NET WORK LISER

Discussion Paper Series

IZA DP No. 18522

April 2026

AI-Enhanced Test Preparation and Student Performance: Evidence from an Introductory Economics Class

Stefani Milovanska-Farrington
University of Tampa and IZA@LISER

Caleb Tomberlin
William and Mary

IZA@LISER NETWORK
ISSN: 2365-9793

IZA DP No. 18522 April 2026

AI-Enhanced Test Preparation and Student Performance: Evidence from an Introductory Economics Class

Abstract

The emergence of artificial intelligence (AI) tools has offered new ways of teaching and learning. Businesses have also highlighted the importance of AI literacy in the workplace as AI is transforming operations and entire industries. Given the importance of obtaining AI skills and the opportunities it provides, it is useful for students to gain experience interacting with the emerging technology. Yet, the optimal ways to incorporate AI in each class so that students' knowledge acquisition in coursework does not suffer are still unclear. This paper examines the causal effect of AI-enhanced test preparation on student performance in an economics class. In a difference-indifferences framework, we compare the changes in students' test scores after relative to before utilizing AI to enhance learning between students who completed a guided AI assignment to prepare and those who did not. The findings provide evidence that learning through AI does not necessarily improve students' performance on formal exams. This does not mean that students should not learn how to use AI tools, but rather that they may not prepare for exams in all courses while simultaneously improving their AI skill set.

JEL classification

A20, A22, I21

Keywords

artificial intelligence (AI), AI assignment, ChatGPT, learning tools, student performance, test preparation

Corresponding author

Stefani Milovanska-Farrington
smilovanska@ut.edu

1. Introduction

Students' learning styles constantly evolve, and educators need to adapt to respond to the needs of new generations of students. One of the newest trends has been triggered by the emergence of artificial intelligence (AI) tools. While offering numerous opportunities for both teachers and learners, these tools have made educators anxious about the risk of students utilizing AI to complete graded coursework, raising academic integrity violation concerns (Sullivan et al. 2023). Students have also expressed concerns about unintentional plagiarism[1]. However, businesses have been adopting AI technologies to improve efficiency, and have been highlighting the importance of gaining experience interacting with AI, especially for the generations that are currently entering the labor force or would be doing so in the next few years. A 2024 study showed that 83% of U.S. professionals believe that students need AI skills upon entering the workforce[2].

To prepare students for this changing world in which AI literacy is necessary to fulfill the needs of businesses, educators could intentionally incorporate interaction with AI tools into the curriculum and the activities they assign to their students. This could improve students' career opportunities for jobs that utilize emerging technologies, initiate an open conversation about the responsible use of AI tools, facilitate learning, and improve students' educational outcomes and experiences.

The purpose of this research is to explore the effectiveness of AI tools to improve exam preparation in an economics class. Specifically, we examine the causal effect of completing an assignment that required interaction with AI tools aimed at providing additional practice for an exam, on students' test performance. To collect data for this study, we asked a sample of students to utilize ChatGPT to generate additional practice questions and detailed solutions that resemble the ones offered by an instructor for preparation for an exam in an introductory economics class. A group of students were not assigned the treatment; another group was required to complete the AI assignment; and a third set of students was allowed to choose whether to use AI for their test preparation or not. AI-enhanced learning was applied only to the last exam for the semester. To estimate the effect of interest, we adopt a difference-in-differences method. It enables us to compare the test results of the students who enhanced their test preparation through AI tools to the group of students who did not use AI, before and after the use of AI as a learning aid.

One could expect that we would find evidence of the effectiveness of using AI tools to improve students' test performance. Interestingly, using data from an introductory Economics class, this was not the case. We discuss the implications and the generalization of these findings in the results section of the article.

The rest of the paper is structured as follows. Section 2 summarizes the relevant literature on AI and its use as a learning aid. We present the intervention and describe the data collection method in Section 3. The empirical approach is explained in Section 4. Section 5 summarizes the results. Sections 6 concludes the article.

[1] Are AI skills a key part of career preparation in college?
[2] Employers Want New Grads with AI Experience, Knowledge

2. Literature review

Research suggests that AI is an excellent learning tool. For example, at UniDistance Suisse, an AI tutor app was provided to help students learn in a psychology class. The tutor app was able to generate course-related questions based on each student's individual level and abilities. The results showed that the class that used the AI tutor achieved higher grades. More specifically, there was an average class grade improvement of up to 15 percentile points compared to a similar course that did not use the tutor app (Baillifard et al. 2024). Furthermore, Alamri (2024) highlights the power of AI in offering personalized learning, improving academic outcomes, and increasing student engagement (Alamri 2024). For example, there are AI-powered, adaptive learning systems that provide real-time feedback and customized learning and have been shown to increase student engagement and performance (Zawacki-Richter et al. 2019).

AI could also improve the efficiency of analyzing data and processing solutions, but it may lack the understanding and creativity of a human (Luckin et al. 2016). This emphasizes the importance of implementing balanced AI integration in which AI complements rather than completely replaces human interaction (Wu 2023). Ravenswood School is an example of a program that emphasizes the importance of integrating AI into education responsibly. While much of the education sector is avoiding the use of AI, Ravenswood has found positive results from incorporating AI into its teaching. Educators' productivity has increased since they started to use AI to create innovative assignments, along with tailoring lesson plans. Additionally, students have learned the limitations and practicalities of AI, helping them to keep up with technological advancements (Ravenswood School 2024).

Although there is research that shows that overall class grades increase after using AI-assisted learning, there is not much research on whether using AI as a study tool can improve test scores, especially in business and STEM courses. The National Library of Medicine suggests that the information on whether using AI leads to higher exam performance is currently limited (Sakelaris et al. 2025).

3. Description of the experiment and data

3.1. Class experiment

The purpose of this study is to examine whether utilizing AI to enhance student learning could improve students' performance on exams. To test this, all students in three sections of a Principles of Microeconomics class were given a handout with five questions that resembled the ones the instructor was planning to give on the third exam for the semester. Then, a sample of these students was asked to complete an assignment that required the use of ChatGPT to enhance their preparation for the test. They were given the following instructions.

For *each of the five questions* in the handout, follow Steps 1 to 4:

Step 1: Ask Chat GPT to create two more questions similar to the ones from the handout. Use the following ChatGPT prompt: "*Suppose that you are a teacher in a Microeconomics class. Could you create two more questions similar to the one that follows and provide their detailed solutions and answers following the steps from the solution of the sample question below, please? [Copy-Paste the sample question here.]*"

- **For the questions in which there is a graph (Question 3 in the handout)**: Please add a sentence asking ChatGPT to also create graphs relevant to the two exam questions it generates. You can use the following ChatGPT prompt: *"Suppose that you are a teacher in a Microeconomics class. Could you create two more questions similar to the one that follows and provide their detailed solutions and answers following the steps from the solution of the sample question below, please? Please provide the accompanying graphs too. [Copy-Paste sample question 3 here.]"*.

Step 2: Save the entire interaction with ChatGPT. You need to submit this as part of your assignment.

Step 3: Solve one of the two questions that ChatGPT has created and compare your solutions with the ones that the chatbot has provided. This also has to be submitted as part of your assignment.

Step 4 (reflection): Write 2-3 sentences to reflect on the usefulness of AI in your test preparation. Was your solution the same as the one ChatGPT had provided? If not, feel free to interact further with the chatbot to help you reach the same final answer to the practice question. Note that this additional interaction with ChatGPT is optional. The abovementioned reflection, however, is mandatory. In addition, the reflection (Step 4) could be completed only once, rather than separately for each of the five questions in the handout.

Data were collected from three sections of a Principles of Microeconomics class in the Spring of 2025. The instructor, the material taught, the materials students were given for preparation, and the three exams given to the students were the same to make sure that the only difference between the students in the various sections of the class was the completion of the AI assignment. In addition, we provided students with the exact ChatGPT prompt because students might have had different prior experiences in interacting with ChatGPT, which could potentially affect the results of our study.

There was no AI assignment before the first and the second midterm exams. Students' test results on these two exams show their pre-treatment performance (performance in the absence of the AI intervention).

Before the third exam, the three sections of the class were treated differently as follows:

- Section 1 of the class (Control group): No treatment, i.e., no AI-enhanced learning assignment. This is the control group in our empirical analysis. We compare this group separately with the second and the third sections of the class.

- Section 2 of the class (Treatment group 1): All students were *required* to complete the AI assignment as described above before the third exam.
- Section 3 of the class (Treatment group 2): Students were given the aforementioned instructions and the option to complete the AI assignment. However, it was *optional* for this section of the class. The instructor explained the potential benefits of AI-enhanced learning and offered 3 points (7.5%) extra on the third exam (with a maximum of 40 points) to the students who chose to complete the assignment. Note that in the empirical analysis for this research, we utilize the exact score students received without the extra credit added so that the scores of the students who did and did not earn extra credit were comparable.

The intervention performed in the third section of the class allows us to examine whether there is any pattern that students voluntarily choose to do the extra work. For instance, it is possible that students who generally perform well on exams self-select into doing all mandatory and recommended work (which could be contributing to higher exam scores), or students whose performance was not excellent on earlier exams could be trying to improve their score in the class.

Test results on the third exam indicate students' post-intervention performance. Because not all students received the treatment, we are able to estimate the impact of AI-enhanced learning on students' test scores.

3.2. Data

Once we collected the data, we dropped two students who did not submit the required AI assignment. Data from all three exams are available for 95 students: 28 students in the section that was not assigned an AI assignment (29.47% of the sample); 33 students in the section that was assigned a required AI assignment (34.74% of the sample); and 34 students in the section for whom the AI assignment was optional (35.79% of the sample). This has produced a highly balanced panel of 285 observations (3 exams/time periods for each of the 95 students). Table 1 provides summary statistics of these data. About 83% of the students have declared a business major (economics, finance, marketing, accounting, management, information systems, or entrepreneurship). Freshmen, sophomores, and juniors represent 74.7%, 21.1%, and 4.2% of the sample, respectively. The mean scores on the first two exams (pre-treatment) are very similar – 78.71% (exam 1) and 78.18% (exam 2). The average score decreases to 71% on the third exam. Although there are differences between the scores of the three sections of the class, the pattern is similar.

Table 1: Summary statistics

	Mean	Std. Dev.	min	max
Panel A: All sections				
Required AI assignment	.347	.477	0	1
Optional AI assignment	.358	.48	0	1
Exam 1 score (out of 40)	31.484	4.685	22	40
Exam 2 score (out of 40)	31.274	4.979	17	40
Exam 3 score (out of 40)	28.4	5.078	19	40
COB Major	.832	.375	0	1
Class standing				
Freshman	.747			
Sophomore	.211			
Junior	.042			
Exam 1 score (%)	78.711	11.712	55	100
Exam 2 score (%)	78.184	12.448	42.5	100
Exam 3 score (%)	71	12.694	47.5	100
Panel B: By class section				
Section that was not assigned an AI assignment:				
Exam 1 score (%)	75.446	12.504	57.5	100
Exam 2 score (%)	77.411	12.92	52.5	100
Exam 3 score (%)	71.429	12.919	47.5	97.5
Section that was assigned a required AI assignment:				
Exam 1 score (%)	80.227	10.756	57.5	97.5
Exam 2 score (%)	77.576	10.829	52.5	95
Exam 3 score (%)	70.303	11.335	50	95
Section that was assigned an optional AI assignment:				
Exam 1 score (%)	79.926	11.505	55	97.5
Exam 2 score (%)	79.412	13.503	42.5	100
Exam 3 score (%)	71.324	13.804	47.5	100
Subsample of the section that was assigned an optional AI assignment & completed it:				
Exam 1 score (%)	79.107	13.204	55	97.5
Exam 2 score (%)	77.679	15.075	52.5	100
Exam 3 score (%)	70.357	16.084	47.5	100
Subsample of the section that was assigned an optional AI assignment & did not complete it:				
Exam 1 score (%)	80.5	10.228	60	97.5
Exam 2 score (%)	80.625	12.27	42.5	92.5
Exam 3 score (%)	72	12.054	55	95

Figures 1 and 2 graphically show the average scores on each exam for each of the three sections (Figure 1) and for the subsamples of students who completed and did not complete the AI assignment in the class section, who could choose whether to complete the assignment or not (Figure 2). In the latter class section, the average test scores of the students who voluntarily chose to complete the AI assignment were lower on all three tests than the mean scores of the students who chose not to complete the assignment. This could indicate that students who needed extra points more had a greater incentive to complete the extra work, but did not necessarily do it diligently

enough to affect their scores on the third exam, or students spent time interacting with AI at the expense of solving problems on their own, or the assignment simply did not help students understand the material better.

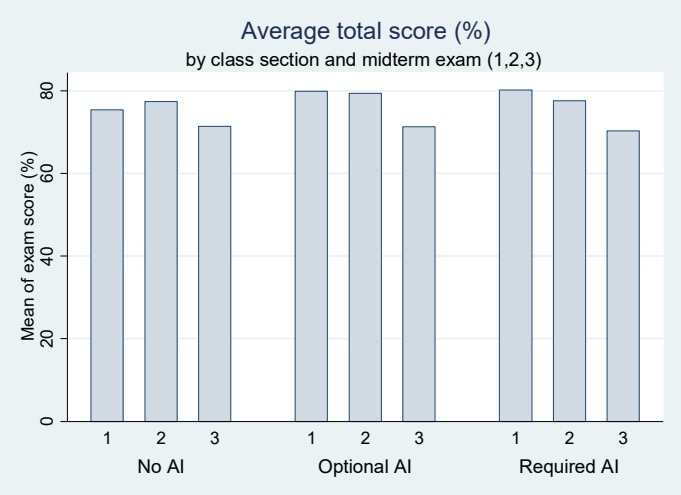

Figure 1: Mean test scores by class section and exam

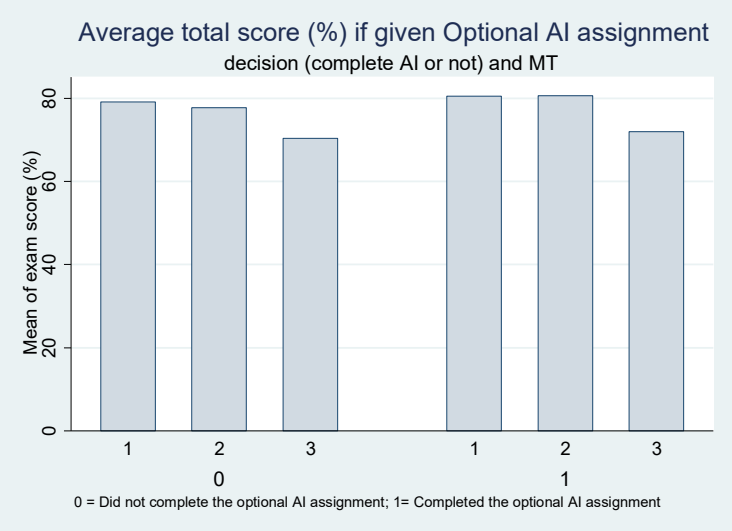

Figure 2: Mean test scores of subsamples of the class that were assigned an optional AI assignment, by decision to complete vs. not complete the optional AI assignment

4. Method

We employ a difference-in-differences (DD) method to examine whether assigning students to complete a guided AI assignment leads to better performance on the actual test. The treatment group includes students who were given the assignment (mandatory or optional, depending on the specification). Students who were not asked to do the AI assignment represent the control group.

We estimate DD regressions of the following form:

$$Score_{it} = \alpha + \beta_1 Treated_i + \beta_2 Post_t + \beta_3 Treated_i * Post_t + \boldsymbol{X_i'\gamma} + \varepsilon_{it} \quad \text{(Eq. 1)}$$

In this equation, the dependent variable $Score_{it}$ is the actual test score of student i on exam t. The variable $Treated_i$ is a binary variable indicating treatment (equal 1 if yes, and 0, otherwise). We use different specifications in which this shows whether a student has been assigned a required AI assignment, or an optional AI assignment, or has completed an AI assignment. $Post_t$ is an indicator which takes the value of 1 for the exam that provided the AI assignment opportunity (the last, i.e., the third midterm exam), and 0 for the prior exams (first and second). The DD framework also includes an interaction term between the treatment group ($Treated_i$) and the post-intervention exam ($Post_t$). Subscripts i identify students. Subscripts t indicate the exam ($t = 1, 2, 3$). X_i is a set of control variables for student i, including class standing (freshman, sophomore, or junior) and an indicator for a business major. The error term is denoted with ε_{it}.
In this specification, the DD estimate in front of the interaction term captures the causal effect of the AI assignment on students' test scores. It reflects the differential change in students' test scores after encouraging the use of AI as a learning tool compared to the same outcome (test score) prior to the AI intervention for students who were given the opportunity to do the AI assignment, relative to those who did not have the option.

The identifying assumption of this method, the parallel trends assumption, suggests that all students would have moved on the same trend in their test scores in the absence of the AI intervention, i.e., there would have been no significant difference in the change in the test scores of the treated and control groups of students in the absence of the AI intervention. If this assumption is violated, the results would be biased. To test for parallel trends, one could run a placebo test in which data from the first and the second exams are used. Neither of these exams provided an AI assignment.

5. Results

In Table 2, the treatment group consists of the students in the section that was required to submit (and submitted) the AI assignment. We compare these students to those who were not given the extra AI task (Columns 1 and 2), or those for whom it was optional (Columns 3 and 4). We also use data from MT 3 (post-intervention) and MT 2 (Columns 1 and 3) or MT 1 (Columns 2 and 4). We do not find any statistically significant difference in the changes in the exam scores after and before the intervention between the students who were and were not subject to the AI intervention. As the summary statistics suggested as well, students generally performed worse on the last exam compared to the previous ones. One possible explanation could be that students tend to get tired towards the end of the semester, which could affect their class performance.

More interestingly, the insignificant effect of the AI assignment could be explained in several ways. First, students could have completed the assignment as requested without attempting to actually understand the questions and their answers or solutions better. Having more questions available does not necessarily imply better performance if students are not practicing or engaging with the material more. In the classes data came from, students were advised to answer questions on their own before looking at the answers. Just seeing questions that look familiar may give students unwarranted confidence that they know the material better than they actually do. Second, it is

possible that the prompt we gave students to use in ChatGPT was imperfect, or AI tools are not ideal for STEM subjects. In these cases, the questions or more likely their solutions might have confused the students. We tried to alleviate this issue by giving AI tools sample steps to answer each type of questions, but future investigation could explore whether and how much using different prompts affects the quality of the AI-generated content, and how this quality impacts students' performance.

Table 2: Regression Results - Treatment: Required AI

	(1)	(2)	(3)	(4)
	No AI vs. Required	No AI vs. Required	Optional AI vs. Required	Optional AI vs. Required
VARIABLES	MT2 vs. MT3	MT1 vs. MT3	MT2 vs. MT3	MT1 vs. MT3
Required AI	5.345*	0.272	-1.784	1.908
	(2.901)	(3.220)	(3.165)	(2.685)
Post	-4.018	-5.982*	-8.088**	-8.603***
	(3.192)	(3.411)	(3.286)	(3.041)
Required AI*Post	-5.906	-1.291	0.816	-1.321
	(4.221)	(4.414)	(4.310)	(4.048)
COB Major	5.424*	6.248**	1.987	-1.515
	(3.167)	(3.087)	(3.205)	(3.100)
Freshman	6.494**	-1.679	-6.818*	5.765
	(2.596)	(3.551)	(3.844)	(3.773)
Sophomore	12.33***	0.0631	-1.993	13.32***
	(3.383)	(4.375)	(4.606)	(4.195)
Observations	122	122	134	134
R-squared	0.186	0.107	0.123	0.195

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The aforementioned findings are confirmed in the specifications in Table 3 where the treatment groups comprised of students who completed the AI assignment. Again, we did not find any statistically significant improvement or deterioration in the scores of the students who completed versus those who did not complete the AI assignment.

Table 3: Regression Results - Treatment: Completed AI

VARIABLES	(1) Optional AI vs. Required MT1 vs. MT3	(2) Optional AI vs. Required MT2 vs. MT3	(3) Completed optional AI MT1 vs. MT3	(4) Completed optional AI MT2 vs. MT3
Completed AI	1.016	0.408	-0.186	1.500
	(3.738)	(4.319)	(4.315)	(4.958)
Post	-8.750	-7.321	-8.750	-7.321
	(5.550)	(5.919)	(5.657)	(6.045)
Completed AI*Post	-0.637	-0.462	0.250	-1.304
	(5.936)	(6.331)	(6.640)	(7.123)
COB Major	-1.312	1.725	-0.760	2.219
	(3.138)	(3.253)	(4.433)	(4.646)
Freshman	5.365	-6.166*	-7.490**	-7.364**
	(3.578)	(3.646)	(3.271)	(3.181)
Sophomore	12.60***	-1.054		
	(3.898)	(4.134)		
Observations	134	134	68	68
R-squared	0.193	0.120	0.165	0.143

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In Table 4, we compare the section that was given an optional AI-driven assignment and the students who were not offered this opportunity. In Table 5, the treatment group consists of students who completed an optional AI assignment. We compare students who voluntarily chose to complete the optional AI assignment with the students who were enrolled in the same class section but chose to not submit this assignment (Columns 1 and 2 of Table 5), or the section that did not have an AI task at all (Columns 3 and 4), or all students except the ones in the section that was required to submit the AI assignment (Columns 5 and 6). As previously noted, students who had performed worse on the previous tests tended to be more willing to complete the AI assignment, perhaps due to the extra credit offered for submitting it or the willingness to improve. Yet, completing the assignment, even voluntarily, did not seem to yield any difference between those students who submitted the assignment relative to their classmates who did not, or did not have an AI-enhanced preparation assigned, or both.

Table 4: Regression Results - Given optional AI vs. not

VARIABLES	(1) No AI vs. Optional MT1 vs. MT3	(2) No AI vs. Optional MT2 vs. MT3
Optional AI	4.343	2.217
	(3.076)	(3.471)
Post	-4.018	-5.982*
	(3.218)	(3.416)
Optional AI*Post	-4.585	-2.106
	(4.467)	(4.748)
COB Major	4.423	4.818
	(3.354)	(3.618)
Freshman	6.320*	2.039
	(3.528)	(5.285)
Sophomore	12.78***	7.762
	(4.007)	(5.537)
Observations	124	124
R-squared	0.155	0.126

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Regression Results - Treatment: Completed optional AI

	(1) Completed optional AI vs. Not complete optional AI	(2) Completed optional AI vs. Not complete optional AI	(3) Completed optional AI vs. No AI or not completed option	(4) Completed optional AI vs. No AI or not completed option	(5) Completed optional AI vs. No AI or not completed option or required AI	(6) Completed optional AI vs. No AI or not completed option or required AI
VARIABLES	MT1 vs. MT3	MT2 vs. MT3	MT1 vs. MT3	MT2 vs. MT3	MT1 vs. MT3	MT2 vs. MT3
Completed Optional AI	-0.186	1.500	3.006	2.675	1.061	2.903
	(4.315)	(4.958)	(3.135)	(3.443)	(2.755)	(3.070)
Post	-8.750	-7.321	-5.595*	-6.429**	-7.500***	-6.800***
	(5.657)	(6.045)	(2.866)	(3.014)	(2.002)	(2.074)
Completed Optional AI*Post	0.250	-1.304	-2.905	-2.196	-1	-1.825
	(6.640)	(7.123)	(4.523)	(4.821)	(3.965)	(4.284)
COB Major	-0.760	2.219	4.258	4.821	2.558	4.287
	(4.433)	(4.646)	(3.433)	(3.565)	(2.759)	(2.707)
Freshman	-7.490**	-7.364**	7.078**	2.233	6.948***	-1.808
	(3.271)	(3.181)	(3.205)	(5.087)	(2.159)	(3.458)
Sophomore			13.30***	7.725	13.17***	2.312
			(3.744)	(5.334)	(2.776)	(3.906)
Observations	68	68	124	124	190	190
R-squared	0.165	0.143	0.147	0.127	0.156	0.113

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In most of the specifications where class standing is statistically significant, sophomores tend to perform better and freshmen worse than juniors. Whether a student majors in a business (COB) discipline or not does not have a statistically significant effect on the performance on the third exam in all specifications except the two where we compare the class section that had no AI assignment versus the one that was required to submit this assignment (Columns 1 and 2 of Table 2). In the latter case, COB majors performed better than non-COB majors.

It is worth discussing what the statistically insignificant effect of preparing for economics exams through AI could imply about students and professors teaching these classes. First, students should distinguish between utilizing AI as a sole learning tool versus augmenting their learning with the new technology. Specifically, our study does not account for the amount of time students spent preparing other than completing the AI assignment, nor do we control for how diligently they prepared with and without AI tools. It could be the case that students spent time interacting with ChatGPT instead of (rather than in addition to) answering questions and practicing the material on their own. It is important that instructors make it clear to their students that AI should complement rather than

substitute traditional learning methods, at least until both students and educators master working with AI and/or AI tools improve performance for business and STEM disciplines. It is also important to mention that even if AI may not be optimal for preparing for business and STEM courses at this time, this does not mean that students should not be learning how to use the tools for career-related reasons after graduation.

6. Conclusion

AI has become a widespread tool that is transforming many industries. Businesses have been changing the way they function, from operations to marketing, through the power of Generative AI. Because of the increasing use of AI in business settings, many employers look to hire applicants with AI expertise. Therefore, the ability to interact with AI successfully could give an advantage to job applicants with the relevant skills and experience, making it necessary for college students to learn how to use AI to remain competitive in the job market.

With this in mind, it is a challenge for college professors to make sure that students have a sufficient understanding of the class material and at the same time, equip the students with the contemporary skills they will need when they start their careers, including AI literacy.

In this research, we explore whether AI-assisted test preparation could serve a two-fold purpose: help students prepare for an exam and improve their performance, and practice working with AI. We find evidence that the skepticism of some educators that AI-supported learning could improve students' class performance could be legitimate in economics classes. Specifically, we do not find evidence of a statistically significant difference in the change in exam scores in an introductory economics class between students who used AI tools to prepare for an exam and students who did not enhance learning with AI. Yet, we do believe that students should take advantage of the opportunities that AI has to offer, just not for exam preparation for business and STEM courses. Teachers should warn their students of the threats of using AI, educate the learners how to use the new tools responsibly, and incorporate AI into their curriculum in ways yet to be explored to prepare their students for their future careers.

References

Alamri, A. (2024). The impact of artificial intelligence on student learning and engagement: A systematic review. *Education Sciences, 15(3),* 343. https://doi.org/10.3390/educsci15030343

Baillifard, A., Gabella, M., Banta Lavenex, P., & Martarelli, C. S. (2024). Effective learning with a personal AI tutor: A case study. *Education and Information Technologies, 30*(1), 297–312. https://dl.acm.org/doi/10.1007/s10639-024-12888-5

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education. [Google Scholar]

Ravenswood School. (2024). Using AI to enhance Socratic learning: A case study. *AI in Education*. Retrieved from https://www.ai-in-education.co.uk/resources/ravenswood-case-study-using-ai-to-enhance-socratic-learning

Sakelaris, P. G., Novotny, K. V., Borvick, M. S., Lagasca, G. G., & Simanton, E. G. (2025). Evaluating the use of artificial intelligence as a study tool for preclinical medical school exams. *Journal of Medical Education and Curricular Development, 12*, 23821205251320150. https://doi.org/10.1177/23821205251320150

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching, 6*(1), 1-10. https://doi.org/10.37074/jalt.2023.6.1.17

Wu, Y. (2023). Integrating generative AI in education: How ChatGPT brings challenges for future learning and teaching. *Journal of Advanced Research in Education, 2*(4), 6–10. Available online: https://www.pioneerpublisher.com/jare (accessed on 29 January 2025). [CrossRef]

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0176-8