Firm Entry and AI Exposure: Evidence from US Industries

Guillermo Gallacher*

April 24, 2026 – Version 0.4

Preliminary. Comments welcome.

Abstract

I study the relationship between sectoral AI exposure and firm entry in the United States, focusing on the impact of ChatGPT's release in November 2022. I first document a positive association between a sector's AI exposure and its post-2022 business application growth. I then characterize its timing: the effect emerges tentatively in early 2023, fades through early 2024, and then re-emerges persistently and grows through the end of the sample (latest observation is March 2026). While I cannot rule out confounding from other sector-specific trends, this dynamic pattern is consistent with AI adoption and capabilities progressively lowering entry costs and/or raising entrepreneurial returns in exposed sectors.

*Email: guillermogallacher@gmail.com. Web: guillgall.github.io. I thank Daisoon Kim for helpful comments.

1 Introduction

After the COVID-19 shock, aggregate business applications surged (Decker and Haltiwanger, 2023). The release of ChatGPT in November 2022 seems to have increased firm entry, too, especially from 2025 onwards, as seen in Figure 1.

This expansion is not uniform across sectors: industries such as Professional Services have increased more than the average, while others such as Construction have lagged behind. From this plot a simple observation arises: while the former sector is exposed to AI, the latter is less so. Does this pattern hold across all sectors, though?

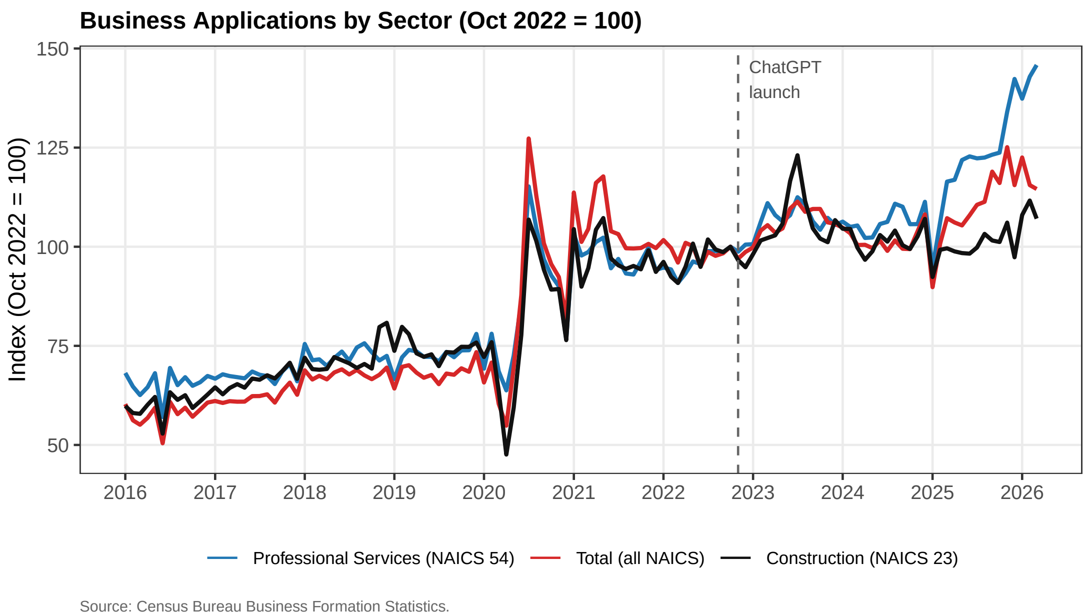

Figure 1: Business applications indexed to October 2022 = 100. Professional Services (NAICS 54), Total (all NAICS), and Construction (NAICS 23). Seasonally adjusted. FRED interactive version.

In this paper I show that it does: I first document a positive association between a sector's AI exposure and its post-2022 business application growth. I then characterize its timing: the effect emerges tentatively in early 2023, fades through early 2024, and then re-emerges and grows persistently thereafter (latest observation is March 2026).

While I cannot rule out confounding from other sector-specific trends, this dynamic pattern is consistent with AI adoption and capabilities progressively lowering entry costs and/or raising entrepreneurial returns in exposed sectors.

Contribution to the literature Most of the "economics of AI" debate focuses on workers and incumbent firms: which jobs are at risk, and what happens to productivity and wages. A related question has received less attention: does AI alter the incentives to start a business?

Chatterji et al. (2025) provide a conceptual treatment of how transformative AI reshapes firm organisation, adoption incentives, and competitive dynamics, a natural complement to the reduced-form evidence presented here.

I combine the AI and firm entry dimensions, by using business formation data as in Decker and Haltiwanger (2023), and papers cited therein. Currently, popular AI macro frameworks, such as Acemoglu (2025), abstract from firm entry. My findings point towards entry being an important dimension in understanding the macroeconomic impact of AI.

2 Data

2.1 Business Formation Statistics

The Census Bureau BFS provides monthly, seasonally adjusted counts of new Employer Identification Number (EIN) applications at the 2-digit NAICS level, derived from IRS Form SS-4 filings. EIN applications are a leading indicator of actual business formation, but the latter has considerable publication lags (Decker and Haltiwanger, 2023). I use BFS applications as my main proxy for firm entry, and report robustness across different BFS metrics. The monthly 2-digit series runs through March 2026 (the latest available at time of writing). The weekly 3-digit series, used in Section 4.2, lags by several months and currently extends through December 2025.

2.2 AI Exposure

Following Felten et al. (2021), I use the AI Industry Exposure (AIIE) index, which maps occupational AI exposure scores to 4-digit NAICS industries through BLS employment weights. The underlying data are publicly available at github.com/AIOE-Data/AIOE. I aggregate to 2-digit (and 3-digit) NAICS using employment-weighted averaging, with 2019 BLS OES industry employment counts as weights, consistent with Felten et al. (2021)'s own use of 2019 employment in constructing the index. Table 1 reports the AIIE 2-digit score alongside summary statistics of the BFS series for each sector in the sample.

Table 1: AI Industry Exposure and Business Formation Statistics by Sector

Sector	NAICS	AIIE	Oct 2022 SA BA (thousands)	Mar 2026	% Change
Finance & Insurance	52	2.04	18.1	19.3	+6.7%
Professional, Scientific, Technical Svcs	54	1.71	52.6	76.7	+45.8%
Management of Companies	55	1.70	3.2	6.3	+95.1%
Educational Services	61	1.34	6.2	8.0	+29.0%
Information	51	1.11	8.2	11.4	+38.3%
Health Care & Social Assistance	62	0.59	25.1	33.1	+31.8%
Real Estate & Rental	53	0.35	24.7	26.1	+5.9%
Wholesale Trade	42	0.33	10.6	8.6	-18.0%
Utilities	22	0.20	0.6	0.6	+11.2%
Retail Trade (NAICS 44–45)	RET	-0.05	69.9	77.2	+10.4%
Other Services	81	-0.25	39.9	42.4	+6.2%
Manufacturing (NAICS 31–33)	MNF	-0.41	7.5	7.0	-7.1%
Admin. & Support Services	56	-0.58	31.4	36.2	+15.3%
Arts, Entertainment, & Recreation	71	-0.61	13.4	15.1	+13.1%
Mining, Quarrying, Oil & Gas	21	-0.66	0.5	0.5	-3.2%
Transportation & Warehousing (48–49)	TW	-0.95	36.1	32.4	-10.3%
Construction	23	-1.10	42.9	46.0	+7.1%
Accommodation & Food Services	72	-1.12	23.4	26.8	+14.3%
Agriculture, Forestry, Fishing	11	-2.00	4.7	3.9	-17.0%

Notes: AIIE = AI Industry Exposure index from Felten et al. (2021), constructed from O*NET task descriptions and aggregated from 4-digit to 2-digit NAICS by employment-weighted averaging. SA BA columns report seasonally adjusted monthly business applications (thousands) in October 2022 (last pre-ChatGPT month) and March 2026 (latest available). % Change is the percentage difference between the two, computed from unrounded values. Sectors sorted by AIIE score descending.

3 Results

3.1 Regression

Figure 2 plots the percentage change in business applications from October 2022 to March 2026 against each sector's AI exposure at the 2-digit NAICS level, illustrating the cross-sectional positive correlation.

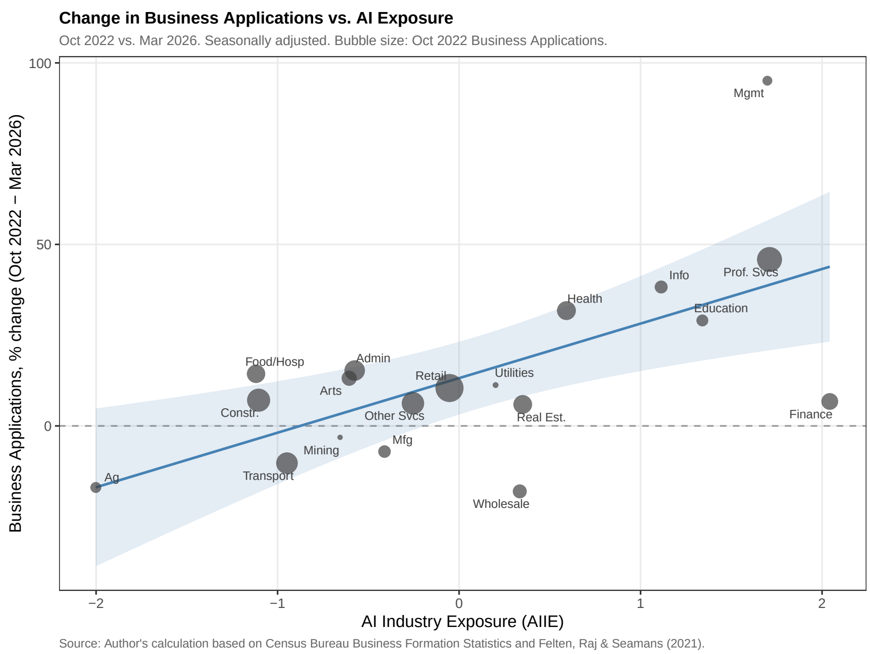

Figure 2: Change in business applications versus AI exposure by 2-digit NAICS sector.

Table 2 formalises the scatter with a cross-sectional regression of the form

$$\%\Delta\text{BA}_i \;=\; \alpha + \delta\,\text{AIIE}_i + \varepsilon_i, \tag{1}$$

where $\%\Delta\text{BA}_i$ is the percentage change in business applications from October 2022 to March 2026 for sector i, and AIIE_i is the AI Industry Exposure index. Column (1) is unweighted OLS; column (2) weights observations by pre-period mean business applications, giving larger sectors proportionally more influence. The slope is positive and statistically significant in both specifications.

Table 2: Cross-Sectional Regression: AI Exposure and Change in Business Applications, October 2022 to March 2026

	OLS	WLS
AI Exposure (AIIE)	15.030***	11.900***
	(4.380)	(3.258)
Constant	13.159**	13.531***
	(4.754)	(3.283)
Sectors	19	19
R^2	0.409	0.440

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
OLS and WLS (weighted by sector mean BA, Oct 2022) estimates. Outcome: percentage change in business applications from October 2022 to March 2026. Each observation is a 2-digit NAICS sector ($N = 19$).

3.2 Event Study

The cross-sectional regression captures the average differential across the full post-November 2022 window but does not reveal when the effect emerges. To characterise its timing, I estimate a panel event study that interacts AI exposure with month-specific indicators:

$$\log(\mathrm{BA}_{it}) \;=\; \alpha_i + \gamma_t + \sum_{\tau \neq 0} \beta_\tau \left(\mathbf{1}\{t - t_0 = \tau\} \times \mathrm{AIIE}_i\right) + \varepsilon_{it}, \tag{2}$$

where t_0 denotes October 2022 ($\tau = 0$; the last fully pre-ChatGPT month), $\mathbf{1}\{t - t_0 = \tau\}$ is an indicator for the month lying τ periods from the reference, and $\tau = 0$ is the omitted period so that $\beta_0 = 0$ by construction. The ChatGPT launch (November 2022) corresponds to $\tau = +1$. Each β_τ captures the differential log change in business applications between sectors with higher and lower AI exposure in event-time τ, relative to the same differential in October 2022. This dynamic specification is the unrestricted analogue of the pooled DiD in Appendix A (eq. (3)), which imposes a single post-period coefficient.

Figure 3 presents the event-study coefficients. The pre-period estimates ($\tau < 0$) are small and statistically indistinguishable from zero, supporting the parallel-trends assumption.

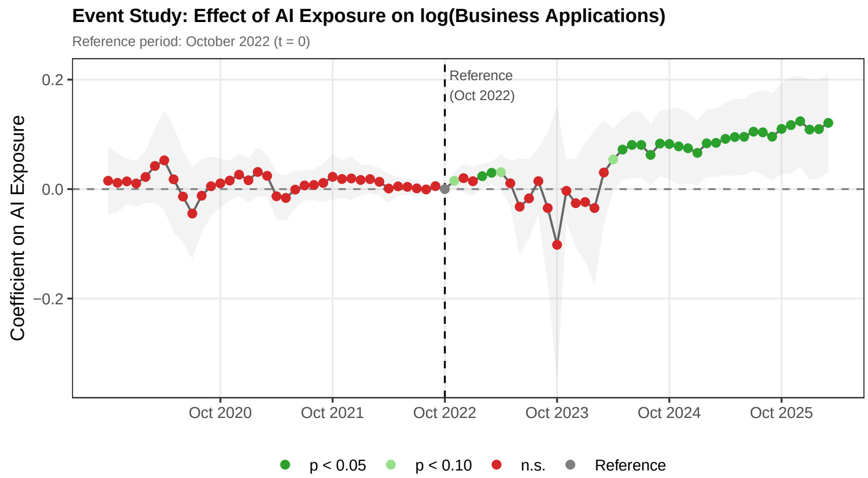

Figure 3: Event-study estimates: effect of AI exposure on log(Business Applications) relative to October 2022 ($\tau = 0$). Shaded band is the 95 % confidence interval based on sector-clustered standard errors.

A handful of coefficients are statistically significant in the first months immediately following the ChatGPT release, suggesting an early, tentative response. These effects then recede into insignificance over roughly a year, before re-emerging persistently from $\tau = +18$ (April 2024, $p < 0.10$) and $\tau = +19$ (May 2024, $p < 0.05$) onward. From May 2024, every subsequent monthly coefficient is statistically significant at the 5 % level or better, and point estimates trend upward through the end of the sample (Figure 3). Figure 9 in the appendix shows the same specification using November 2022 ($\tau = +1$) as the reference period; the qualitative pattern is unchanged.[1]

[1] Figure 8 in the appendix replicates the dynamic pattern using monthly cross-sectional regressions rather than a panel model: for each month t, I regress the percentage change in business applications since October 2022 on AIIE across the 19 sectors, obtaining a time path of OLS slopes $\widehat{\beta}_t$. The coefficient path mirrors the panel event-study estimates, providing a specification-check that does not rely on two-way fixed effects.

4 Robustness Checks

4.1 Alternative Business Applications Metrics

I probe robustness by substituting the outcome with the full suite of available BFS series spanning the formation funnel.

On the applications side, I use corrected business applications (CBA), which purge filings unlikely to yield an employer business; high-propensity applications (HBA), which restrict to applicants with a high predicted probability of hiring; and applications with planned wages (WBA), indicating explicit intent to hire.

On the formations side, I use seasonally-adjusted formations within four and eight quarters (SBF4Q, SBF8Q). The realised formation counts (BF4Q, BF8Q) are excluded because the Census has not yet published these beyond 2022 due to publication lags.

Figure 4 replicates Figure 1 separately for each series (Professional Services, high AI exposure; Construction, low AI exposure; and the total), so that any sector-level divergence and its timing are directly visible. Across all series, Professional Services pulls away from Construction after ChatGPT's launch, with the gap widening from mid-2024 onward, broadly mirroring the pattern in Figure 1. Notably, this divergence is present in HBA and the formation series (SBF4Q, SBF8Q) as well as in BA, suggesting it reflects genuine heterogeneity in entry activity rather than noise from marginal applications. The pre-2022 co-movement of the two sectors across panels also supports the parallel-trends assumption underlying the event study.

One notable feature of the aggregate series is that while BA has surged since 2024, WBA and CBA have declined in aggregate, hinting that AI-driven entry may be tilted toward leaner, wage-light ventures, consistent with an AI-enabled "solopreneur" boom. Importantly, this aggregate decline does not contradict the cross-sectional evidence: Table 3 shows positive and significant AIIE gradients for both WBA and CBA, meaning AI-exposed sectors still outpace low-exposure ones even as the aggregate series drifts down. Whether this compositional shift matters for the paper's conclusions is ultimately an empirical question, addressed below.

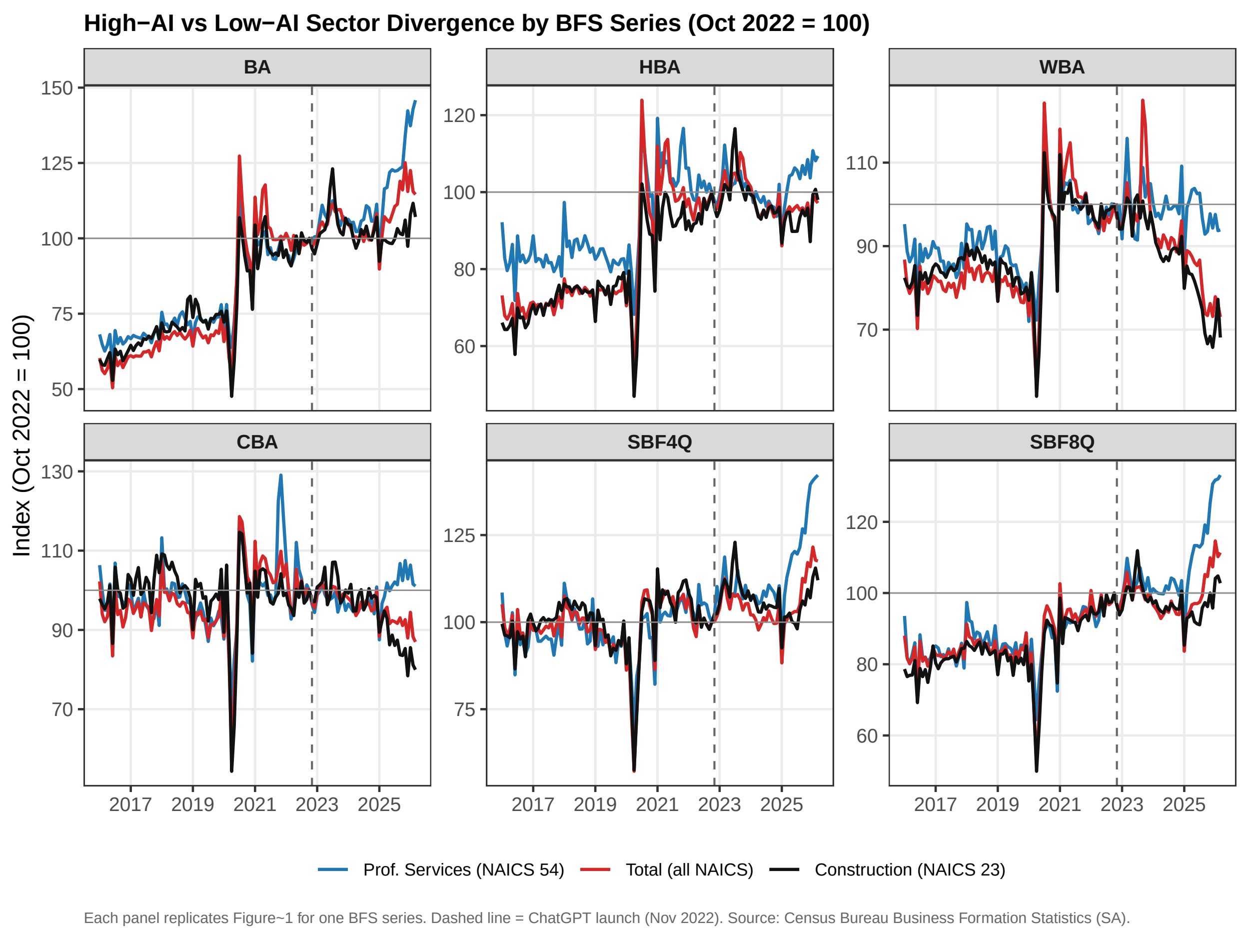

Figure 4: Each panel shows Professional Services (NAICS 54), Total, and Construction (NAICS 23), indexed to October 2022 = 100. Dashed vertical line marks the ChatGPT launch. Seasonally adjusted. Source: Census Bureau Business Formation Statistics.

Figure 5 presents event-study coefficients for all six series. The pre-period estimates are uniformly small and insignificant, ruling out differential pre-trends. Post-launch, coefficients are positive across all series and periods, with no evidence of a reversal. Statistical significance is strongest and most sustained for BA, CBA and WBA, where green dots accumulate consistently from mid-2023 onward. The formation series (SBF4Q, SBF8Q) and HBA show the same positive direction but some individual months fall short of conventional significance thresholds even as the overall trend is positive.

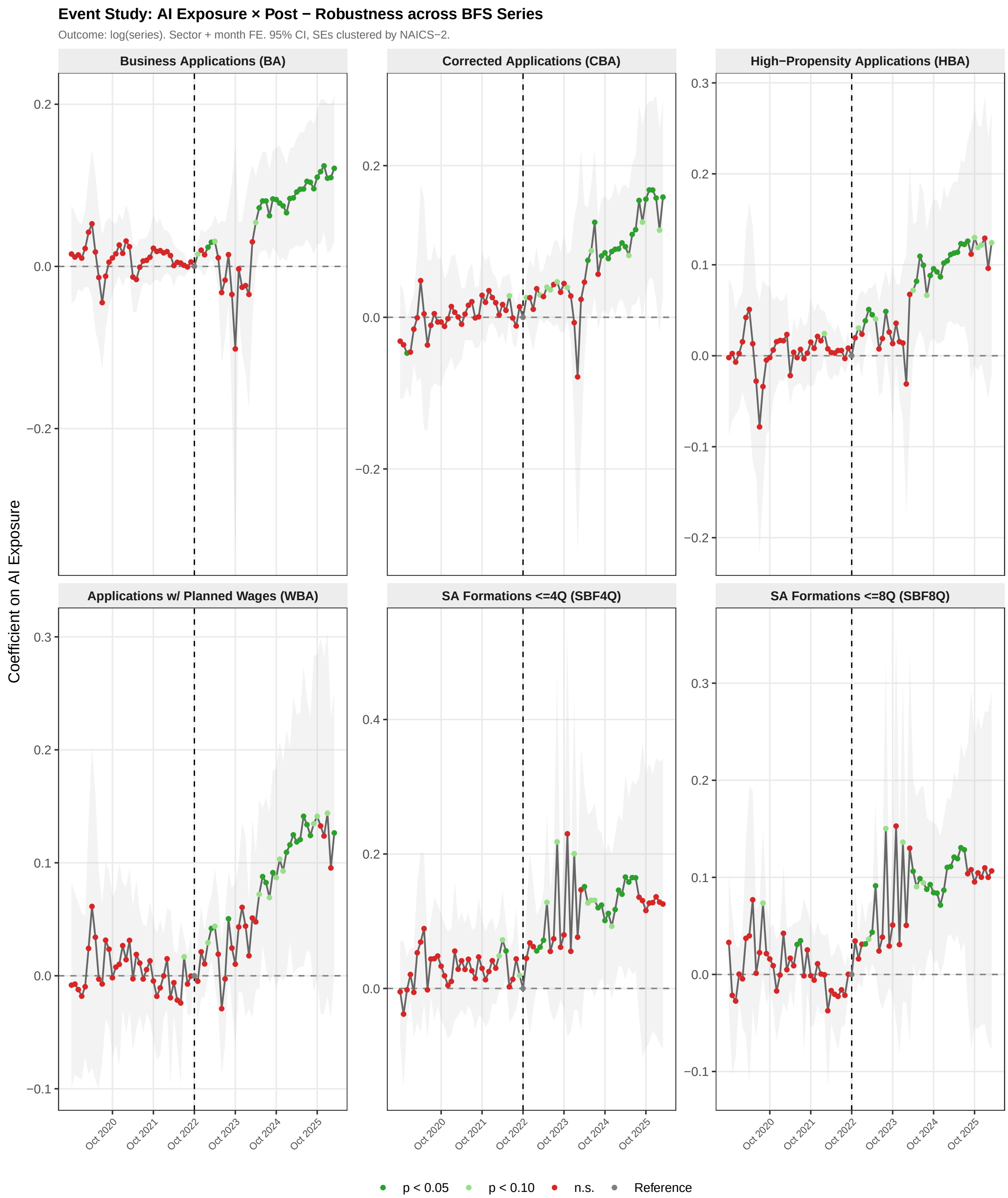

Figure 5: Event-study estimates for six BFS series. Dashed vertical line marks October 2022 ($\tau = 0$, reference period).

Table 3 formalises the cross-sectional gradient with a simple OLS regression of the per-

Table 3: Cross-Sectional OLS: AI Exposure and Change in BFS Series (October 2022 – March 2026)

	(1) BA	(2) HBA	(3) WBA	(4) CBA	(5) SBF4Q	(6) SBF8Q
AI Exposure (AIIE)	15.030***	11.911*	10.918**	14.977**	25.367*	18.193*
	(4.380)	(5.708)	(4.458)	(5.168)	(14.409)	(9.939)
Constant	13.159**	-12.604*	-27.626***	-15.641**	28.184*	15.977
	(4.754)	(6.196)	(4.839)	(5.610)	(15.640)	(10.788)
Sectors	19	19	19	19	19	19
R^2	0.409	0.204	0.261	0.331	0.154	0.165

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

Each column is an OLS regression of the percentage change in the indicated BFS series since October 2022 on the 2-digit NAICS AIIE index. The end date is March 2026 for BA, CBA, WBA, and HBA; the latest available date is used for formation series (SBF4Q, SBF8Q) if that precedes March 2026. Heteroskedasticity-robust standard errors in parentheses.

centage change in each series on AIIE (Figure 10 in the appendix shows the corresponding scatter plots). The AIIE coefficient is positive and statistically significant across all six series (at least at the 10 % level).

4.2 3-digit NAICS

Figure 6 replicates the motivating scatter at the finer 3-digit NAICS level, comparing seasonally adjusted mean weekly business applications in October 2022 with the latest available month across 83 sectors, using weekly BA data.[2]

[2]The BFS adopted NAICS 2022 codes, while the AIOE index (Felten et al., 2021) and the WFH teleworkability measure (Dingel and Neiman, 2020) use NAICS 2017. The 2022 revision reorganised retail (44–45) and information (51) sub-sectors, renaming or merging eight 3-digit codes. I remap affected AIOE and WFH scores to the new codes using the official Census Bureau 2022-to-2017 NAICS concordance (census.gov/naics/concordances). Without this crosswalk the inner join with BFS silently drops those eight sectors.

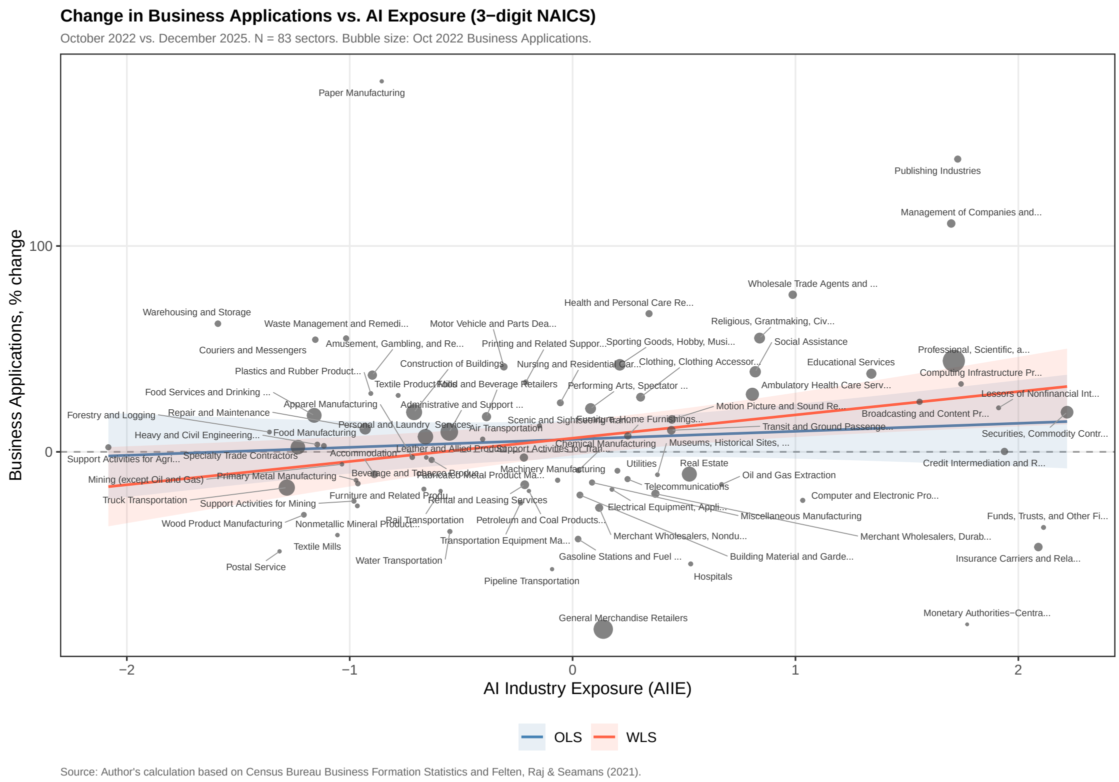

Figure 6: Change in weekly business applications versus AI exposure, 3-digit NAICS sectors. October 2022 vs. latest available month (seasonally adjusted). Bubble size proportional to October 2022 mean weekly Business Applications. Blue line: unweighted OLS. Red line: WLS weighted by October 2022 Business Applications. Shaded bands: 95 % confidence intervals.

The positive cross-sectional gradient visible at the 2-digit level persists at the 3-digit level but only after weighting (Figure 6). Table 4 formalises this. The unweighted OLS slope is near zero and statistically insignificant, reflecting substantial noise among the many small 3-digit sectors. Once sectors are weighted by their pre-period application volume, however, the slope rises sharply to a similar coefficient to the monthly estimates reported before: among the sectors that account for the bulk of economic activity, higher AI exposure is strongly associated with faster application growth.

Figure 7 presents the event-study coefficients from estimating equation (2) on the weekly 3-digit NAICS panel. After seasonal adjustment, the event study uses weekly event time ($t = 0$ is the week of 28 November 2022, the ChatGPT launch week; the omitted reference

Table 4: Cross-Sectional Regression: AI Exposure vs. Change in Weekly Applications (3-digit NAICS)

	OLS	WLS
AI Exposure (AIIE)	3.887	11.312***
	(4.630)	(3.947)
Constant	6.087	6.641
	(4.626)	(4.217)
Sectors (3-digit)	83	83
R^2	0.009	0.092

* p < 0.1, ** p < 0.05, *** p < 0.01
OLS and WLS (weighted by pre-period mean weekly BA) at 3-digit NAICS. Outcome: percentage change in mean weekly business applications from October 2022 to December 2025.

period is $t = -1$, the week of 21 November 2022), sector and year-week fixed effects, and observations weighted by each sector's pre-period mean weekly BA. Weighting by sector size is motivated by the cross-sectional evidence in Table 4: the relationship between AI exposure and application growth is concentrated among larger sectors, so weighting ensures that identification comes from the same margin.

Compared to the 2-digit estimates, the 3-digit event study shows no initial increase immediately after ChatGPT's launch, but a positive and persistent association does emerge and strengthen over time.

One notable difference is a positive pre-period association during the pandemic years, which is absent at the 2-digit level. This likely reflects the fact that AI-exposed sectors (predominantly knowledge-intensive industries) disproportionately benefited from the shift to remote work, generating an entry boost that predates the AI shock. To the extent this constitutes a differential pre-trend, it would bias the event-study estimates upward; however, the year-week fixed effects and the sector-level seasonal adjustment absorb much of this variation, and the pre-ChatGPT window in the event study itself is flat. I discuss this in the next subsection.

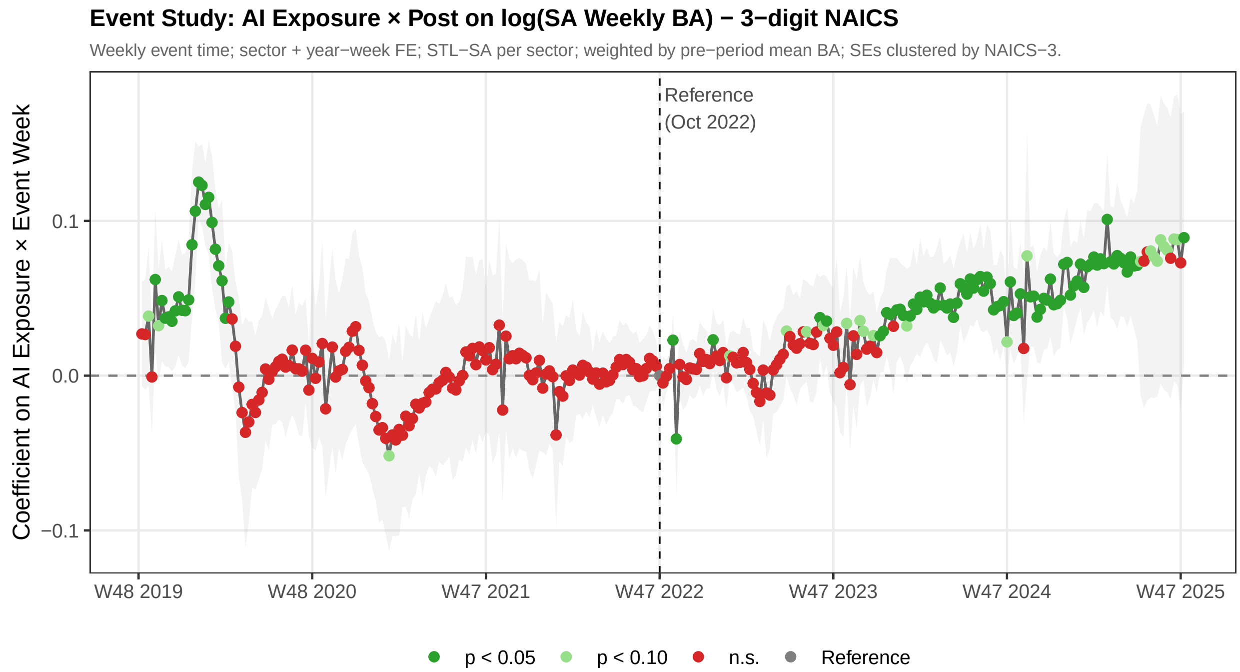

Figure 7: Event-study estimates at 3-digit NAICS. Outcome: log seasonally adjusted weekly business applications (STL decomposition, multiplicative, per sector). Weekly event time; $t = 0$ is the week of 28 Nov 2022 (ChatGPT launch). Dashed vertical line marks the reference week ($t = -1$, week of 21 Nov 2022). Observations weighted by pre-period mean weekly BA. Sector and ISO year-week fixed effects; SEs clustered by NAICS-3.

4.3 Remote-Work Confounding

A potential concern is that remote-work capacity (work-from-home, or WFH) rather than AI exposure drives the results. A first piece of evidence against this is provided by the pre-period of the event study (Figure 3): in the months and years preceding the ChatGPT launch, high- and low-AI-exposure sectors show no differential trend in business applications (the 3-digit event study shows some pandemic-era co-movement in 2020, which I return to below).

Importantly, the AIIE index is highly correlated with the Dingel and Neiman (2020) WFH teleworkability measure across sectors (Gallacher, 2024), but this correlation is structural rather than coincidental: sectors where AI can substitute or augment tasks are precisely those where work is computer-mediated and location-flexible. Remote-work capacity is almost a byproduct of being AI-exposed, not an independent channel. Treating WFH as a separate confounder would amount to controlling for a characteristic that is largely implied by the treatment itself.

That said, Tables 6 and 7 in the appendix report three specifications (AI only, WFH only, and both jointly) at both 2-digit (OLS) and 3-digit NAICS (WLS weighted by sector size). At the 2-digit level (Table 6), the R^2 is similar whether AI or WFH enters alone, and adding the second regressor barely changes it, as each variable explains most of what the other does. The AIIE point estimate is stable when WFH is added, while the WFH coefficient loses precision, a near-collinearity artefact with only 19 sectors. The 3-digit replication (Table 7) covers 82 sectors (one sector lacks WFH teleworkability data from Dingel and Neiman (2020) and is dropped) but the same pattern holds: depending on the specification, one index crowds out the other.

In this light, the positive pre-period association visible in Figure 7 around 2020 is not surprising: the onset of the pandemic triggered a temporary boom in firm entry in precisely those sectors that are both AI- and remote-work-intensive. This 2020 effect is likely distinct from the post-ChatGPT divergence that is the focus of this paper.

4.4 US States

Figure 11 in the appendix reports a state-level robustness check using the AI Geographic Exposure (AIGE) index of Felten et al. (2021) and state $\times$ month panel. The state-level estimates are statistically insignificant, likely because state aggregation dilutes the within-sector AI variation that drives the main results; the sector-level estimates remain the primary evidence.

4.5 Language Modeling AIIE

As a final robustness check I replace the broad AIIE with the updated LM AIIE of Felten et al. (2023), which restricts the underlying task battery to language-model capabilities (reading comprehension, translation, and text generation). The two indices are highly correlated and results are very similar, as documented in the appendix (Figure 12).

5 Interpretation

The empirical findings are consistent with two distinct but complementary mechanisms through which AI can stimulate firm entry in exposed sectors.

1. The first is a reduction in fixed costs: language models lower the cost of tasks (legal drafting, coding, market research, content creation) that are disproportionately required to start and operate a firm in AI-exposed industries.

2. The second is a productivity gain: workers in exposed sectors produce more output per hour when augmented by capable language models, raising the expected returns to entrepreneurship and tilting the free-entry condition in favor of entry.

Decomposing the two channels would require auxiliary data on sectoral prices, output, or labor productivity.

The tentative response in early 2023, the fade through early 2024, and the persistent rising gradient thereafter could be due to endogenous and/or exogenous factors:

1. Endogenous experimentation and adoption: the pattern is consistent with an initial burst of experimentation following ChatGPT's launch, a period of consolidation as early adopters integrated new tools into workflows, and a second wave from mid-2024 as both adoption broadened and model capabilities made a discrete jump.

2. Exogenous AI advances ("Agentic AI"): importantly, the re-emergence and acceleration from mid-2024 coincides with a qualitative shift in AI capabilities, from con-

versational assistants toward agentic systems capable of executing multi-step tasks autonomously. Agentic AI could lower entry costs more than earlier language models did, by automating not just individual tasks but entire workflows such as customer acquisition, legal compliance, or software deployment.

A third implication concerns the composition of entry rather than its level. The dispersion panel in Figure 4 shows that while total business applications (BA) have surged in exposed sectors since 2024, applications with planned wages (WBA) and corrected applications (CBA) have declined in aggregate, suggesting that the marginal entrant in AI-exposed sectors is wage-light relative to historical norms. This is consistent with the productivity channel: if AI raises output per worker, the equilibrium firm in an exposed sector requires fewer employees to be viable, shifting entry toward leaner, solo ventures. The observation that the divergence between Professional Services and Construction is visible in high-propensity applications (HBA) and the formation series (SBF4Q, SBF8Q) as well as in raw BA rules out a purely compositional explanation based on marginal or speculative filings.

6 Directions for Future Research

A natural next step is to validate the application-based results against realized firm entry from the Business Dynamics Statistics or Business Employment Dynamics data at the sector level, once post-2022 vintages become available. A study of actual application-to-employer transitions, post-entry survival, and employment growth in AI-exposed sectors must wait for the availability of administrative micro data; as Decker and Haltiwanger (2023) note, this will require the confidential Longitudinal Business Database (LBD), which tracks post-entry dynamics and job-to-job flows. Incorporating alternative AI exposure measures (Eloundou et al., 2024) would further strengthen robustness. Extending the analysis to other countries would shed light on whether the US pattern reflects idiosyncratic institutional features or a broader regularity in how general-purpose technologies reshape the incentives to start a business. Job posting data matched to new entrants could directly test whether AI-exposed sectors see fewer hires per new firm, providing micro-level evidence on the wage-light entry margin documented here (Bahaj et al., 2024). Finally, embedding these findings in a quantitative multisector macro model with endogenous firm entry (Bilbiie et al., 2012) would allow one to assess the aggregate implications of AI-driven entry booms concentrated in high-exposure sectors.

7 Conclusion

This paper documents a positive cross-sectional relationship between sectoral AI exposure and business applications in the United States following the release of ChatGPT in November 2022. The effect is not immediate: it emerges tentatively in early 2023, fades through early 2024, and then re-emerges persistently and grows through the end of the sample. The dynamic pattern, interpreted through the lens of Section 5, traces a plausible AI adoption path and is consistent with both fixed-cost reduction and productivity gains in exposed sectors, though alternative explanations based on unobserved sector-level trends cannot be fully ruled out.

References

Daron Acemoglu. The simple macroeconomics of AI. *Economic Policy*, 40(121):13–58, 2025. doi: 10.1093/epolic/eiae042.

Saleem Bahaj, Sophie Piton, and Anthony Savagar. Business creation during COVID-19. *Economic Policy*, 39(119):611–648, 2024. doi: 10.1093/epolic/eiae008.

Florin O. Bilbiie, Fabio Ghironi, and Marc J. Melitz. Endogenous entry, product variety, and business cycles. *Journal of Political Economy*, 120(2):304–345, 2012. doi: 10.1086/665825.

Aaron Chatterji, Daniel Rock, and Eduard Talamàs. Transformative AI and firms. In Ajay K. Agrawal, Erik Brynjolfsson, and Anton Korinek, editors, *The Economics of Transformative AI*, chapter 7. University of Chicago Press, 2025. URL https://www.nber.org/chapters/c15327.

Ryan A. Decker and John Haltiwanger. Surging business formation in the pandemic: Causes and consequences? *Brookings Papers on Economic Activity*, Fall:1–74, 2023. URL https://www.brookings.edu/wp-content/uploads/2023/09/Decker-Haltiwanger_16820-BPEA-FA23_WEB.pdf.

Jonathan I. Dingel and Brent Neiman. How many jobs can be done at home? *Journal of Public Economics*, 189:104235, 2020. doi: 10.1016/j.jpubeco.2020.104235. URL https://github.com/jdingel/DingelNeiman-workathome.

Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702):1306–1308, 2024. doi: 10.1126/science.adj0998.

Edward W. Felten, Manav Raj, and Robert Seamans. Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal*, 42(12):2195–2217, 2021. doi: 10.1002/smj.3325. URL https://github.com/AIOE-Data/AIOE.

Edward W. Felten, Manav Raj, and Robert Seamans. Occupational heterogeneity in exposure to generative AI. SSRN Working Paper No. 4414065, 2023. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4414065.

Guillermo Gallacher. Remote work and AI: Exposed, but not complementary? Evidence from the USA. Working paper, 2024. URL https://guillgall.github.io/files/rw_ai.pdf.

A Pooled Difference-in-Differences Estimates

For completeness, I also estimate the effect of AI exposure (AIIE_i) on business applications (BA_{it}) using the pooled difference-in-differences (DiD) model

$$\log(\text{BA}_{it}) \;=\; \alpha_i + \gamma_t + \beta\,(\text{Post}_t \times \text{AIIE}_i) + \varepsilon_{it}, \tag{3}$$

where i indexes 2-digit NAICS sectors, t indexes calendar months, Post_t is an indicator equal to one from November 2022 onwards, α_i are sector fixed effects absorbing time-invariant differences in application rates, γ_t are month fixed effects absorbing aggregate cyclical variation, and standard errors are clustered by sector. The coefficient β identifies whether sectors with higher AI exposure experienced a differential change in business formation after ChatGPT, restricting the post-period response to a single constant.

Table 5 reports the DiD estimates. All three columns pool the entire post-November 2022 period into a single *Post* indicator, with columns (1)–(3) using log business applications, the indexed BA outcome, and raw BA levels respectively.

The pooled *Post* coefficient is statistically insignificant because the treatment effect is time-varying: pooling the full post-November 2022 window into a single indicator averages near-zero early-period effects with the large, persistent effects that only emerge from early 2024 onwards, attenuating the estimate toward zero. The event study in Section 3.2 makes this dynamic explicit. The negative sign on the raw BA levels coefficient in column (3) reflects the same attenuation: in levels the pre-trend noise dominates the pooled estimate, and the sign should not be interpreted as a negative effect.

Table 5: Difference-in-Differences Estimates

	(1) log(BA)	(2) BA Index	(3) BA Level
Post × AI Exposure	0.049	7.255	-312.365
	(0.053)	(7.304)	(1823.743)
Observations	4959	4959	4959
R^2	0.982	0.689	0.834
Sector FE	✓	✓	✓
Month FE	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Standard errors clustered by 2-digit NAICS sector.

B Extra Figures and Tables

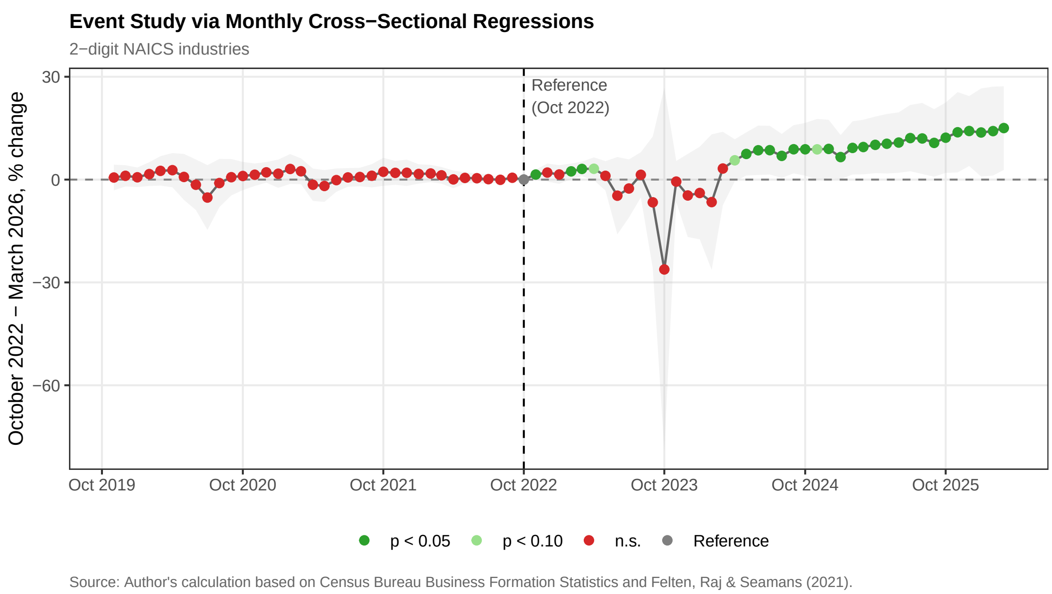

Figure 8: Monthly cross-sectional regressions of the percentage change in business applications (relative to October 2022) on AI Industry Exposure across 2-digit NAICS sectors. Each point is the OLS slope $\widehat{\beta}_t$ from a separate OLS regression for that month; shaded band is the 95 % confidence interval using HC1 robust standard errors. Points coloured by significance.

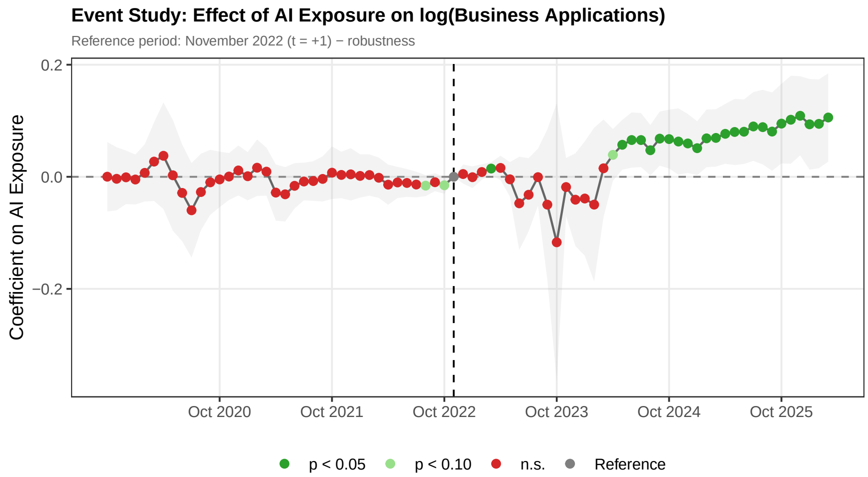

Figure 9: Event-study estimates using November 2022 ($\tau = +1$) as the reference period. All other specification details are identical to Figure 3. Shaded band is the 95 % confidence interval based on sector-clustered standard errors.

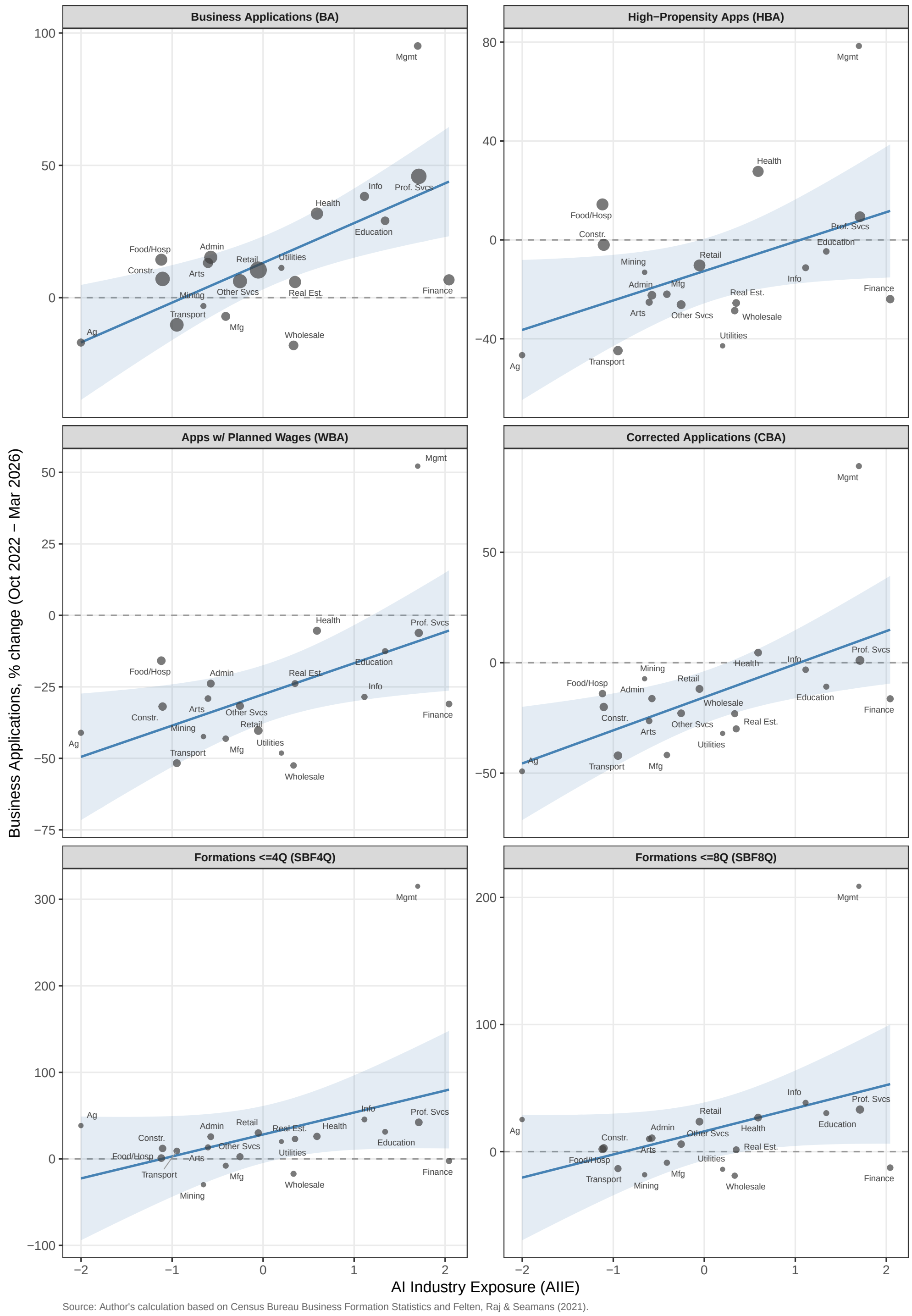

Figure 10: Cross-sectional scatter of percentage change in each BFS series since October 2022 against AI Industry Exposure (AIIE). Each dot is a 2-digit NAICS sector, sized by its October 2022 Business Applications. Blue line: OLS fit with 95% confidence band. Seasonally adjusted. Sources: Census Bureau Business Formation Statistics; Felten, Raj & Seamans (2021).

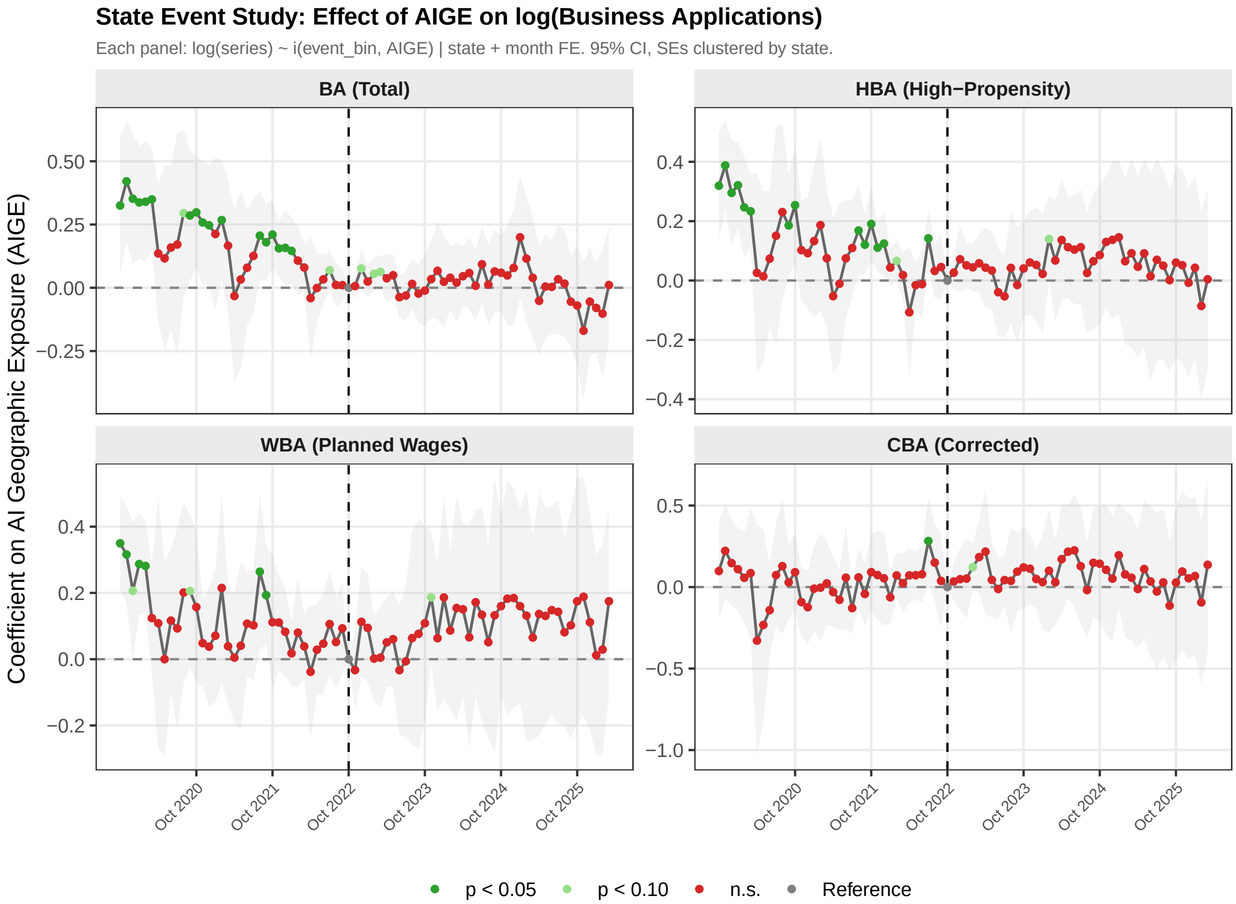

Figure 11: State-level event-study estimates for four BFS series. Each panel plots the interaction coefficient between AI Geographic Exposure (AIGE) and a month-relative-to-October-2022 indicator, estimated on the state $\times$ month panel. BA = total applications; HBA = high-propensity applications; WBA = applications with planned wages; CBA = corrected applications. Business formation series (BF4Q, BF8Q, etc.) are not published at state level by the Census Bureau. Shaded bands are 95 % confidence intervals based on state-clustered standard errors. All specifications include state and month fixed effects. Dashed vertical line marks October 2022 ($\tau = 0$, reference period).

Table 6: Robustness: Controlling for Remote-Work Confounding (October 2022–March 2026)

	(1) AI only	(2) WFH only	(3) AI + WFH
AI Exposure (AIIE)	15.030***		11.050
	(4.380)		(11.213)
WFH Share		64.102***	19.249
		(19.892)	(49.681)
Constant	13.159**	-7.359	6.951
	(4.754)	(8.336)	(16.748)
Sectors	19	19	19
R^2	0.409	0.379	0.415

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Outcome: % change in business applications, October 2022 to March 2026. Each observation is a 2-digit NAICS sector ($N = 19$). OLS. Column (1) replicates Table 2. $r(\text{AIIE}, \text{WFH}) = 0.91$; joint inclusion in column (3) inflates standard errors via multicollinearity.

Table 7: Robustness: WFH Control at 3-digit NAICS (October 2022–December 2025)

	(1) AI only	(2) WFH only	(3) AI + WFH
AI Exposure (AIIE)	15.768***		-7.354
	(3.076)		(6.673)
WFH Share		76.202***	103.019***
		(11.475)	(26.896)
Constant	-17.848***	-44.871***	-54.132***
	(3.291)	(5.334)	(9.949)
Sectors (3-digit)	82	82	82
R^2	0.247	0.355	0.365

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Outcome: % change in mean weekly business applications, October 2022 to December 2025. 3-digit NAICS; $N = 82$ (one sector lacks WFH teleworkability data). WLS weighted by October 2022 mean weekly applications.

Table 8: Robustness: WFH Control with Language Modeling AIIE (October 2022–March 2026)

	(1) LM AIIE only	(2) WFH only	(3) LM AIIE + WFH
LM AIIE (Language Modeling)	16.161***		12.437
	(4.541)		(10.092)
WFH Share		64.102***	17.663
		(19.892)	(42.471)
Constant	12.421**	-7.359	6.878
	(4.703)	(8.336)	(14.173)
Sectors	19	19	19
R^2	0.427	0.379	0.433

* p < 0.1, ** p < 0.05, *** p < 0.01
Outcome: % change in business applications, October 2022 to March 2026. Each observation is a 2-digit NAICS sector ($N = 19$). OLS. $r(\text{LM AIIE}, \text{WFH}) = 0.89$; joint inclusion in column (3) inflates standard errors via multicollinearity.

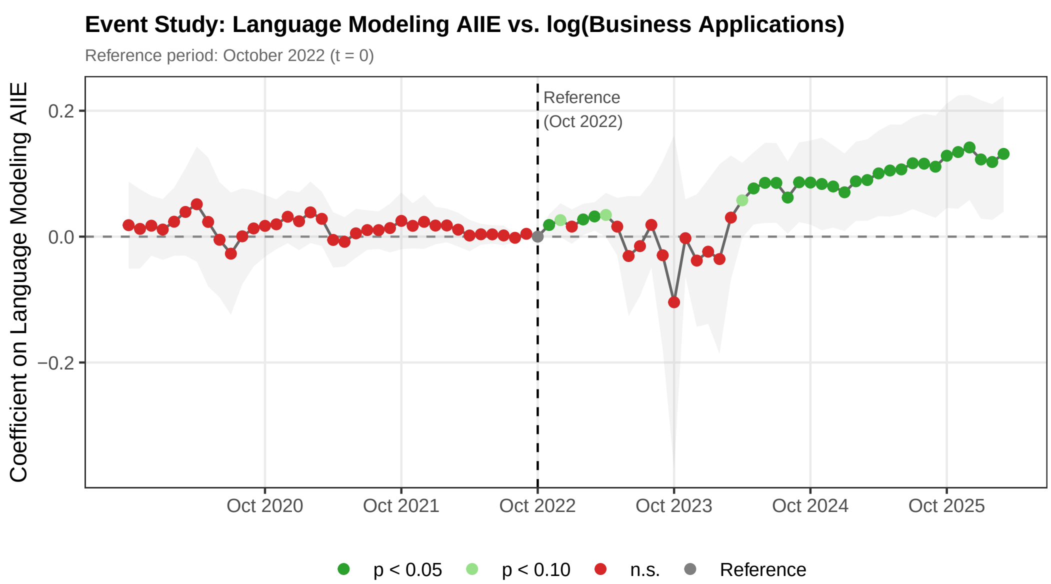

Figure 12: Event-study estimates using the Language Modeling AIIE of Felten et al. (2023) as the treatment variable, in place of the broad AIIE used in the main specification. The Language Modeling AIIE restricts the AI task battery to language-model capabilities only. Specification identical to Figure 3: two-way FE panel with sector and month fixed effects; standard errors clustered by 2-digit NAICS sector (19 clusters). Shaded band is the 95 % confidence interval; points coloured by significance level; dashed vertical line marks October 2022 ($\tau = 0$, reference period).

Table 9: Robustness: WFH Control with Language Modeling AIIE, 3-digit NAICS (October 2022–December 2025)

	(1) LM AIIE only	(2) WFH only	(3) LM AIIE + WFH
LM AIIE (Language Modeling)	15.253***		-6.468
	(3.276)		(5.804)
WFH Share		76.202***	97.729***
		(11.475)	(22.458)
Constant	-18.496***	-44.871***	-51.923***
	(3.391)	(5.334)	(8.270)
Sectors (3-digit)	82	82	82
R^2	0.213	0.355	0.365

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Outcome: % change in mean weekly business applications, October 2022 to December 2025. 3-digit NAICS; $N = 82$ (one sector lacks WFH teleworkability data). WLS weighted by October 2022 mean weekly applications. $r(\text{LM AIIE}, \text{WFH}) = 0.87$ at 3-digit level.