Marginal adopters may be less effective users
- Higher usage rates for men, actually lower use effects
- Younger students adopt more slowly but seem to be more effective users

Robustness to Treatment Definition

- Main definition based on 10% and 40% thresholds of AI-Compatibility index
- Some grade components are of ambiguous AI-compatibility (e.g. presentations)
- Estimate AI effect using 3 alternative methods of defining treatment
  - Vary AI-compatibility index thresholds (6 alternatives)
  - Vary inclusion of home grade components
  - LLM classifications based on syllabus course description
    - Core concepts (neg corr w/AI-Compatibility)
    - Reading and writing (pos corr w/AI-Compatibility)

Alternative AI-Compatibility Index Thresholds

Dependent Variable: Grade (0-100)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Before:							
2018-2019*Treat	0.682	0.324	0.643	0.352	0.425	0.928*	0.883
	(0.567)	(0.530)	(0.623)	(0.492)	(0.499)	(0.534)	(0.579)
2019-2020*Treat	-0.066	-0.205	-0.030	0.013	0.134	0.144	0.149
	(0.438)	(0.414)	(0.445)	(0.468)	(0.481)	(0.390)	(0.427)
2020-2021*Treat	0.166	0.050	0.066	-0.156	0.102	0.438	0.402
	(0.330)	(0.326)	(0.341)	(0.352)	(0.360)	(0.299)	(0.337)
After:							
2022-2023*Treat	0.970***	0.913***	0.755***	0.784***	0.780***	0.635**	0.987***
	(0.289)	(0.280)	(0.281)	(0.285)	(0.297)	(0.254)	(0.284)
2023-2024*Treat	1.484***	1.205***	1.644***	1.407***	1.428***	1.187***	1.539***
	(0.406)	(0.405)	(0.406)	(0.437)	(0.445)	(0.357)	(0.392)
AI-compatibility Index Thresholds							
Control (Upper)	10	10	10	10	10	30	15
Treatment (Lower)	40	30	50	60	70	40	40
R^2	0.477	0.471	0.480	0.485	0.489	0.472	0.476
N	500,611	543,442	475,756	452,800	439,900	639,005	533,505

Alternative Definitions of AI Compatibility Using Grade Components

	Dependent Variable: Grade (0-100)				
Treatment Definition:	(1) Has Home Exam	(2) No In-Class Exam	(3) Presentation in Control	(4) ≥20% Home Components	(5) Control: ≤ 10% Home Comps
Before:					
2018-2019*AI-compatible	0.270	0.505	0.354	-0.692	-0.676
	(0.542)	(0.497)	(0.569)	(0.570)	(0.559)
2019-2020*AI-compatible	-0.361	0.118	-0.023	-0.113	-0.001
	(0.438)	(0.468)	(0.361)	(0.340)	(0.322)
2020-2021*AI-compatible	-0.452	0.225	0.186	-0.042	0.030
	(0.334)	(0.342)	(0.299)	(0.308)	(0.284)
After:					
2022-2023*AI-compatible	0.787***	0.734**	0.659**	0.576**	0.487**
	(0.290)	(0.295)	(0.263)	(0.257)	(0.247)
2023-2024*AI-compatible	1.294***	1.495***	1.635***	1.026**	0.793**
	(0.419)	(0.420)	(0.378)	(0.404)	(0.384)
R^2	0.495	0.494	0.467	0.466	0.464
N	397,637	433,270	553,631	620,355	669,291

Alternative Definition of AI-Compatibility based on LLM Classifications

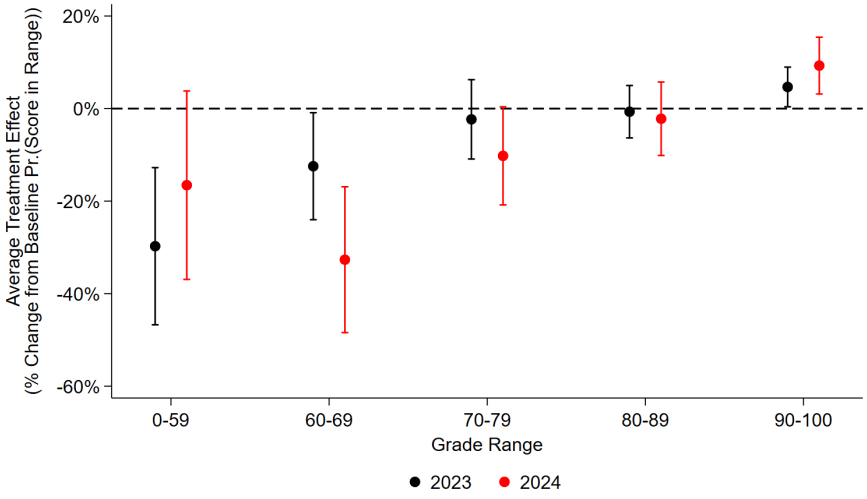
	(1) Core Concepts	(2) Reading & Writing	(3) Keywords	(4) 1^{st} PC
Before:				
2018-2019*AI-Compatibility LLM Index	0.472	-0.202	0.152	0.261
	(0.301)	(0.250)	(0.372)	(0.304)
2019-2020*AI-Compatibility LLM Index	-0.013	-0.176	-0.296	-0.213
	(0.221)	(0.171)	(0.223)	(0.210)
2020-2021*AI-Compatibility LLM Index	0.002	-0.151	0.118	0.007
	(0.164)	(0.138)	(0.185)	(0.170)
After:				
2022-2023*AI-Compatibility LLM Index	0.507***	0.348***	0.337**	0.501***
	(0.152)	(0.119)	(0.166)	(0.149)
2023-2024*AI-Compatibility LLM Index	-0.046	0.151	-0.064	0.009
	(0.200)	(0.176)	(0.215)	(0.207)
R^2	0.505	0.501	0.502	0.503
N	396,465	396,465	396,465	396,465

Note: Classification fixed as of 2020-2021.

Estimating Distributional Effects

- So far: increase in average grade, heterogeneity across contexts
- Care also about heterogeneity over grade distribution
  - Does AI help the strongest or weakest students the most?
  - Potential for persistence into labor market, changes in income inequality
  - Spread of grades facilitates signaling to employers, grad programs
- Estimate Equation 1 with binary dep var for each grade range
- Estimate DID at course-semester level predicting grade at 25^th, 50^th, 75^th pctiles

Effect on Probability of Scoring in Grade Range



Distributional Effects: DID Estimation

$$g_{ct}^{p} = \sum_{\substack{t=2019 \\ t\neq 2022}}^{2024} \alpha_t \, \mathrm{D}_t \times \text{AI-Comp}_{ct} + \sum_{\substack{t=2019 \\ t\neq 2022}}^{2024} \rho_t \, \mathrm{D}_t + \omega \text{AI-Comp}_{ct} + \Omega' X_{ct} + \epsilon_{ct}$$

g_{ct}^{p} – grade at percentile p in course c and year t
X_{ct} – course chars: class size, dept, degree, credits, language, mandatory
ϵ_{ct} – idiosyncratic error, clustered by course

Effect of AI Availability on the Grade Distribution

Dependent Variable: Grade at:	(1) p25	(2) p50	(3) p75
Before:			
2018-2019*AI-compatible	0.875	-0.157	-0.694
	(1.077)	(0.784)	(0.678)
2019-2020*AI-compatible	0.081	-0.443	-0.594
	(1.114)	(0.788)	(0.705)
2020-2021*AI-compatible	1.620	0.270	-0.507
	(1.206)	(0.893)	(0.523)
After:			
2022-2023*AI-compatible	3.014***	1.272**	0.738
	(0.933)	(0.612)	(0.453)
2023-2024*AI-compatible	2.383**	1.125	0.583
	(0.978)	(0.752)	(0.511)
Dep. Var. Avg.	84.72	90.18	94.14
R^2	0.426	0.401	0.312
N	10,076	10,076	10,076

Distributional Effects

- Reduced probability of failing or getting a low grade
- Largest increase in performance at the bottom of the distribution
  - Consistent with results in many work contexts (Noy and Zhang (2023); Brynjolfsson et al. (2025); Kanazawa et al. (2025); Cui et al. (2025))
  - Consistent with results in learning contexts, though learning is impaired (Bastani et al. (2025); Lehmann et al. (2024))
- Compression of distribution reduces signal value of grades
- Costs to employer screening, job match quality

Estimating the AI Experience Effect

$$g_{ict}^{AI\text{-}comp} = \alpha_1 I_{ct}^{2ndYear} * I_i^{AIexper} + \theta_1' X_{ct} + \rho_{1i} + \varepsilon_{ict} \quad (1)$$

$$g_{ict}^{AI\text{-}incomp} = \alpha_2 I_{ct}^{2ndYear} * I_i^{AIexper} + \theta_2' X_{ct} + \rho_{2i} + \varepsilon_{ict} \quad (2)$$

$I_{ct}^{2ndYear}$ – 1 if the course is an advanced course being taken in the second year and 0 otherwise
$I_i^{AIexper}$ – 1 if the student was treated (took at least one AI-compatible introductory course during his first year) and 0 otherwise
X_{ct} – course chars: class size, dept, degree, credits, advanced course, language, mandatory
ρ_{1i} and ρ_{2i} – student FE, incl main effect of being treated
ϵ_{ict} – idiosyncratic error, clustered by course

AI and the Link between HC and Performance

Rank-Rank Analysis (Supportive Evidence)

- Do student rankings in AI-incompatible courses (where grades reflect HC) predict their rankings in AI-compatible courses (where grades reflect HC+AI)?
- Does this predictive power change after the introduction of ChatGPT?
- Hypotheses:
  - If rank relationship remains stable: performance is driven largely by HC which is likely to be highly correlated across courses.
  - If ranking gaps increase: AI may be inflating performance for some students in AI-compatible courses, weakening the link between HC and ranking.

Student and Faculty Adjustment to AI Regime

and Workforce Implications

- Ultimately want to understand workforce effects of HC changes
  - Ability to cope with job tasks more/less amenable to AI assistance
  - HC changes depend, in part, on how faculty respond to AI presence
- Faculty could adjust in at least 3 primary ways
  1. Change *composition* of assignments to be less AI-compatible (in-person grade points)
  2. Change *content* of at-home assignments to *reduce AI replacement*
  3. Change content of assignments to *encourage AI augmentation*, learning using AI. Could encourage both basic and AI-specific HC development
- What have faculty done thus far? Where does student adoption stand?
- Workforce implications?
- Use survey, admin data to assess

Margins of Faculty and Student Adjustment

- **Faculty** self-reported adjustment on margins above:
  1. Nearly 1/3: increased in-person graded work (ummm, not in admin data!)
  2. 2/5 adjusted coursework to impede AI replacement
  3. 1/5 encourage AI augmentation
- Faculty: Nearly 40%: no course changes. 78%: students would benefit from further changes
- **Students'** self-reported adjustment:
  - 79% adoption by July 2024; 100% adoption by spring 2025
  - Use on 61% of home assignments
  - Reduces effort by at least 50% on nearly 50% of home assignments
- Slow faculty response $\rightarrow$ AI effect on grades and HC development largely unchecked

Summary Statistics from Student Survey

Variable	Mean	SE	N
Demographics			
Male	0.52	0.05	91
Jewish	0.86	0.04	91
Age	26.21	0.50	91
New Immigrant	0.29	0.05	91
English Survey Version	0.19	0.04	91
Survey Responses			
Early Adopter (by July 2023)	0.28	0.05	90
Adopter (by July 2024)	0.79	0.04	91
AI Use: Percent of Assignments	60.66%	2.78%	91
Usage Reduces Effort by at least 50%	47.25%	3.27%	91
Beliefs about AI			
AI Replaces Learning (Scale 1-5)	2.63	0.13	91
Importance of Lost Knowledge (Scale 1-5)	2.38	0.11	91
Expected Access to AI for Job Tasks (Scale 1-5)	4.30	0.09	91
Anticipated effect of AI in coursework on students' labor market productivity (Scale -5 to 5)	2.96	0.22	91

Summary Statistics from Faculty Survey

Variable	Mean	SE	N
Demographics			
Male	0.41	0.06	79
Age	43.75	0.83	75
Tenure Track	0.81	0.04	81
Survey Responses			
Currently Using Paid Version of AI (June 2025)	0.54	0.06	81
Early Adopter for Teaching (by July 2023)	0	0	81
Adopter for Teaching (by July 2024)	0.14	0.04	81
Adopter for Teaching (by June 2025)	0.44	0.06	81
Course Changes Made to Address Student AI Use			
No Changes	0.38	0.05	81
Made Coursework More Difficult to do with AI	0.37	0.05	81
Adjusted Coursework to Improve Learning with AI	0.19	0.04	81
Increased In-Class Work	0.30	0.05	81
Beliefs on AI, Learning, and the Labor Market			
Adjustments for AI sufficient to assess students	0.44	0.07	50
Students would benefit from further assignment changes	0.78	0.06	50
Anticipated effect of AI in coursework on students' labor market productivity (-5 to 5)	0.61	0.29	77

Does AI help workers? Mostly yes, so far

- Adoption
  - At least 10% of tasks for 80% of workers (Eloundou et al. (2024))
  - 2024: 39% of US 18-64 pop using gen AI (Bick et al. (2024))
- Occupational exposure / applicability: coding, coding, coding (Bick et al. (2024); Handa et al. (2025); Tomlinson et al. (2025))
- Distributional effects
  - AI tends to boost productivity most for less-skilled workers (Noy and Zhang (2023); Brynjolfsson et al. (2025); Kanazawa et al. (2025); Cui et al. (2025))
  - AI boosts most-skilled entrepreneurs in Kenya (Otis et al. (2024))
- AI worse on some tasks. Nuanced judgment (Dell'Acqua et al. (2023)), variety of ideas (Doshi and Hauser (2024))

▸ Student Evidence

AI is changing work AND learning

- Many jobs and tasks are affected by AI
  - At least 10% of tasks for 80% of workers (Eloundou et al. (2024))
  - 2024: 39% of US 18-64 pop using gen AI (Bick et al. (2024))
- Research focus on today's workers (educated without AI)
  - Occupational exposure/ applicability: coding[3]
    (Bick et al. (2024); Handa et al. (2025); Tomlinson et al. (2025))
  - Distributional effects: largest boost for less-skilled workers (Noy and Zhang (2023); Brynjolfsson et al. (2025); Kanazawa et al. (2025); Cui et al. (2025))
- Many questions re **future workers:** Which jobs to train for? Skills needed? How to educate in AI regime, signal skill?
- RQ: How does AI availability affect student performance and signaling?

AI, Student Performance, Productivity of Future Workers and Firms

- **"80% of workers** have occupations where generative AI is relevant for at least **10% of tasks"** (Eloundou et al., 2024)
- How will AI adoption impact future workers?
  - Early adoption may improve **future performance on the job**. (Dell'Acqua et al. (2023); Acemoglu (2024))
  - Or AI could act as a **crutch**, inflating student grades without enhancing HC (Bastani et al. (2024); Lehmann, Cornelius and Sting (2024); Darvishi et al. (2024)).
- **The signal value** of a grade may change with the availability of AI tools
  - A mismatch between ability and grades could reduce **job match quality**
- Research question: How does AI availability affect student performance and grades' signal value?
- Main findings: Grades increase, driven by bottom quartile; compression of distribution

AI, Student Performance, Productivity of Future Workers and Firms

- **"80% of workers** have occupations where generative AI is relevant for at least **10% of tasks"** (Eloundou et al., 2024)
- How will AI adoption impact future workers?
  - Early adoption may improve **future performance on the job**. (Dell'Acqua et al. (2023); Acemoglu (2024))
  - Or AI could act as a **crutch**, inflating student grades without enhancing HC (Bastani et al. (2024); Lehmann, Cornelius and Sting (2024); Darvishi et al. (2024)).
- **The signal value** of a grade may change with the availability of AI tools
  - A mismatch between ability and grades could reduce **job match quality**
- Research question: How does AI availability affect student performance and grades' signal value?
- Main findings: Grades increase, driven by bottom quartile; compression of distribution