

1 Motivation

There is anecdotal evidence that college students are moving away from fields exposed to AI capabilities [computer science majors](#), and [radiologists](#).

Expectations about demand for a profession, which turn out to be wrong ex post, can lead to shortages or gluts in a profession.

For example, in 2016, AI researcher Geoffrey Hinton, who would go on to win the Nobel Prize in 2024, said “[people should stop training radiologists now](#)”, and that “within five years [by 2021] deep learning is going to do better than radiologists.” Radiologists are still employed in 2025 and some companies report a “[supply-demand gap](#)” for radiologists.

Classes of Questions

- Are AI-exposed education fields suffering a decline or deceleration in enrollment since the launch of ChatGPT? **Evidence today.**
- If so, is this an ex post reaction to AI-induced decline in hiring/wages, or primarily about *ex ante expectations* about future automation risk?
 - For example, we might find that the slowdown in computer science enrollment is a reaction to labor market deterioration, whereas the shortage in radiology was induced by medical students moving away from a field they expected to be automated.
- Most speculatively, can we predict which professions might have shortages or gluts if automation ends up being less of a big deal than we thought? This might be viewed as the labor market side of an AI bubble.

This idea is in the air among economists: [\[Brynjolfsson et al., 2025\]](#) released a paper on declining headcount of early career software developers and customer service employees since ChatGPT was launched. In the conclusion, they speculate that “The adoption of new technologies typically leads to heterogeneous effects across workers, resulting in an adjustment period as workers reallocate from displaced forms of work to new forms with growing labor demand (Autor et al., 2024). **Such endogenous adjustment may already be happening with AI, with emerging evidence of shifts in college majors away from AI-exposed categories such as computer science.**”

2 Literature

2.1 College Major Choice Broadly

Altonji et al Handbook chapter (2016) Missing Manual, Dynarski et al, NBER WP (2013)

2.2 AI and Jobs

Acemoglu/Restrepo Tasks frameworks, and earlier. Felten et al, 2021.
Felten et al, 2023.

Timmerman, 2025
[Source](#).

2.2.1 AIs in Medicine

Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology

3 Data on College Major

Ideally want very up-to-date college major choice. The premium is on recency rather than micro data, I think. I'll say what I did first, then what else is out there.

4 Sational Student Clearinghouse (NSC)

Many papers use this. It is reported at about a 6-month lag at the state-major level.

4.1 Decomposing a change in enrollment from t to $t + 1$

We will be interested in changes in a major over time $\Delta_{m,t \rightarrow t+1} = \text{Enrollments}_{m,t+1} - \text{Enrollments}_{m,t}$.

In principle, this can come from:

- Existing majors: People in that major could have left Computer Science (and either switched to another major or dropped out)
- More people graduated with that degree than declared that major
 - This could be due to overall trends in college attendance
 - Or due to major-specific trends.

What about a deceleration in a major, i.e. $\Delta_{m,t+1 \rightarrow t+2} - \Delta_{m,t \rightarrow t+1} > 0$? This could be due to:

- If inflow growth slows or outflow growth speeds up (e.g., a bulge graduating, rising switch-outs, rising leaves, or fewer new declarants than last year).

appens

I cannot separate these without a panel of micro data, but it's useful to be clear on where variation could come from.

4.2 Aggregates

I have aggregates. In fact, at the state-level, but I am not sure what I would do with states.

4.3 Micro Data

Not sure that I need this, although literally aggregating to the year-level, like I am doing now seems too coarse; we're talking basically 5 datapoints (for each major). Seeing it at the institution level would be cool (E.g. are people just dropping CS majors at good colleges). For about \$1000 they will give you enrollment data for any list of students. But as far as I can tell, there is no product for just "all students".

4.4 Three sources of measurement error [[Dynarski et al., 2013](#)]

Matching Errors. To get student-level data, NSC has to match its records to those from other data sources. So there will inevitably be match errors between those. **Roverage (both across and within universities)**

Across universities: The NSC does not get enrollment numbers from all universities. This means totals will not match the true number of college students. "As of fall 2011, the NSC reports that they cover 93 percent of postsecondary enrollment." [[Dynarski et al., 2013](#)]. This number is up to 97% now as of 2025.

Within universities: "It is not necessarily reasonable to assume that NSC participating institutions report degree information for all of their students" **FERPA blockage** "FERPA is a federal law that protects the privacy of student education records. The law applies to all schools that receive funds from the U.S. Department of Education. Under FERPA, both students and schools can block their enrollment and degree information. So, NSC cannot release student-level information if the record is "FERPA-blocked." "

4.5 RPS

4.6 ACS

At least a 2-year lag. And doesn't even include enrollments, as far as I can tell.

4.7 IPEDS

Roughly a 1-year lag. "For example, the most recent data release on September 23, 2025, included provisional data from the Fall 2024 data collection" This is more comprehensive than

4.8 National Survey of College Graduates

This data is biennial.

5 Analyses

5.1 How to define AI-exposed majors.

We calculate a static measure of AI exposure by college major. Doing so requires a way to map from occupations, where past researchers have quantified AI exposure, to majors.

Table 1: Construction of Major-Level AI Exposure Scores

Step	Description
<i>Data Sources</i>	
(1)	Occupation-level AI exposure: Felten et al. (2023) Language Modeling AIOE scores for 774 SOC occupations
(2)	Occupation-major linkage: IPUMS ACS 2013-2017 (5-year estimates) for ages 22-35, linking field of degree to current occupation. We use this vintage to ensure SOC code definitions align with those in Felten et al. (2021).
(3)	Enrollment data: National Student Clearinghouse 2023 undergraduate enrollment by 4-digit CIP code
<i>Construction of Major-Level AI Exposure</i>	
(1)	Link majors to occupations: For each major c , identify all ACS individuals whose field of degree maps to that major
(2)	Assign occupation-level exposure: Each individual i receives the AI exposure score of their occupation: $\text{AIOE}_{\text{occ}(i)}$ from Felten et al. 2023.
(3)	Aggregate to major level: Calculate major-level exposure as the weighted average AIOE across all occupations that major’s graduates enter: $\text{AI_Exposure}_c = \frac{\sum_{i \in c} w_{i,c} \times \text{AIOE}_{\text{occ}(i)}}{\sum_{i \in c} w_{i,c}}$ where $w_{i,c}$ is the weight on person i used for major c .

Since ACS uses FOD (Field of Degree) codes while enrollment data uses CIP (4-digit Classification of Instructional Programs) codes, we harmonize using a many-to-many NCES crosswalk (ii) weight by empirical enrollment shares in each CIP when an ACS FOD maps to many CIPs.

This all goes into the calculation of $w_{i,c}$.

For example, if a FOD maps to multiple CIPs, which we denote $\text{CIP}(\text{FOD}_i)$, then we split that person using empirical enrollment shares in each of those CIPs, from 2023. PERWT_i is the normal person weight from ACS.

$$w_{i,c} = \text{PERWT}_i \times \frac{\text{Enrollment}_c}{\sum_{c' \in \text{CIP}(\text{FOD}_i)} \text{Enrollment}_{c'}}$$

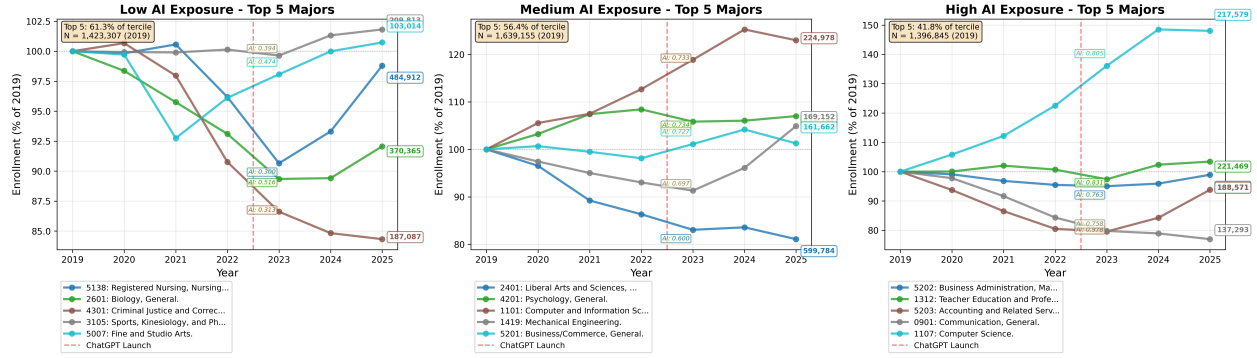


Figure 1: Enrollment by AI occupational exposure over time. Data from NSC.

5.1.1 Is AI exposure the measure we want?

One challenging thing is whether we are looking under the lamp post with an easy-to-get measure of AI-induced unemployment. But the Felten et al measure is not designed to capture that, as they say, “we remain agnostic to whether occupational exposure to generative AI is likely to result in automation vs. augmentation of occupations”.

5.2 Descriptives

5.2.1 Quantities of enrollment

Qualitatively, we see a deceleration in computer science and computational and information science between 2024 and 2025, as reported on in the popular press. But other trends that are likely unrelated to AI, and have been occurring since at least 2019, e.g. the continued decline in Communication (high AI exposure) and Criminal Justice (low AI exposure) majors.

Try to get Spring and fall numbers for each this; in principle, this double your sample size within major

5.2.2 Wages by AI exposure over time

be careful about when you measure wages here. Need it to be recent enough to be useful and whether you want this at the college major level or the occupation level.

The most recent wage \times major data is the 2024 ACS data.

5.3 Motivating DiDs

It seems like if anything was happening, it was not happening until this year. So we should not expect significant results. But these are the kinds of DiDs we might want to run. (Note, I have run these, and none of the results were significant)

5.4 DiD w.r.t. AI Exposure

5.5 DiD w.r.t. AI Exposure, controlling for wages

6 A simple model of major choice in response to substitutes

The key features we desire in a model are:

- Two majors; roughly, one should be “low-risk”, it pays off the same in each state, and one should be high-risk, it pays off nothing in the state you get automated in, and a lot otherwise.
- Think of this as nursing and computer science.
- This risk is basically macro-risk, not uncertainty about your own type.
- Expectations about labor market outcomes for each major; this need not be your expectation about right when you graduate, it could be further down the line. We could summarize this into a “terminal value”, which compresses all future wages.
- An ability to “simulate” what happens when expectations change.
- A delay between when you decide your major and get your first job (I am not sure this is important)

For now, assume people have equal ability to do any major. And everyone has the same ex-ante beliefs.

So, two majors. C and N . We assume one major maps directly into one profession and there is no unemployment. Two potential states of the world, low demand L (AIs automate computer science jobs) and H (humans keep those jobs). With probability $p_H = 1 - p_L$. We can thus think of H as a status quo, and L as a deviation from that status quo.

Preferences We assume people have preferences only over their final wage. Conditional on employment they each provide 1 unit of labor inelastically. One way to think about this is people are extremely elastic to their choice of major, but completely inelastically supply labor once school is done.

Furthermore, assume, to start, that wages are only determined by employer demand. And employer demand only varies between the two states, H and L .

Demand for each major upon graduation are as follows.

$$D_R(H) > D_S(L) = D_S(H) > D_R(L)$$

For simplicity, we thus denote demand for safe major in either state as D_S

Under risk neutrality, the following indifference in expected wages must hold, although we have not explained how wages are determined yet.

$$p_H w_R(H) + p_A w_R(A) = w_S(H) \tag{1}$$

6.1 Demand

We now quantify the distribution of majors, assume a mass 1 of students choose between the two majors, with n_R choosing computer science, and $n_S = 1 - n_R$ choosing nursing.

6.1.1 Completely Elastic Demand

We assume demand for employment in each degree is perfectly elastic at wage $w_R(H) > w_S > w_R(L)$. This fully pins down the equilibrium. This means any split of majors, and in fact, any wages, are an equilibrium as long as Equation ?? is satisfied.

6.1.2 Inelastic Demand

(I do not think you can have inelastic demand and inelastic supply without just assuming everything about the final eqm)

6.1.3 Partially Inelastic Demand

We now specify demand a little more narrowly, for all three majors we have a linear demand curve of the form

$$D = a - wb$$

And we specify demand for each major as a shift in the demand curve, with the same constant elasticity b .

$$D_{H,L} = a_H - w_{R,H}b, D_S = a_S - w_Sb, D_{R,L} = a_L - w_{R,L}b$$

Where $a_H > a_S > a_L$. On the supply side, wages must satisfy the following indifference condition.

$$E[w_R] = w_S$$

$$p_H w_{R,H} + p_L w_{R,L} = w_S$$

$$p_H \left(\frac{n_R - a_H}{-b} \right) + p_L \left(\frac{n_R - a_L}{-b} \right) = \frac{n_S - a_S}{-b}$$

$$p_H(n_R - a_H) + p_L(n_R - a_L) = (n_S - a_S)$$

$$n_R + p_H(-a_H) + p_L(-a_L) = (n_S - a_S)$$

And quantities must sum to 1.

$$n_R + p_H(-a_H) + p_L(-a_L) = (1 - n_R - a_S)$$

Denoting $E[a]$ as expected demand intercept for the risky major.

$$-E[a] = 1 - 2n_R - a_S$$

$$a_S - E[a] - 1 = -2n_R$$

$$\frac{1 + E[a] - a_S}{2} = n_R$$

The deviation, from 50-50, depends on how much higher expected demand is for the risky major relative to the safe major.

$$\frac{1 + a_S - E[a]}{2} = n_S$$

Wages are a little more interesting because the ex post realizations matter, and no indifference conditions need to hold ex post. Using the demand curves and the n_R above,

$$\begin{aligned} w_S &= \frac{a_S - (1 - n_R)}{b} = \frac{E[a] + a_S - 1}{2b}, \\ w_{R,H} &= \frac{a_H - n_R}{b} = \frac{a_S - 1 + a_H + (1 - p_H)(a_H - a_L)}{2b}, \\ w_{R,L} &= \frac{a_L - n_R}{b} = \frac{a_S - 1 + a_L - p_H(a_H - a_L)}{2b}, \end{aligned}$$

where $E[a] = p_H a_H + (1 - p_H) a_L$.

Under $a_H > a_S > a_L$ and $0 < p_H < 1$, we have

$$w_{R,H} > w_S > w_{R,L}.$$

Ex post wages in the risky sector depend on realized demand, and the more confident people were ex ante of a bad state, the higher ex post wages are in the high state occur.

$$\frac{\partial w_{r,H}}{\partial p_H} = \frac{(a_L - a_H)}{2b} < 0$$

This is not a deviation from rational expectations, it is just that as $p_H \rightarrow 0$, entering a profession becomes more like a lottery ticket. This is the idea of a shortage of workers entering the risky profession, but wages rising ex post when the risk is not realized, (or - not modeled - if this is realized later than expected). Running example here is radiologists. So far, I think this is nothing new. This is the same idea as in finance there can be a “bubble” with rational expectations.

7 Promise of this idea

I find area interesting, but it seems very crowded, and it is not obvious I have a strong advantage in data, etc.. Clearly many people are asking the question.

Is this an interesting question even if there is a null result? It clearly seems more interesting if AI exposed majors were falling off a cliff.

Are we too early for this to be studied concretely? A few extra years might make this more convincing.

References

- Erik Brynjolfsson, Bharat Chandar, and Ruyu Chen. Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence. Working paper, Stanford Digital Economy Lab, November 2025. URL <https://digitaleconomy.stanford.edu/publications/canaries-in-the-coal-mine/>.
- Susan M Dynarski, Steven W Hemelt, and Joshua M Hyman. The missing manual: Using national student clearinghouse data to track postsecondary outcomes. Technical Report w19552, National Bureau of Economic Research, 2013.