




# **From Theory to Practice: AI's Influence on Teaching Economics**



Justin Wolfers and Joshua Hill  
Online webinar | October 3, 2024

# Economic consequences of new technologies

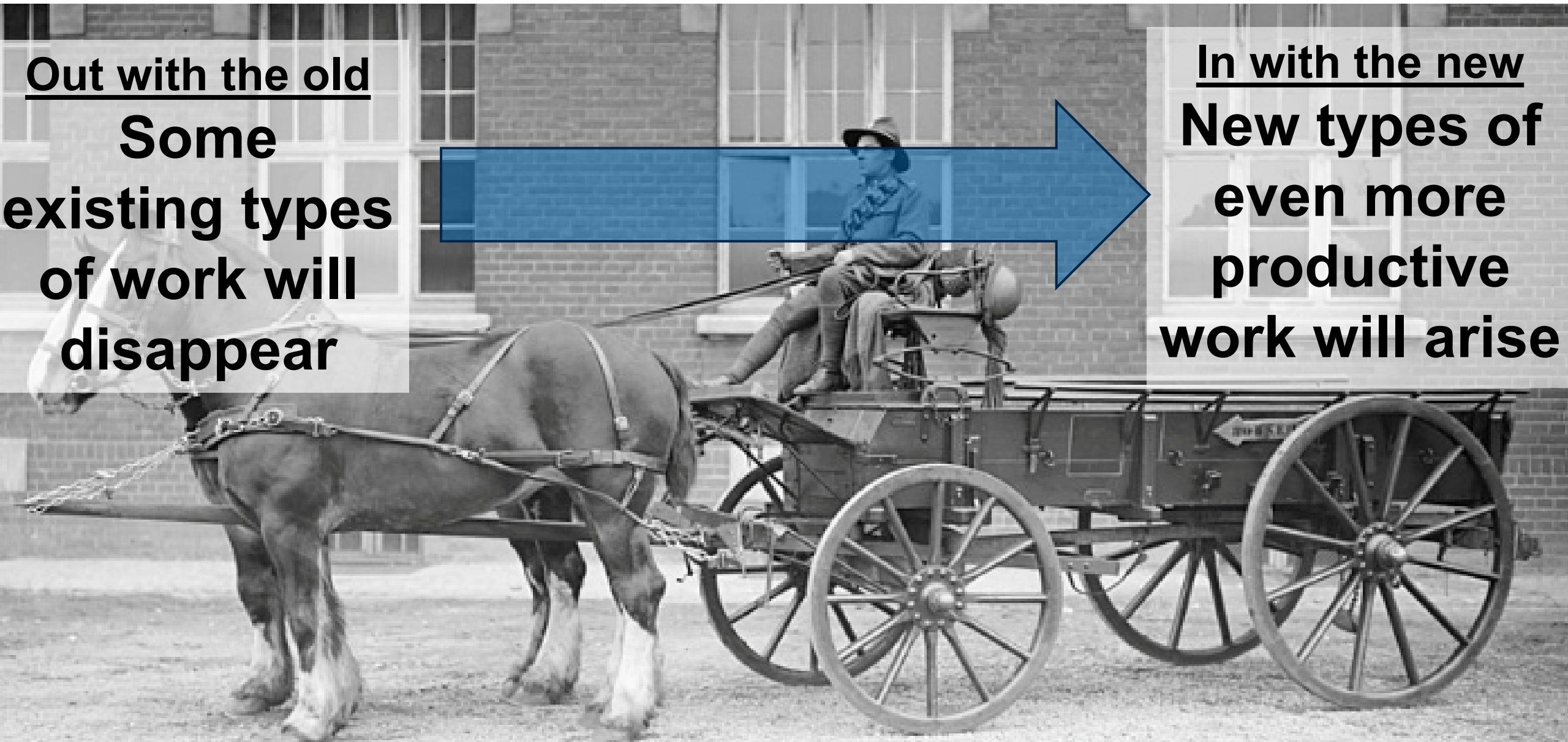
Out with the old

Some  
existing types  
of work will  
disappear



In with the new

New types of  
even more  
productive  
work will arise



# Economic consequences of new technologies

Out with the old

**Some existing types of work will disappear**



High-stakes at-home assessment is dead  
(AI-fueled plagiarism killed it)

The same rules apply to Gen AI

In with the new

**New types of even more productive work will arise**



With new tools we can do new things  
(Let's figure out what)

# The challenge posed by AI

1. Your students are using Chat GPT

## Young people are using AI tools frequently

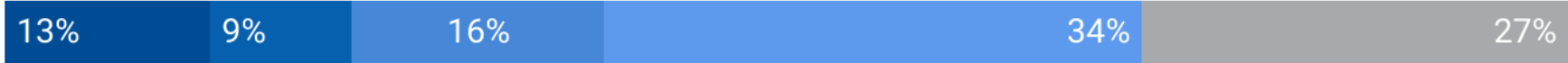
How often do you use or interact with artificial intelligence, or AI?

■ Several times a day ■ About once a day ■ Several times a week ■ Less often ■ Never

Aged 18-29 years



30-49 years



50-64 years



65 years or older



2. Chat GPT is better than you think

3. Your university is likely giving you terrible advice

Chart: @JustinWolfers • Source: KFF Health Misinformation Tracking Poll (June 3-24, 2024) of U.S. adults (n=2,428).

# The challenge posed by AI

---

1. Your students are using Chat GPT

## Share of students who said that they...

Are aware of ChatGPT



2. Chat GPT is better than you think

Have used ChatGPT to help with homework assignments



Used ChatGPT to write an essay



3. Your university is likely giving you terrible advice

Used ChatGPT for an at-home test or quiz



Source: January 2023 Study.com survey of 1,000 students aged 18 or over  
<https://study.com/resources/perceptions-of-chatgpt-in-schools>

# The challenge posed by AI

1. Your students are using Chat GPT

2. Chat GPT is better than you think

3. Your university is likely giving you terrible advice

## ChatGPT scores at the top of the class in standardized tests

Percentile rank of ChatGPT

■ GPT 3.5 ■ Chat GPT 4.0

*And more recent models are even better*

### After grad school

Bar exam (after Law School) 90%

### Graduate admissions

LSAT 88%

GRE Quant 80%

GRE Verbal 99%

GRE Writing 54%

### College level(ish) courses

TUCE Micro (Micro 101) 91%

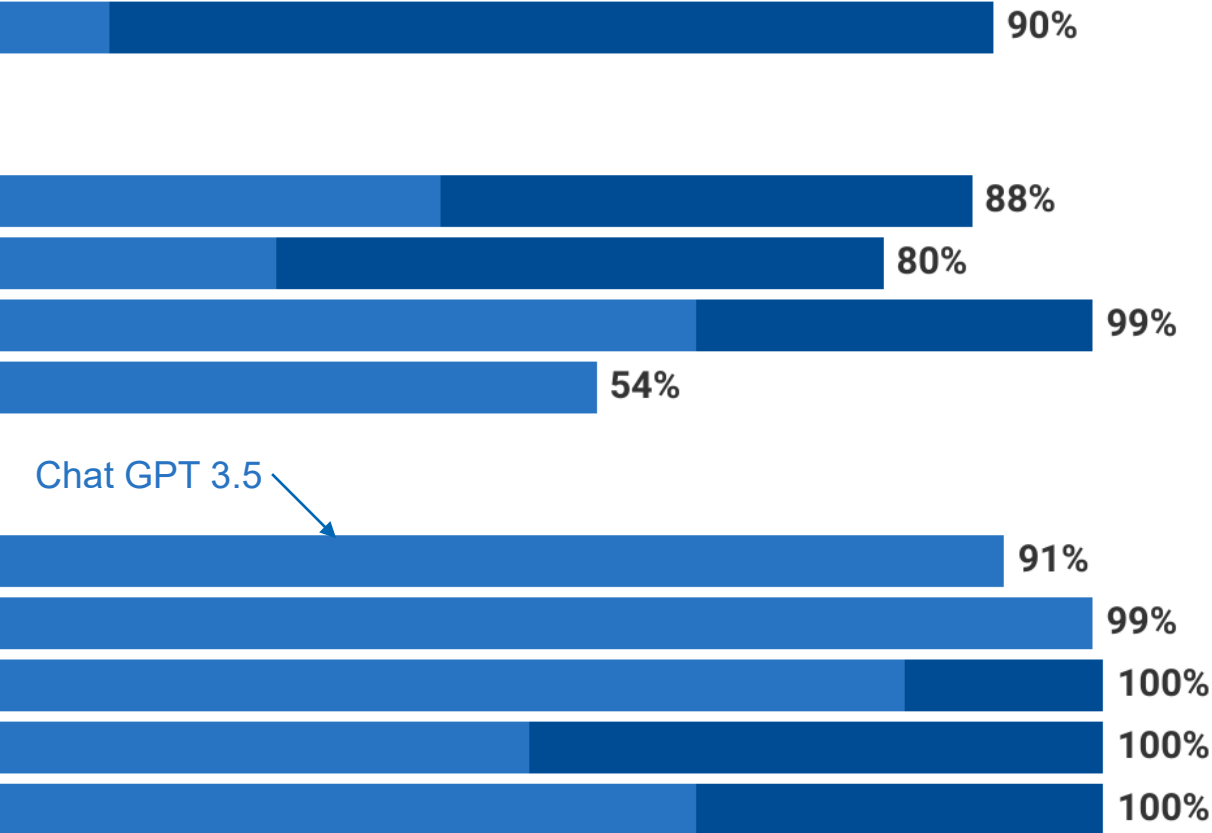
TUCE Macro (Macro 101) 99%

AP Micro 100%

AP Macro 100%

AP Statistics 100%

Chat GPT 3.5



# The challenge posed by AI

1. Your students are using Chat GPT
2. Chat GPT is better than you think
3. Your university is likely giving you terrible advice

**A Harvard student asked her professors to grade ChatGPT's essays. It got mostly A's and B's**



“It is beautifully written!”

“Well written and well articulated paper.”

“Clear and vividly written.”

“The writer's voice comes through very clearly.”

“impressive... attention to detail”

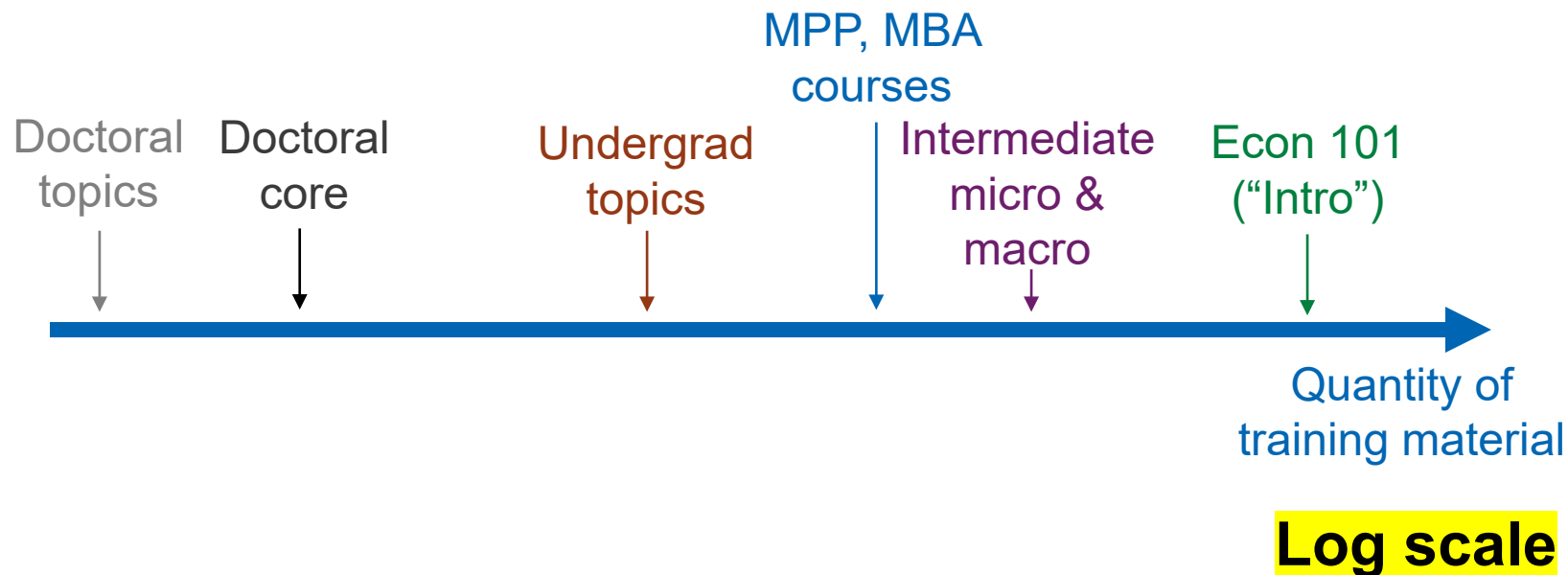
# The challenge posed by AI

---

1. Your students are using Chat GPT

2. Chat GPT is better than you think at intro econ **but not all econ classes**

3. Your university is likely giving you terrible advice



This is why ChatGPT is *nearly perfect* at Econ 101 ...but not other subjects or levels



# The challenge posed by AI

---

1. Your students are using Chat GPT
2. Chat GPT is better than you think
3. Your university is likely giving you terrible advice

## A New Digital Divide: Student AI Use Surges, Leaving Faculty Behind

While both students and faculty have concerns with generative artificial intelligence, two new reports show a divergence in AI adoption.

# The challenge posed by AI

1. Your students are using Chat GPT

2. Chat GPT is better than you think

3. Your university is likely giving you terrible advice

## Universities see shortcomings

- ❑ Responses are **formulaic**
- ❑ **Hallucinates** facts (& citations)
- ❑ Makes (math) **errors**
- ❑ Not up to date
- ❑ Computers lack **creativity**
- ❑ Trained on canon of dead white men

## Students find solutions

- ❑ “Sharpen that essay”
- ❑ Set **temperature** = 0
- ❑ Prompt “use **Python**”
- ❑ Use **Bing AI** not GPT
- ❑ Not true
- ❑ So are our intro econ students

## Prompt engineering

# The challenge posed by AI

---

1. Your students are using Chat GPT

2. Chat GPT is better than you think

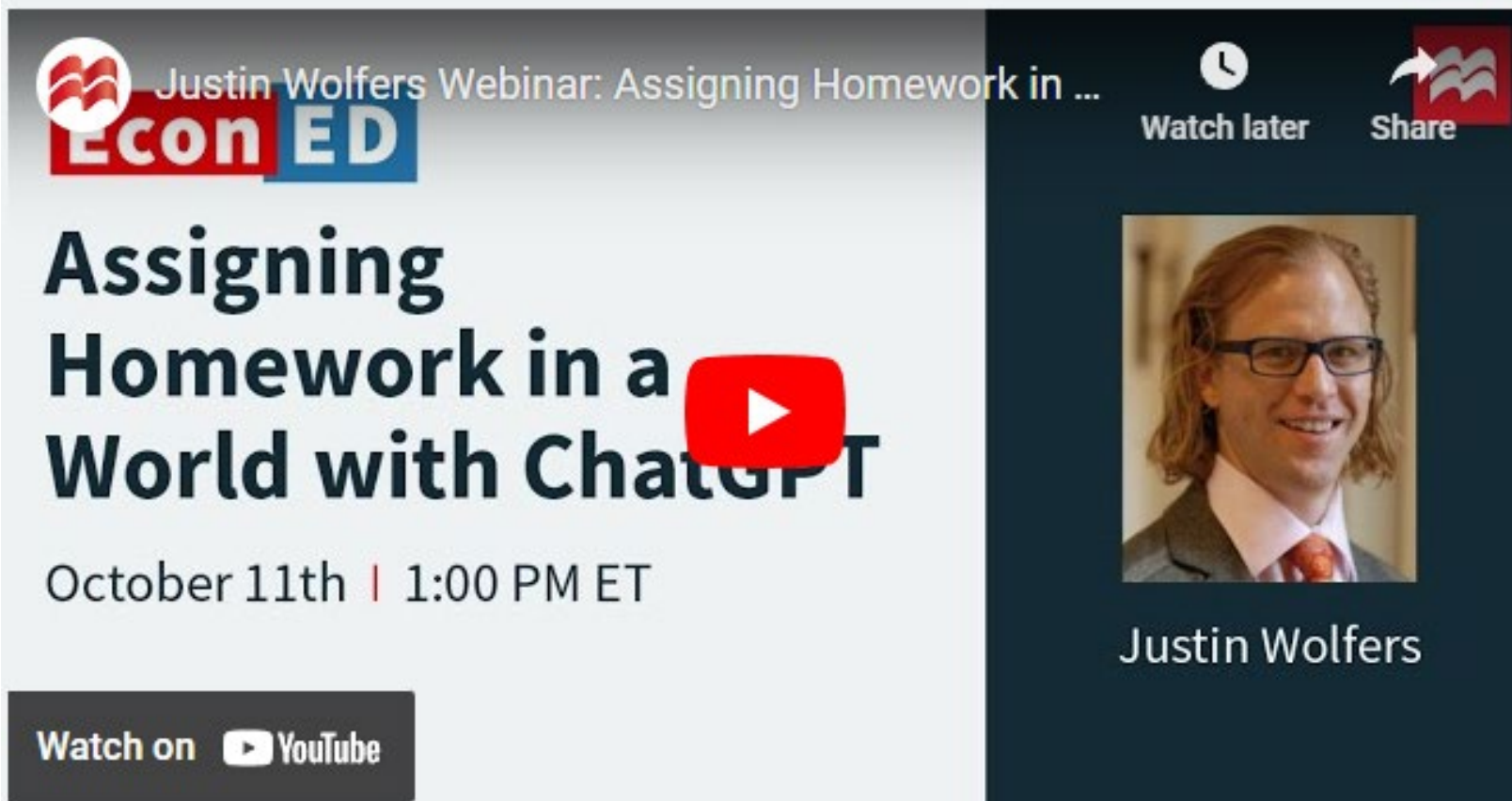
3. Your university is likely giving you terrible advice

## A meta theory

Universities	Students
<ul style="list-style-type: none"><li>❑ Looking for what GPT <b>can not</b> do</li><li>❑ <b>Confirmation bias:</b> Search for ways in which ChatGPT <i>is</i> ineffective</li></ul>	<ul style="list-style-type: none"><li>❑ Looking for what GPT <b>can</b> do</li><li>❑ <b>Problem-solving:</b> Search for ways to <i>make</i> ChatGPT effective</li></ul>

# More on the challenges posed by Generative AI

---



The image shows a YouTube video player interface. On the left, the video title is "Assigning Homework in a World with ChatGPT" in large, bold, black text. Above the title is the "EconED" logo, which consists of a red book icon and the text "EconED" in red and blue. Below the title, the date and time are listed as "October 11th | 1:00 PM ET". At the bottom left, there is a "Watch on YouTube" button. On the right side of the player, there is a "Watch later" button with a clock icon and a "Share" button with a red book icon. Below these buttons is a portrait of Justin Wolfers, a man with glasses and a suit, smiling. The name "Justin Wolfers" is written below the portrait.



An earlier webinar: <https://tinyurl.com/54peshr3>

# Economic consequences of new technologies



Out with the old

**Some  
existing types  
of work will  
disappear**



High-stakes at-home  
assessment is dead  
(AI-fueled plagiarism killed it)

In with the new

**New types of  
even more  
productive  
work will arise**



With new tools we can  
do new things  
(Let's figure out what)



[Opinion](#) | Adrian Wooldridge, Columnist

## **Can Oxford and Cambridge Save Harvard From ChatGPT?**

Their time-tested tutorial system offers top US universities a way to blunt AI cheating and revive real learning.

August 23, 2023 at 12:00 AM EDT

## One uneconomic idea

---



“A tutorial is the weekly meeting all students have with a tutor and usually one or two other students, at which you are expected to talk in depth about your ideas and opinions relating to that week's reading or problems.”

-- The New College [website](#)

An Oxford tutorial

One big idea:

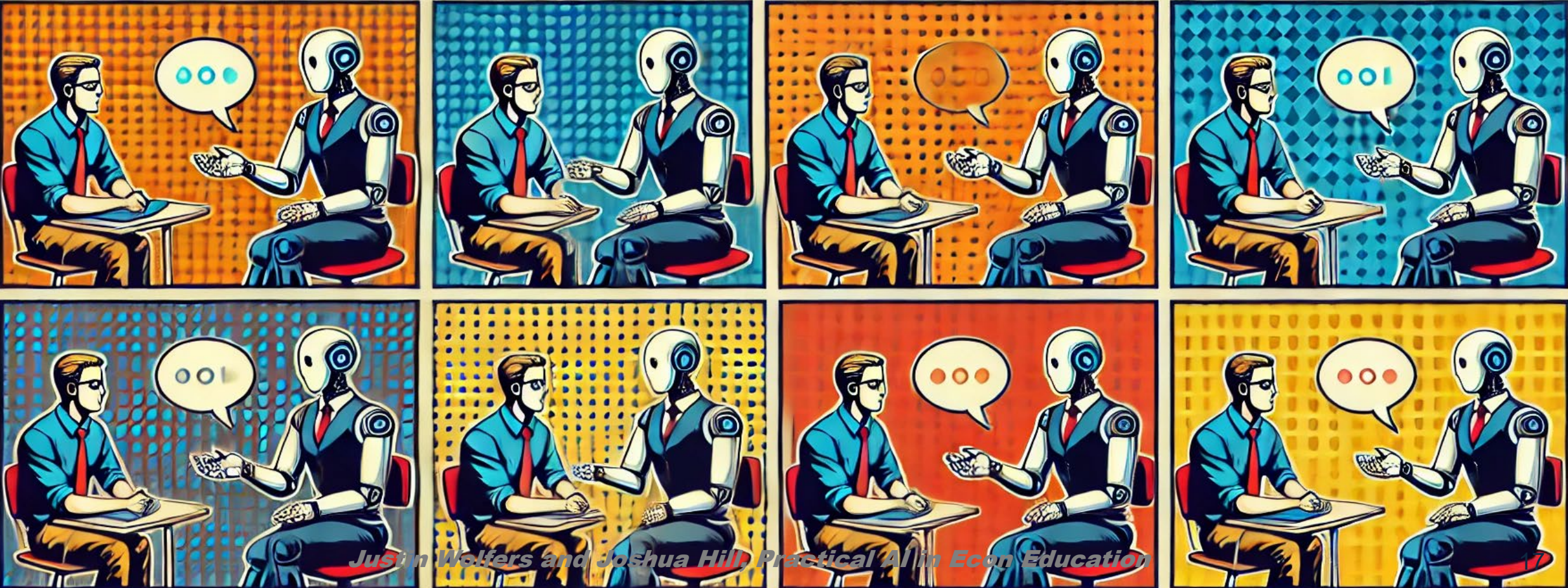
# Personalization at scale





One (feasible) big idea:

# Personalization at scale



## Some simple economics: Why scale matters

### Enhance student learning

University graduates outearn non-grads  
by \$1 million over the course of their lives  
= \$250,000 per year  
= \$125,000 per semester  
≈ \$30,000 per class (eg Econ 101)

If AI makes me just 5% more effective  
→ Student is \$1500 better off

### Save instructor time

I teach 500 students each week  
and AI saves me just 1 minute per  
student each week  
= 500 minutes per week  
= 8.3 hours per week  
x \$60 per hour cost of instructor  
= \$500 per week  
x 16 weeks per semester  
→ Department saves \$8,000



One big idea:

# Personalization at scale

Three (tentative) applications:

## A Socratic tutor

- **Personalized** homework help for each student
- Providing help, but not for the answer

## A practice exam coach

- Provide practice questions for **each** student before the exam
- On the topics **they** want to be drilled on

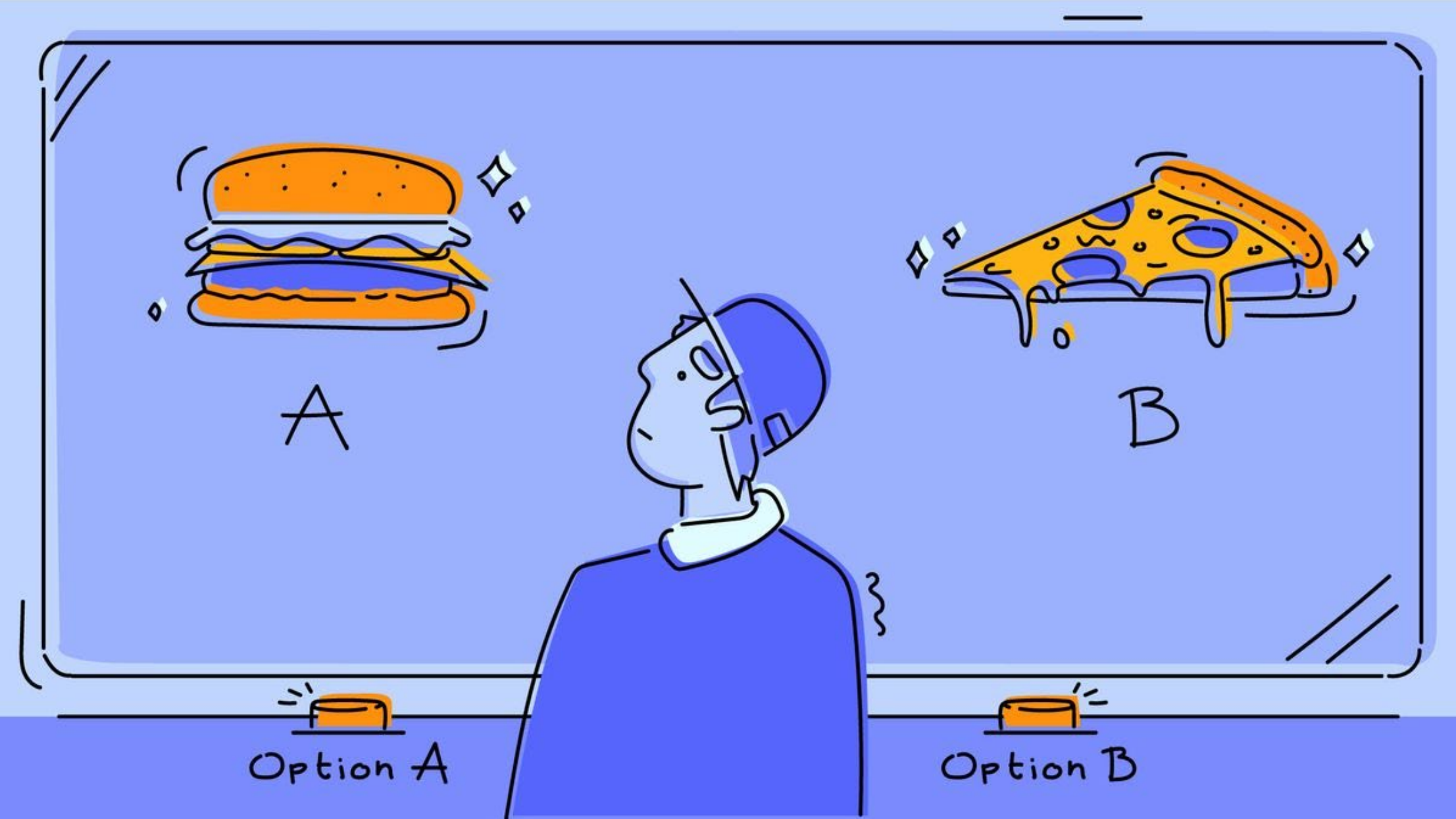
## Polling question generator

- Stimulate class discussions
- On the topics **you** want to teach about

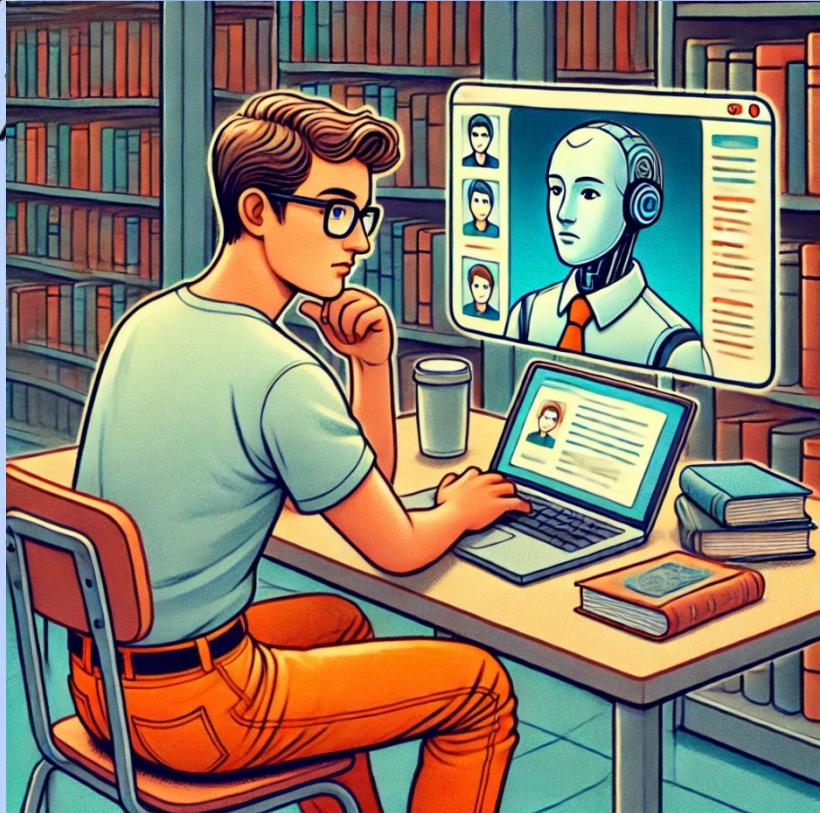
---

...and many more to come

# When evaluating AI tools, don't forget the opportunity cost principle



# AI is not perfect (but neither is your alternative)



Option A

Option B

AI can be at least as good as a pretty good



teaching assistant



Option A



Option B

# AI can be at least as good as the alternative (which may be nothing)



Option A



Option B

# Personalized Tutoring at Scale



# Making a Tutoring Bot is EASY (a 2 minute job!)

-----  
**Test it out**

chatbot tutor for  
courses. Students will  
for help on homework  
stuff they didn't  
from class. Include some  
starters.

**the parameters]**

atural rate?

e of **unemployment...**

# Making a Tutoring Bot is EASY



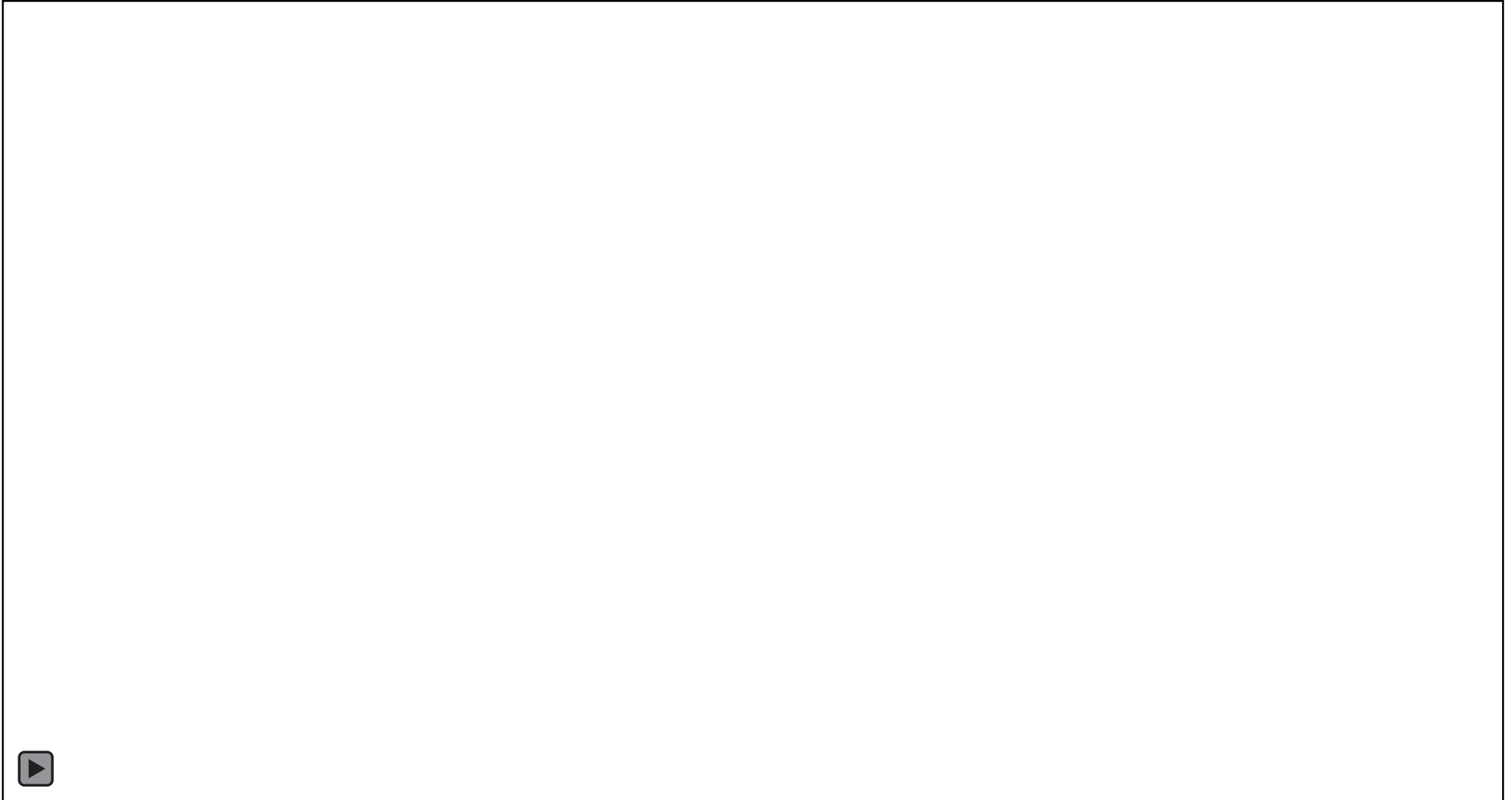
Make me a chatbot tutor for economics courses. Students will come to you for help on homework or to explain misunderstandings from class including some explanation of the...

**problem**

**solve?**

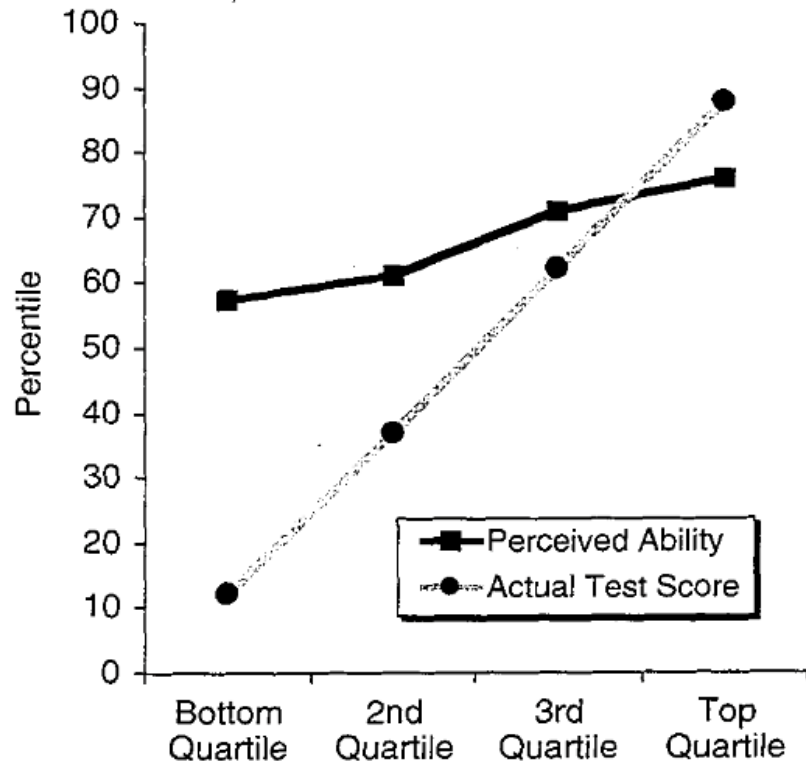
# This is a problem we already have a solution for

---



# Why haven't chatbots taken over the world?

## The Dunning-Krueger effect



“You don’t know what you don’t know”

→ Students don’t know what to do with **undirected** chatbots

# What do people do with undirect chatbots? (Cheat)

---

An open-ended chatbot  
has one ideal use  
case:

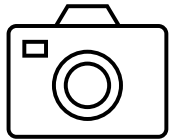
Ctrl-C + Ctrl-V = **Cheating**  
(without thinking)

# What do people do with undirect chatbots? (Cheat)

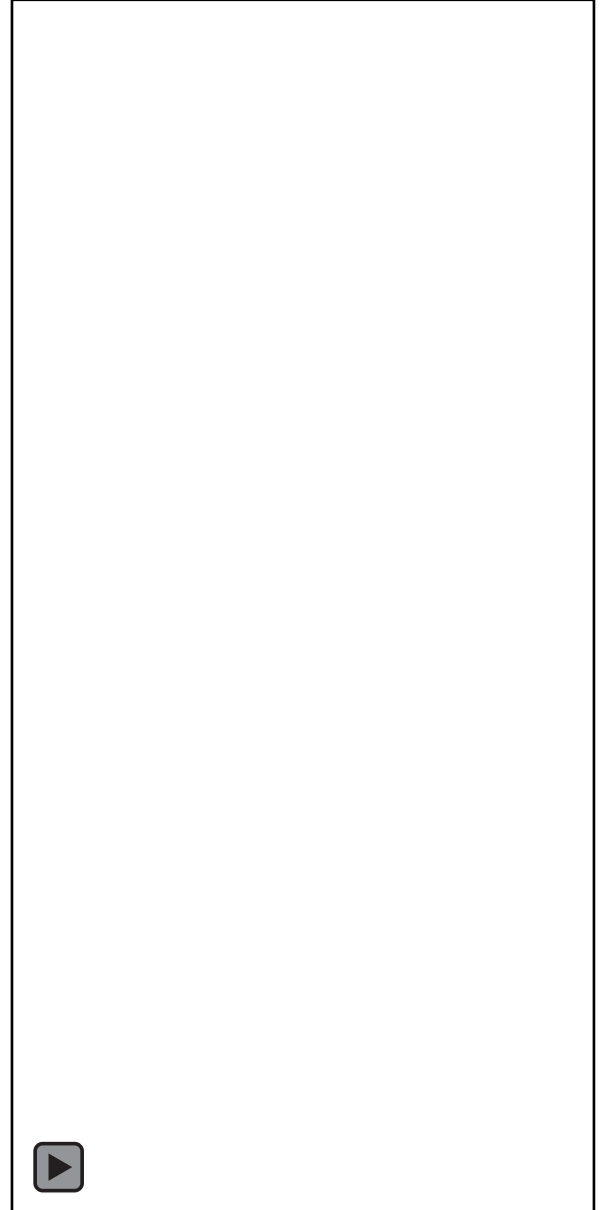
---

An open-ended chatbot  
has one ideal use  
case:

Ctrl-C + Ctrl-V = **Cheating**  
(without thinking)



→ Answer



# Few users truly **engage** with chatbots, preferring to just copy-paste

## □ Optimal use of AI:

Engage with AI as a co-creator

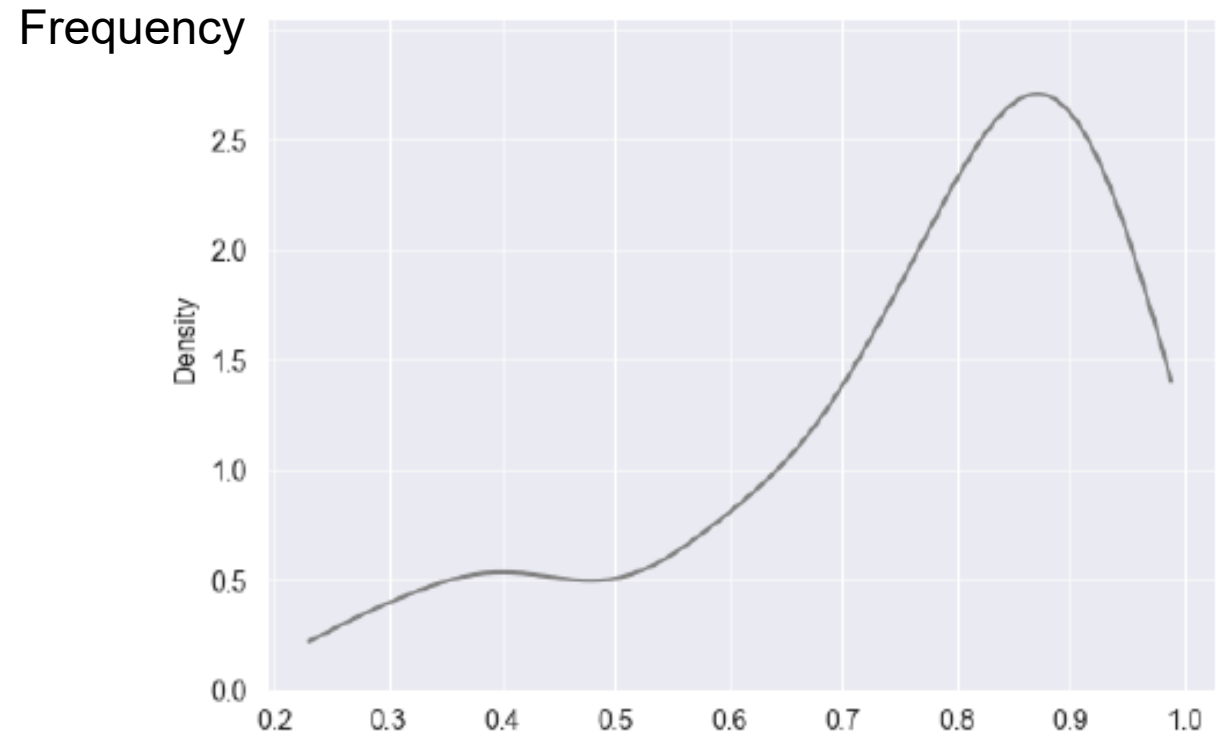
- ▶ A **complement** to thinking

## □ Actual use:

Ctrl-C + Ctrl-V

A **substitute** for thinking

**When BCG consultants were given access to ChatGPT, their responses to consulting problems closely mirrored ChatGPT's output**



**Similarity** of consultant's answer to GPT's response  
(0 = every character is different; 1 = every character is the same)

# Solving the Learning Problem is HARD

- ❑ If students **knew how to ask the questions they needed** to fill their knowledge gaps, they would just use Google.

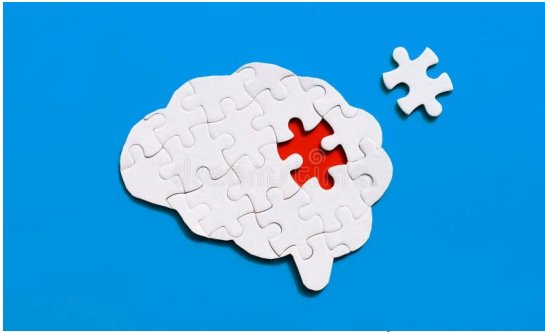
## Knowledge



Abundant  
(low value)



## Understanding knowledge gaps



Scarce  
(high value)

Where are these assets located?

Personalization is essential →

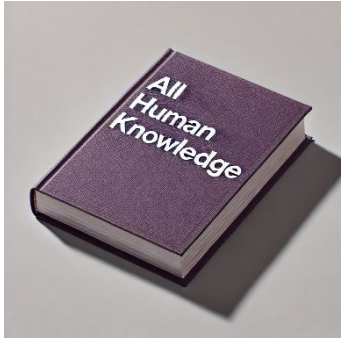




# Solving the **Learning Problem** is HARD

- If students **knew how to ask the questions they needed** to fill their knowledge gaps, they would just use Google.

Knowledge



Understanding knowledge gaps



Abundant  
(low value)

Scarce  
(high value)

- We have to create solutions to the real problem!
  - ▶ Each student comes with **unique** set knowledge and skills.
  - ▶ They need tutoring that can **adapt** to their personal situation.

# Consider your ideal teaching assistant



Chat GPT's image of the ideal teaching assistant

- As painted by Van Gogh
- As painted by Dali
- As painted by Klimt
- As painted by Rembrandt

# Consider your ideal teaching assistant

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Works long hours

*Available when (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*



# Your ideal teaching assistant **knows the answers**

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

*When (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

## ChatGPT scores at the top of the class in standardized tests

Percentile rank of ChatGPT

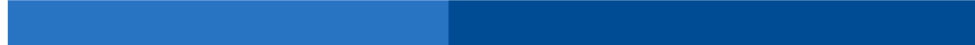
■ GPT 3.5 ■ Chat GPT 4.0

*And more recent models are even better*

### After grad school

Bar exam (after Law School)  90%

### Graduate admissions

LSAT  88%

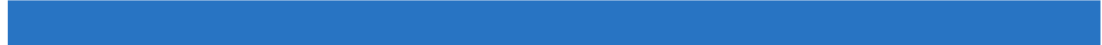
GRE Quant  80%

GRE Verbal  99%

GRE Writing  54%

### College level(ish) courses

TUCE Micro (Micro 101)  91%

TUCE Macro (Macro 101)  99%

AP Micro  100%

AP Macro  100%

AP Statistics  100%

Justin

# Your tutor must be customized to your course

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

*When (and where) each student most need help*

## Multilingual

*Provide extra support as needed*

### Business Cycles — End of Chapter Problem

A politician makes the following comment: “The fundamentals of our economy are very strong. According to market economists, we are producing more than anyone expected and even beyond what they call our natural rate. My goal is to guarantee that we continue to produce more than our potential output throughout the next few decades.”

Hi! 🙌 I'm your AI tutor!

Let's start with a hint:

Potential output is the level of output that occurs when all resources are fully employed.

Before you attempt the question, here are a few ways that I can help:

what is the natural rate in the quote?

What do you think the term "natural rate" might refer to in the context of an economy's output?



idk

The "natural rate" often refers to the level of output an economy can sustain over the long term without increasing inflation. How might this relate to the concept of potential output?



# Empathy helps meet your student where they are

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

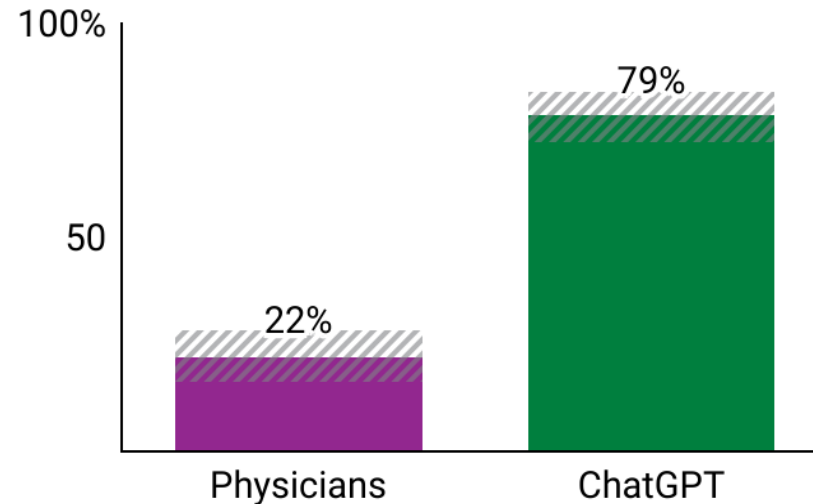
*When (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

## ChatGPT offers **higher quality** medical advice than physicians

*Share of responses evaluated by healthcare professionals as "good", or "very good"*

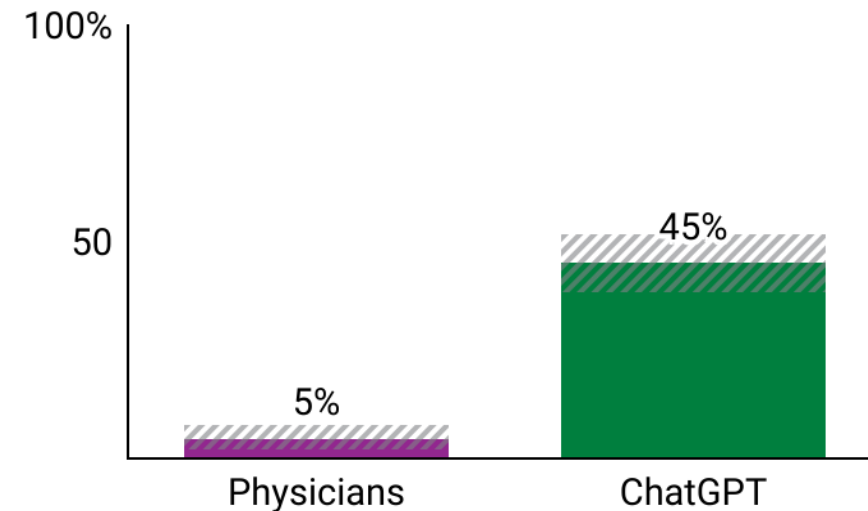


Shaded area shows 95% confidence interval. n=195 randomly drawn patient questions

Chart: @JustinWolfers • Source: Ayers, Poliak, Dredze, et al (2023), "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum", *JAMA Internal Medicine* 183(6).

## ChatGPT offers **more empathetic** medical advice than physicians

*Share of responses evaluated by healthcare professionals as "empathetic", or "very empathetic"*



Shaded area shows 95% confidence interval. n=195 randomly drawn patient questions

Chart: @JustinWolfers • Source: Ayers, Poliak, Dredze, et al (2023), "Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum", *JAMA Internal Medicine* 183(6).

# Empathy helps meet your student where they are

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Works long hours

*Available when (and where) each student most need help*

## Multilingual

*Provide extra support as needed*

**student:** its late at night. i need to go to sleep but i should do this now

**AI Tutor:** I understand that you're feeling tired, but it's important to get a good night's sleep for your overall well-being and learning efficiency. 😴 How about we continue this tomorrow when you're feeling refreshed? You'll be able to grasp these concepts much better after a good rest. What do you think?

**student:** okay

**AI Tutor:** Great! Remember, it's important to balance study with rest. Sleep well and we'll tackle these chemistry problems when you're ready. Good night! 🌙

**student:** thanks

**AI Tutor:** You're welcome! Don't hesitate to reach out when you're ready to continue. Have a good rest! 😊

**student (about 12 hours later):** im back

**AI Tutor:** Welcome back! I hope you're feeling refreshed. 😊 Ready to continue with the next reaction?

# Available **when** each student needs help

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

*When (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

### Students' Peak Demand Falls Outside of Working Hours

Shaded regions indicate 5PM - 9AM.

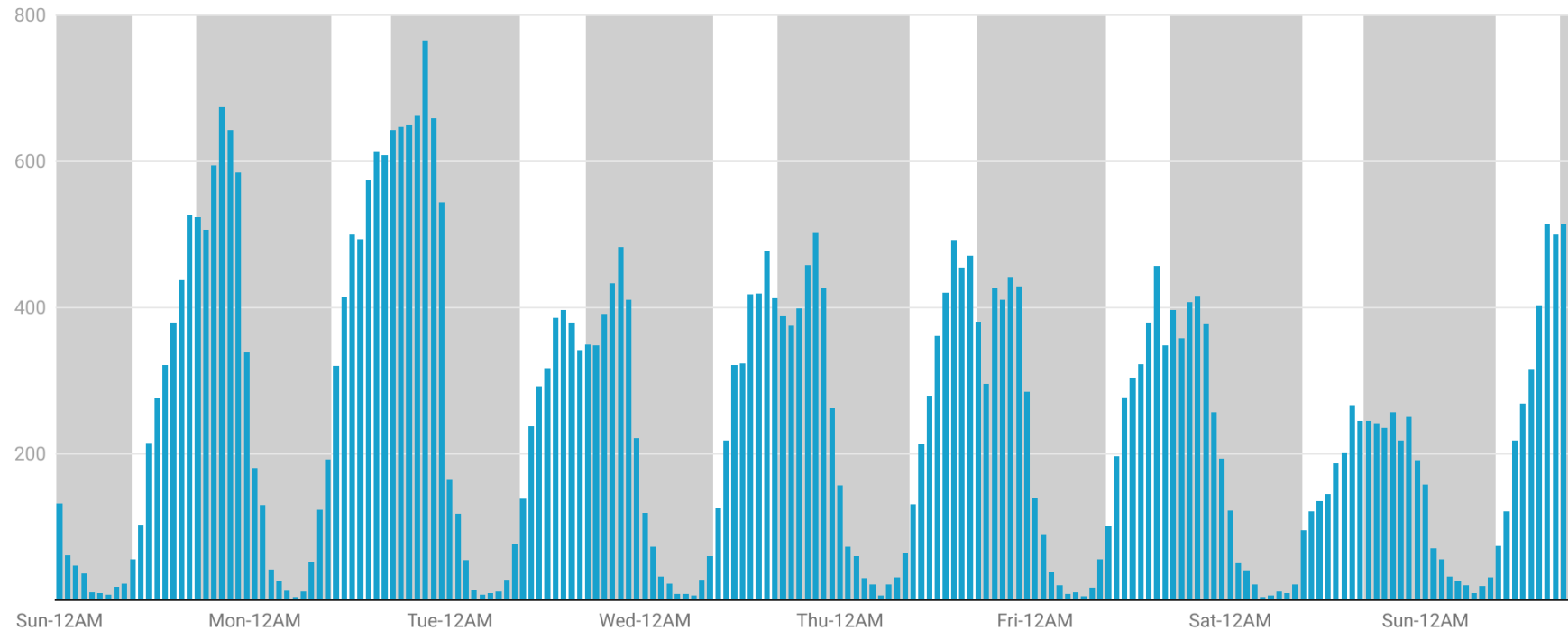


Chart: Joshua Hill • Source: Macmillan Learning



# Available **where** each student needs help

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

*When (and where) each student most needs help*

## Multilingual

*Provide personalized support as needed*



It's hard to get many students to office hours



But it's easier to find them at home

*“first-generation, underrepresented minority students and low-income students are less likely to take advantage of office hours than their classmates.”*

Source: Supiano (2023), “The missed opportunity of Office Hours”, *Chronicle of Higher Education*, 8/21/2023

# Meet each student in the **language** they find most natural

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

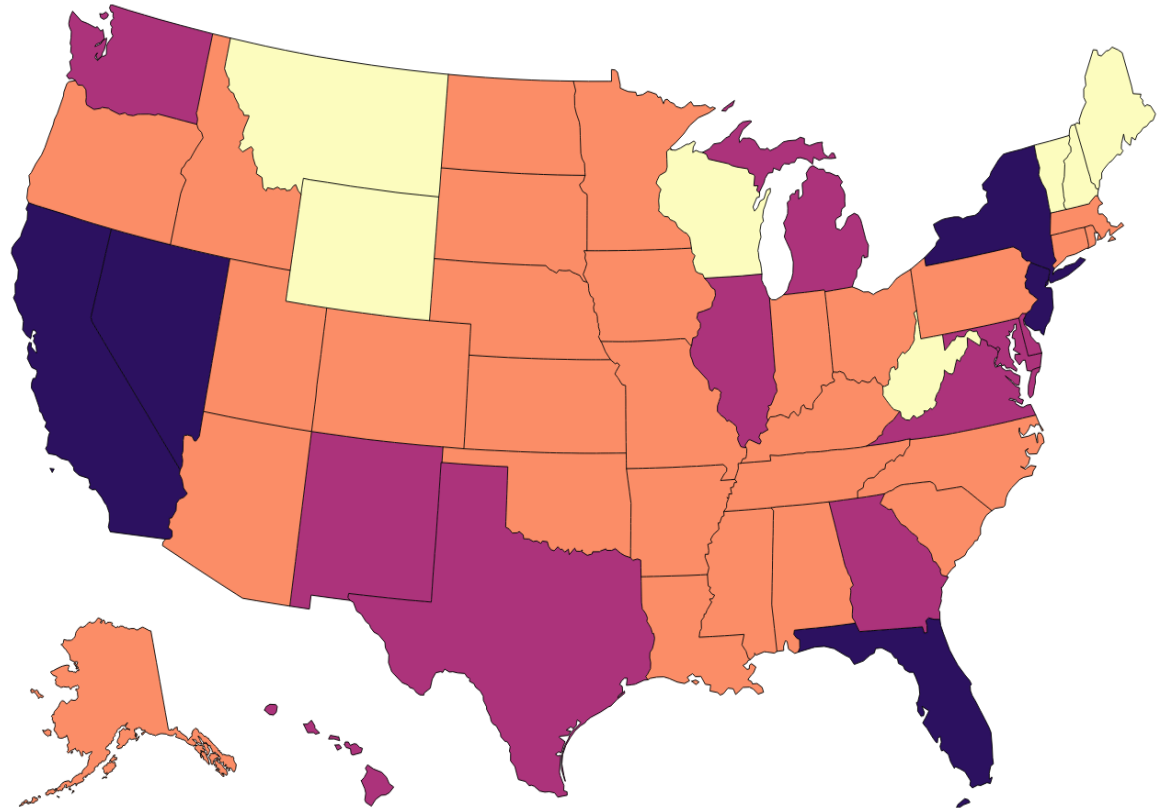
*When (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

## Share of economics majors who speak a language other than English at home

■ < 10% ■ 10%–20% ■ 20%–30% ■ ≥ 30%



Map: @JustinWolfers • Source: American Community Survey, 2018-2022



# Meet each student in the **language** they find most natural

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Available

*When (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

## Language spoken at home by Economics majors

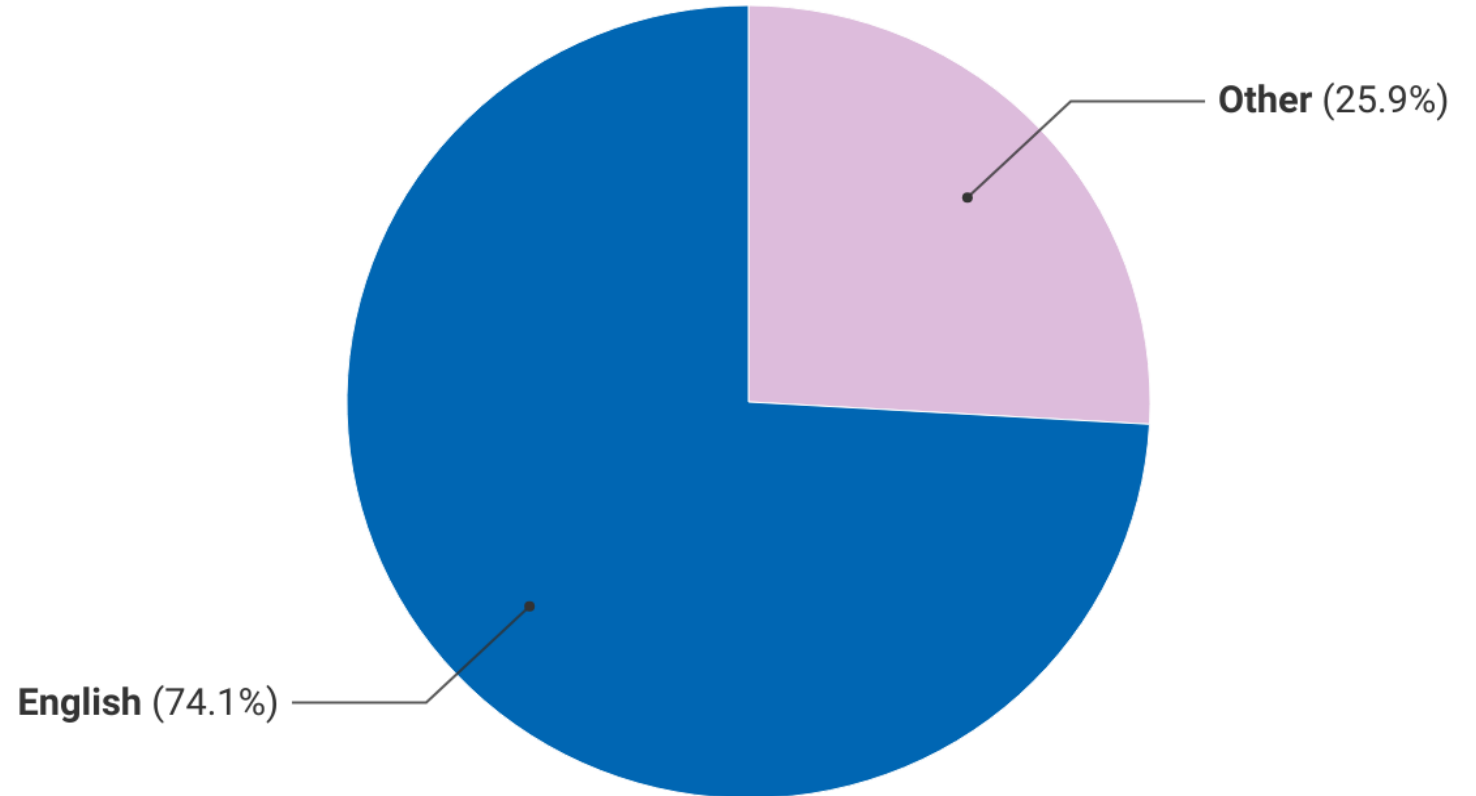


Chart: @JustinWolfers • Source: American Community Survey, 2018-2022

# Meet each student in the **language** they find most natural

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Works long hours

*Available when (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

## Which (non-English) languages are they speaking?

*Language spoken at home by economics majors among those who speak a language other than English at home*

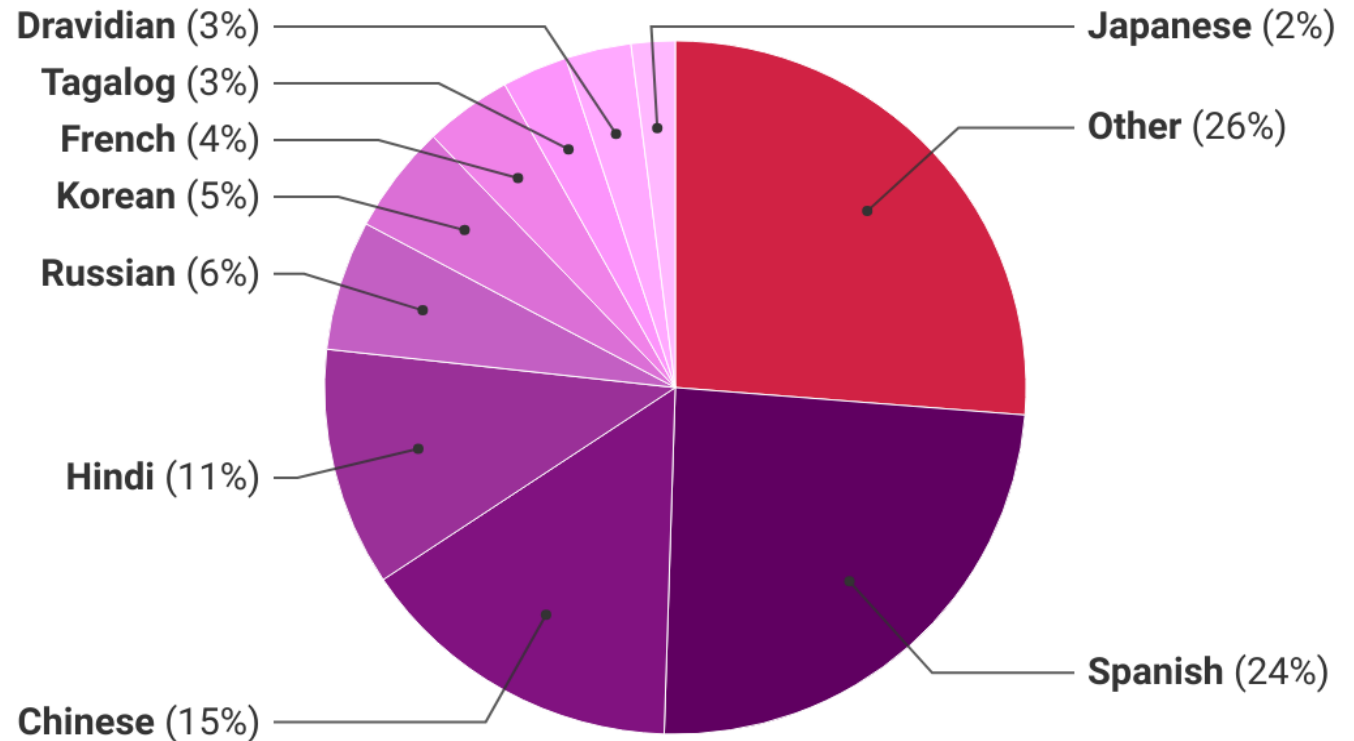


Chart: @JustinWolfers • Source: American Community Survey, 2018-2022

# Meet **every** student in the **language** they find most natural

## Smart

*Perfect knowledge of economics*

## Adaptable

*Uses definitions and tools consistent with your course*

## Empathetic

*Pays attention to what each student needs*

## Works long hours

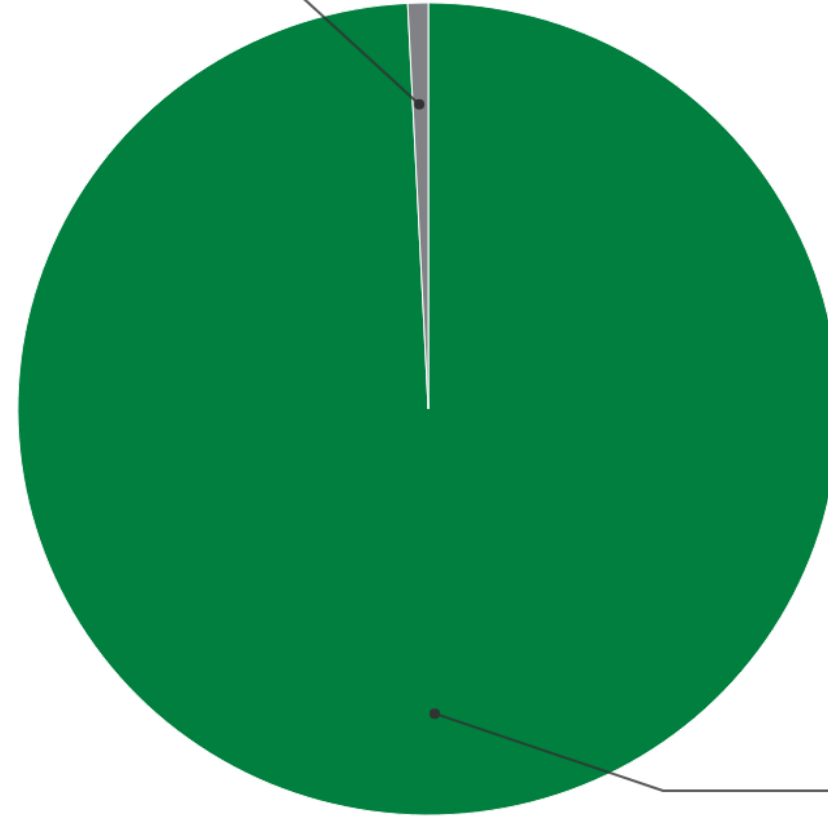
*Available when (and where) each student most need help*

## Multilingual

*Provide personalized support as needed*

## Language spoken at home by Economics majors

Not covered (0.8%)



Covered by ChatGPT  
(99.2%)

Chart: @JustinWolfers • Source: American Community Survey, 2018-2022

# Students need Tutoring not Tools

## ❑ Problem:

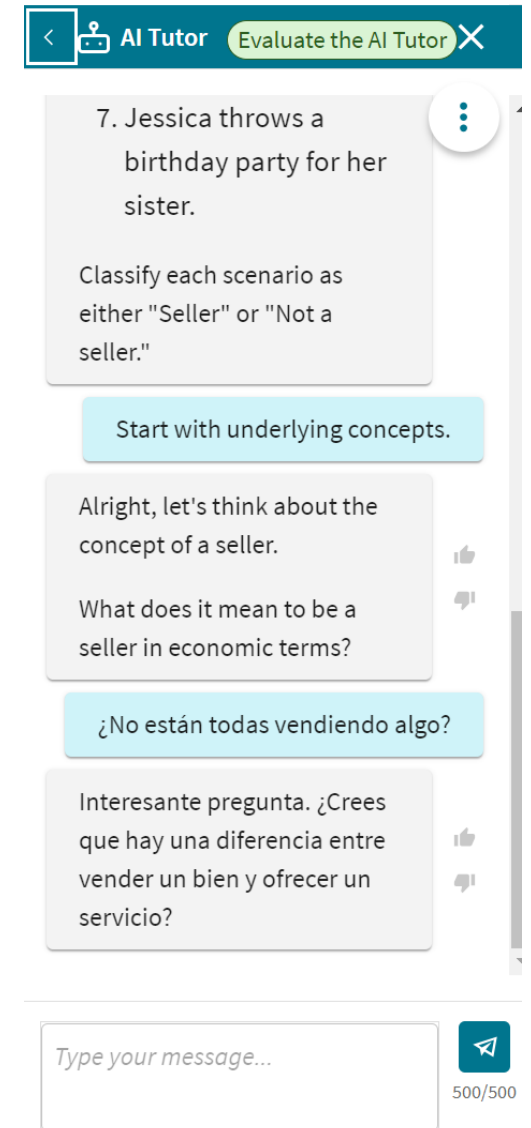
- ▶ Students are **heterogeneous** in their needs and abilities
- ▶ Feedback at scale is (has been) **homogeneous** or **non-existent**.

## ❑ Solution:

- ▶ Create a tool that provides students the agency to **personalize their own feedback** experience.

## ❑ Macmillan Learning's AI Tutor =

- ▶ Top tier LLM +
- ▶ Sophisticated and iteratively developed LLM instructions +
- ▶ High quality assessment and solution content +
- ▶ Student's exact wrong answer and human authored feedback +
- ▶ Interventions to mitigate gaming the LLM



The screenshot shows a chat interface for 'AI Tutor'. At the top, there is a navigation bar with a back arrow, the 'AI Tutor' name, and a button that says 'Evaluate the AI Tutor' with a close icon. The main chat area contains several messages:

- A system message: "7. Jessica throws a birthday party for her sister. Classify each scenario as either 'Seller' or 'Not a seller.'" followed by a light blue intervention box: "Start with underlying concepts."
- A user message: "Alright, let's think about the concept of a seller." with a thumbs-up icon.
- A user message: "What does it mean to be a seller in economic terms?" with a thumbs-down icon.
- A system message: "¿No están todas vendiendo algo?" followed by a light blue intervention box: "Interesante pregunta. ¿Crees que hay una diferencia entre vender un bien y ofrecer un servicio?" with a thumbs-up icon.

At the bottom, there is a text input field with the placeholder "Type your message..." and a character count "500/500". A send button with a paper plane icon is located to the right of the input field.

# Demo



# AI Tutor



# What are econ students saying? \*Direct Student Survey Results

---

It helps me solve problems more efficiently

It can read my responses and point me in the right direction with out giving me the answer, which allows me to think about the question myself.

I like that it helps you work through the problems if you get stuck. Sometimes, I get stuck on things that AI Tutor refuses to help me with.

It is walking you through the steps/concepts very well

It definitely helps clarify confusingly worded questions. It also helps clearly state what information from the course should be considered.

This working around the answers does nothing

NOTHING IT IS NOT HELPFUL AT ALL!

I like that it helps you work through the problems if you get stuck. Sometimes, I get stuck on things that AI Tutor refuses to help me with.

this sucks pls fix it

UGH

# How are econ student doing?

## Econ AI Tutor Usage 8/1/24 – 9/30/24

- ▶ 762,771 messages sent
- ▶ 172,719 conversations
- ▶ 34,203 students
- ▶ 614 courses
- ▶ 200 institutions

### Number of Econ Student Messages per Week

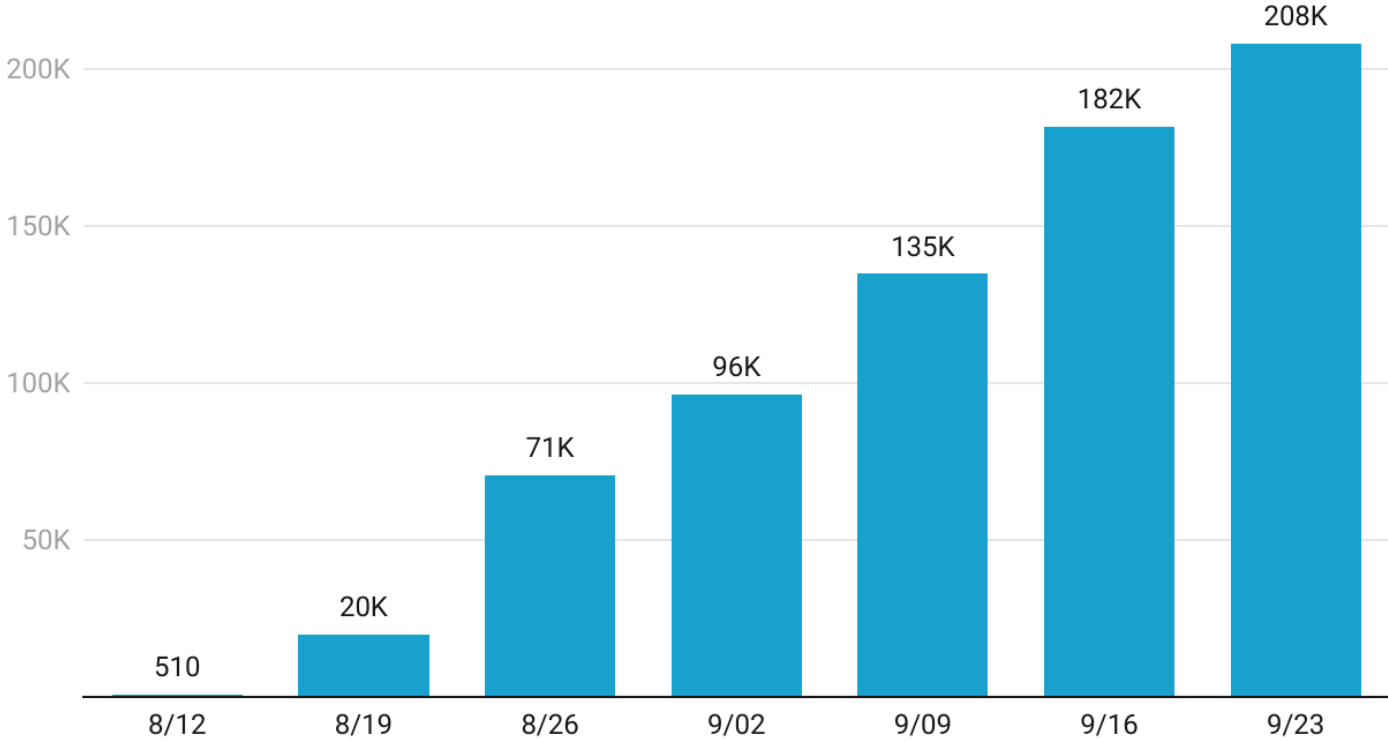


Chart: Joshua Hill • Source: Macmillan Learning

### Individual Message Rating by Econ Students

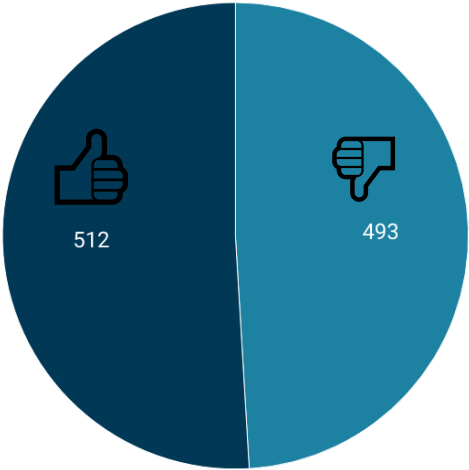


Chart: Joshua Hill • Source: Macmillan Learning

# Solving the **Learning Problem** is HARD

❑ Creating an **effective** and **successful** personalized tutor is **HARD**.

- ▶ Probabilistic systems like LLMs are risky
  - LLMs +students = wrangling cats
  - LLMs turn predictive systems into student realities
- ▶ Mitigating that risk is required!
  - Collaborate with Natural Language Processing experts to develop evaluation and monitoring standards
  - There are often observations that have no clear cause and effect!



# Personalized Exam Prep at Scale

## Idea #2: Midterm exam coach

---



+



# Introducing your midterm econ coach

---



# Demo

# Class Prep at Scale: Just-in-time polling questions

# Educator Opportunity: JIT Polling Questions



## ❑ Problem:

- ▶ Drafting active learning questions for your exact class and content is costly.

## ❑ Solution:

- ▶ Create a tool that allows instructors to generate question drafts with a few inputs.

## ❑ iClicker Question Creator =

- ▶ Top tier LLM +
- ▶ Prompt engineering to nudge the model toward active learning +
- ▶ Inputs from instructors +
- ▶ Instructor review and edit

### Create Question Set Cancel

Which discipline do you want to use to create your questions?

Discipline

Economics

How would you like to create your questions?

Question Topic

Provide a few key words or sentences to help create questions. Do not input personal information.

Impact of a price change of a complement or substitute on the supply and demand of a good. Use real world scenarios. Focus on both shifts and equilibrium changes.

338

What question types do you want to create?

Input the number of each question type you'd like to generate. You can only generate a max of 10 questions at a time.

Multiple Choice	2
Short Answer	1
Numeric	1
Multiple Answer	0
Total Questions	4

Clear Form

Create (4)

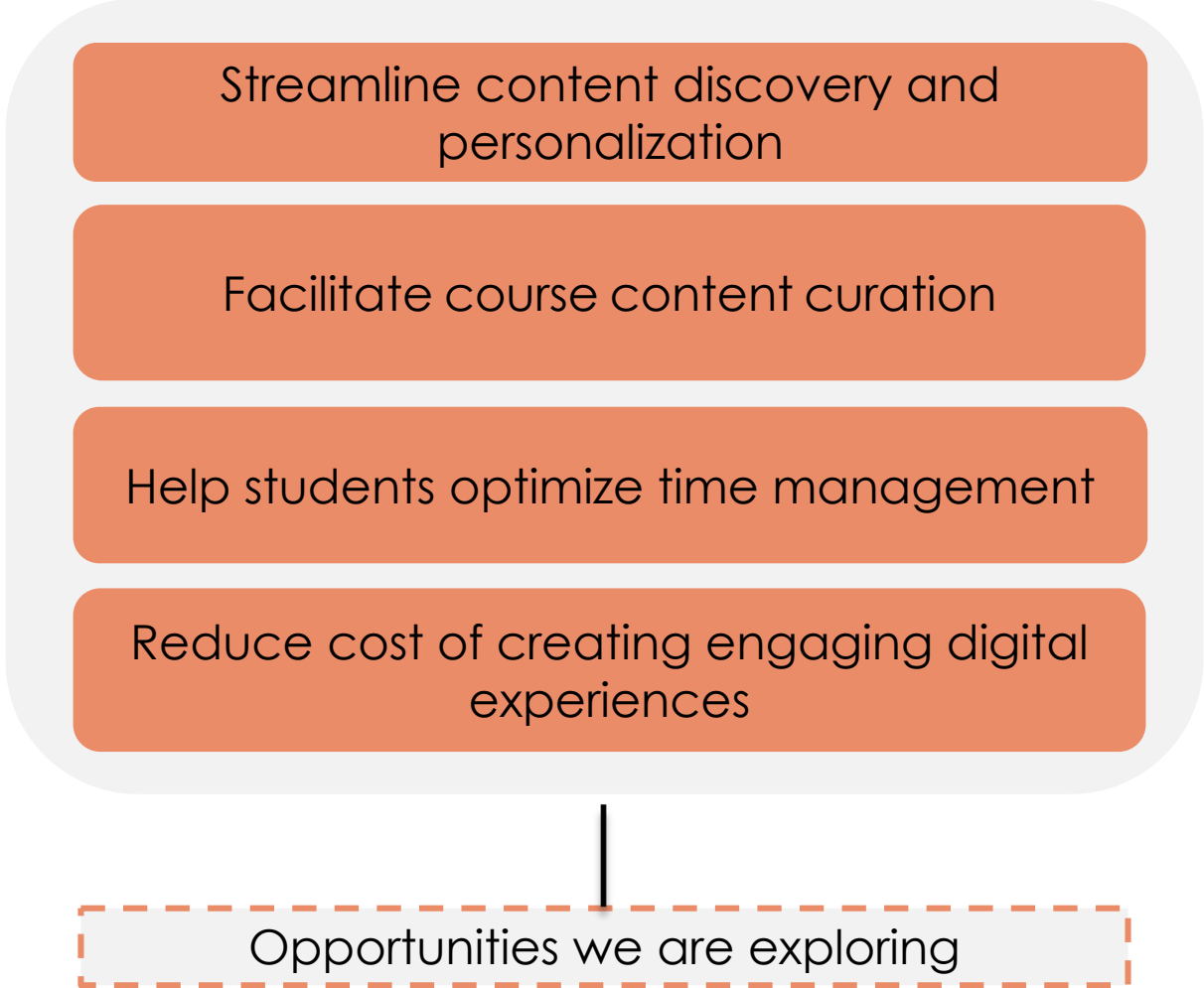
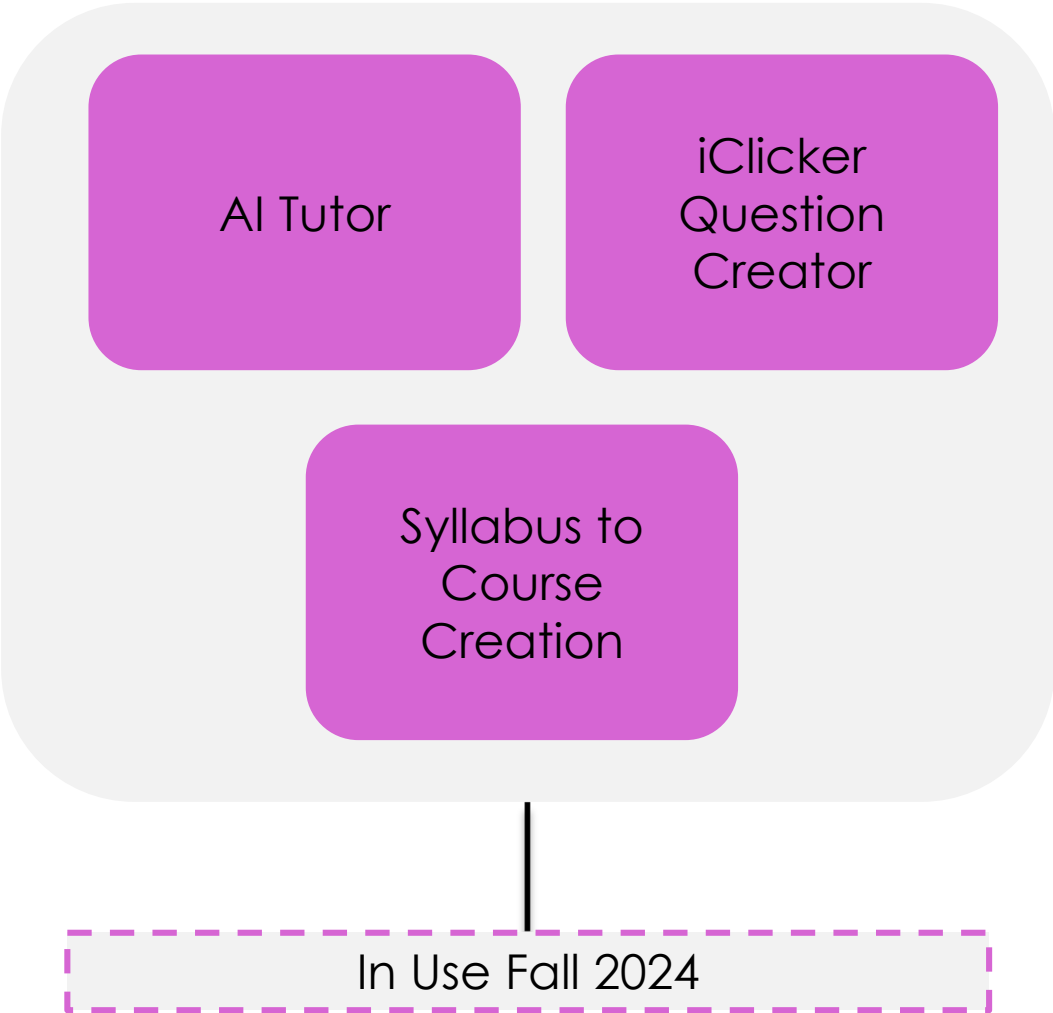


# Demo



iClicker Cloud

# Opportunities We Are Pursuing



# Our goal for you

---



## Yes, there will be some challenges along the way

---



- ❑ We are going to continue to develop AI tools for econ instructors
- ❑ We would love to hear from you about what you would find useful
- ❑ Email:  
[jwolfers@umich.edu](mailto:jwolfers@umich.edu)  
[joshua.hill@macmillan.com](mailto:joshua.hill@macmillan.com)

