Expected Economic Impacts of Open-Source AI Models: Firm-Level Evidence

Presenter: Bernard Yeung
(ABFER, NUS Business School, and SUSTech)

[Preliminary and incomplete. Please do not circulate.]

Co-authors: Y. Huang, S. Lee, and Y. Zhou

2026 November

Disclaimer

- The views expressed in this publication are those of the authors and do not necessarily reflect the views of the BIS or its member central banks.

Motivation

- AI is a general-purpose technology, its economic impact depends on complementary innovations, organizational capabilities, and widespread adoption (Bresnahan & Trajtenberg 1995; Brynjolfsson et al. 2019)
- GPT development is evolutionary and can reach a **tipping point** when accessibility rises and user costs fall
  - Ford Model T (1908; assembly line from 1913): over 40% of U.S. cars by 1917
    $\rightarrow$ roads, petrol stations, trucking, motels (Eli, Hausman & Rhode 2025)
  - The emergence of **open-source AI models** may represent a comparable turning point
- Two business models
  - **Proprietary** (ChatGPT, Claude, Gemini): developers retain control of model weights and monetize via APIs, subscriptions, and enterprise contracts
  - **Open-source** (Llama, Qwen, Mistral, DeepSeek): shared parameter weights let external parties download, modify, fine-tune, and deploy. OS AI modelers capture value through complementary products and services, inducing growth in **inference** and **API services** as well as demand for cloud computing and GPUs (Lerner & Tirole 2005)
- Price of intelligence fell roughly a thousandfold; open-source models cost about 90% less than closed models of similar quality (Demirer et al. 2025; Demirer et al. 2026)

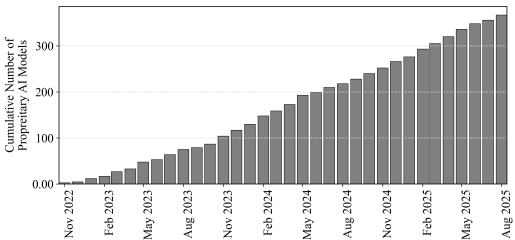
Motivation: Conjectures

- **Conjectures:** open-source AI broadens the economic gains
  - **Access and costs:** lower user costs and greater accessibility let gains reach smaller, more constrained firms, unlike proprietary models
  - **Suitability**: firms with more AI - suitable task benefit more, including resource-constrained firms
  - **Spillovers:** value s pills over along supply chains to low-AI-exposure suppliers and customers, but only after open-source model releases (Cheng & Nault (2007). Garrick & Gertler (2008))
- **But:** "open" does not necessarily mean genuinely open (Widder, Whittaker & West 2024; Azoulay, Krieger & Nagaraj 2024)
  - **Access and costs:** compute, training data, and deployment infrastructure remain barriers for constrained firms
  - **Suitability:** gains depend on firms' deployment capabilities, not on access alone
  - **Spillovers:** control over critical resources may stay concentrated, so gains need not diffuse; released weights also raise safety concerns

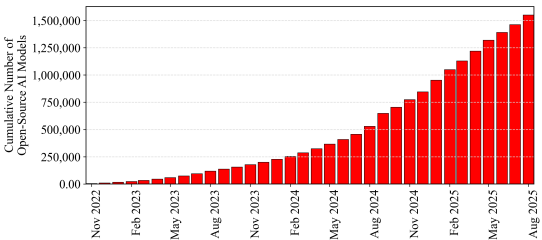
$\Rightarrow$ **This paper:** Examines the **expected** economic impacts using stock market reactions to 67 proprietary and open-source model releases (Nov 2022 – Aug 2025), with early evidence from firms' operating outcomes

Motivation - Number of AI Models: Proprietary vs. Open-Source

Panel A: Proprietary Models

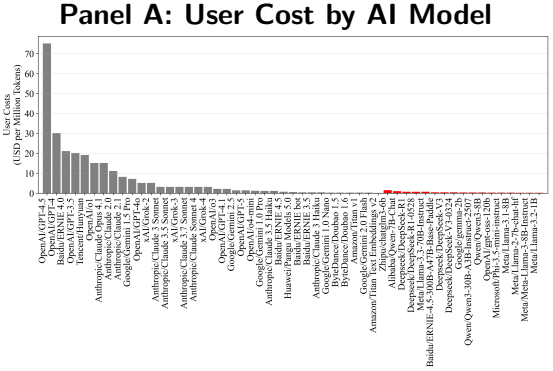


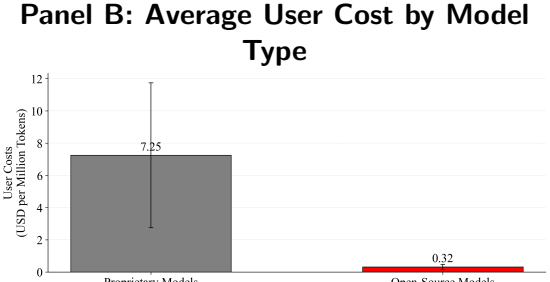
Panel B: Open-Source Models



- Rapid growth of open-source models on Hugging Face vs. slow growth of proprietary models on Epoch AI

Motivation - User Costs of Gen-AI Models: Proprietary vs. Open-Source

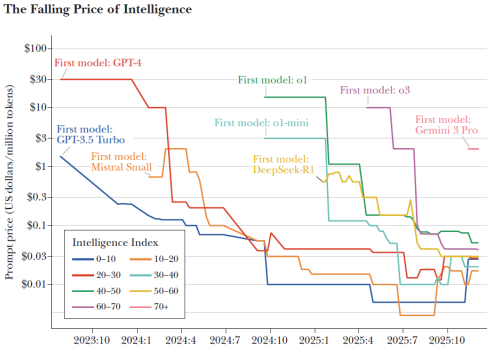




- Much lower user cost of open-source AI models, in terms of USD per million tokens

Still, user costs drop across models





Source: OpenRouter and Artificial Analysis.
Note: This figure shows the minimum prompt price per million tokens for models at each level of intelligence over time. Each line represents a distinct bin of intelligence scores. The y-axis uses a log scale.

Sample: Difference Between Proprietary vs. Open-Source

- **1. Proprietary Models**
  - **Definition:** models for which the code, weights, and training data are not publicly released, and whose use, modification, and redistribution are restricted by the developer
  - **Leading Developers:**
    - *U.S.:* OpenAI, Anthropic, Amazon, xAI, and Google
    - *China:* Baidu, Huawei, Tencent, ByteDance, and Moonshot AI
  - **Method:** Hand-collected initial public release dates of major models
  - **Sample:** 41 proprietary model releases (November 2022 – August 2025)

▸ Model Release Dates

- **2. Open-Source Models**
  - **Definition:** models for which code, weights, and/or training data are publicly shared and that often achieve material improvements in computational efficiency through reducing compute, memory, energy, or latency during training or inference
  - **Platform:** Hugging Face, primary platform for disseminating open-source models
  - **Selection Criteria:**
    - Record cumulative downloads from its release date to August 27, 2025
    - Select from each month the top 5 models by cumulative downloads
    - Restricted to widely recognized commercial developers in U.S. and China to ensure prominence
  - **Sample:** 26 open-source model releases (November 2022 – August 2025)

Research Questions

Intuition: Open-source models foster a competitive market for complementary inference and API services, increasing their availability and quality while reducing the cost of accessing AI capabilities

- What are the nuanced impacts of proprietary vs open-source AI model releases on firm valuations?
- Which types of firms benefit most — and why?
- How do labor-market exposure and supply chain dynamics mediate these effects?

Preview of Findings

- Open-source releases elicit higher & broader stock market reactions than proprietary releases; the latter, on the contrary, are mostly negative with a positive surge on the other end
  - Open-source models provide downside protection for firms with high AI exposure, leading to slightly positive skewness in CARs
- Benefited firms:
  (i) greater financial constraints (ii) higher AI exposure (iii) mostly augmentation-oriented firms
- Positive valuation spillovers extend from high-AI exposure firms to low-AI exposure firms in both upstream and downstream markets

Point: Open-source AI shifts value creation from the foundation model layer, through growth in inferences, to the application layer, where dynamic competition allows smaller firms to profitably exploit newly accessible model capabilities

Contribution

- **AI and firm value:** from frontier model capabilities to the economics of **access and diffusion** (Brynjolfsson et al. 2019; Lerner & Tirole 2005; Demirer et al. 2025)
  - Open-source releases elicit more favorable investor reactions than proprietary releases, particularly among resource-constrained firms
- **AI exposure and technology adoption:** gains depend on how AI **complements** firms' production (Eisfeldt et al. 2023; Felten et al. 2019; Hershbein & Kahn 2018; Nagle 2019; Chen & Wang 2025)
  - Distinguish potential suitability from realized usage, and augmentation from displacement: gains concentrate where AI complements human work
- **Distribution and transmission of gains:** openness eases financial frictions and extends benefits along supply chains (Midrigan & Xu 2014; Cai et al. 2025; Bertomeu et al. 2025; Chu et al. 2019)
  - Constrained firms fare relatively better after open-source releases; spillovers reach low-AI-exposure suppliers and customers only after open-source releases

Data & Empirical Strategy

Data

- Compustat North America & Compustat Global Security Price data
  - Stock price
- Compustat North America & Compustat Global Security Fundamentals data
  - Use of fundamental variables to remove non-stock securities such as ETFs
  - Control variables: log(asset), ROA, debt-to-equity ratio, and Book-to-Market
- Revelio Labs
  - Standard Occupational Classification (SOC) codes for each employee's job position
- OECD Input-Output data
  - Flow of sales between industries
- Factset Revere
  - Flow of sales at the inter-firm level
- 5,656 unique U.S. listed firms

Measurement of AI Exposure

Labor market approach: AI exposure at the firm level based on job postings

- Construction of "AI exposure", following Eisfeldt et al. (2023) and Bertomeu et al. (2025)
  - $AI\ Exposure_{ig}^{raw}$: This represents share of tasks in each occupation that can be made more productive using Gen-AI (potential AI exposure) ▸ Eisfeldt measure details
  - $AI\ Exposure_{ig}^{claude}$: We also use Anthropic Economic Index as an alternative measure; this measures actual use of Gen-AI (realized AI use intensity)
- Human capital measurement: share of employees with majors in (i) humanities (ii) science & engineering (iii) business & economics

Measurement of AI Exposure

- AI exposure along supply chain:

$$AI\ Exposure_j^{upstream} = \sum_{j \neq k} w_{jk}\ AI\ Exposure_k^{raw}$$
$$AI\ Exposure_j^{downstream} = \sum_{j \neq k} w_{jk}\ AI\ Exposure_k^{raw}$$

  - Use the U.S. domestic input-output table from the OECD in 2022 and construct industry weights based on the shares of a focal industry's purchases from upstream suppliers and sales to downstream customer industries.
  - Do this for each focal industry with below-median industry AI exposure, we construct a measure of upstream-industry AI exposure that is based on upstream industries with above-median AI exposure as follows.
  - Weight = fraction of sales (downstream) or purchases (upstream)

Empirical Strategy - Event Studies

- We examine stock market reactions to open-source releases

$$\text{CAR}[-1,+k]_{ig} = \beta_0 + \beta_1 \text{OS}_g + \alpha X_i + \text{FE} + \varepsilon_{ig} \tag{1}$$

- We also examine the heterogeneous effect along different degrees of AI exposure

$$\begin{aligned}\text{CAR}[-1,+k]_{ig} = \beta_0 + \beta_1 \text{OS}_g \times \text{AI Exposure}_{ig}^{raw} \\ + \beta_2 \text{AI Exposure}_{ig}^{raw} + \beta_3 \text{OS}_g + \alpha X_i + \text{FE} + \varepsilon_{ig}\end{aligned} \tag{2}$$

- We also examine the valuation effect along supply chains

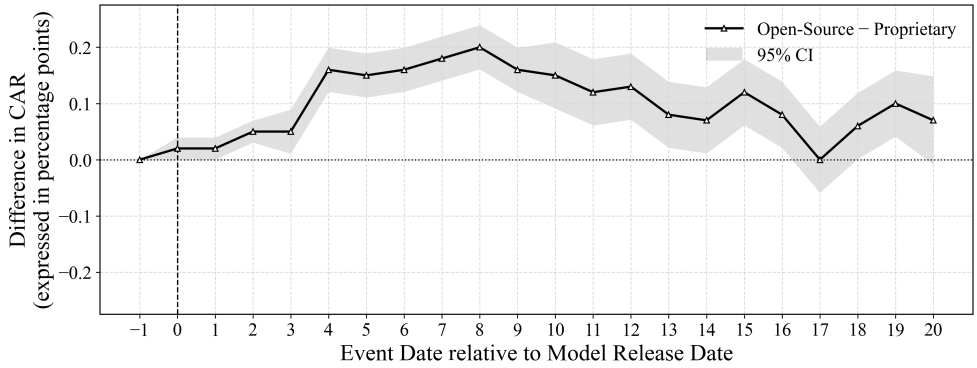
$$\begin{aligned}CAR[-1,+k]_{ijg} = \beta_0 + \beta_1 AI\ Exposure_{jg}^{upstream} + \beta_2 AI\ Exposure_{jg}^{downstream} \\ + \beta_3 AI\ Exposure_{jg}^{ind} + \alpha X_i + FE + \varepsilon_{ijg}\end{aligned} \tag{3}$$

Sample: Focal industries with below-median AI exposure &
Up- and downstream industries with above-median AI exposure

Empirical Results

Investor Expectations of Firm Value: Stock Market Reactions

Market-Implied Valuation Gap: Open-Source vs. Proprietary



- Gap in CAR between open-source releases and proprietary releases is persistently positive

Market Response by Model Type: Open-Source vs. Proprietary

Panel A: Cumulative Abnormal Return

	(1)	$CAR[-1,+5]_{ig}$	(2)
	Proprietary AI Models		Open-Source AI Models
Constant	-0.693***		0.174***
	(0.000)		(0.000)
Controls	N		N
Firm FE	Y		Y
Adj. R^2	0.021		0.010
Obs.	125,505		79,264

Panel B: Cumulative Raw Return

	(1)	$CRR[-1,+5]_{ig}$	(2)
	Proprietary AI Models		Open-Source AI Models
Constant	-0.102***		0.975***
	(0.000)		(0.000)
Controls	N		N
Firm FE	Y		Y
Adj. R^2	0.013		0.008
Obs.	125,505		79,264

Market Response to Open-Source vs. Proprietary AI Models

	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$
OS_g	0.230***	0.349***	0.867***
	(0.0140)	(0.0180)	(0.0310)
Constant	-0.508	-0.984*	1.069
	(0.393)	(0.582)	(1.119)
Controls		log(Assets), D/E, ROA	
Firm FE		Y	
Adj. R^2	0.00300	0.00600	0.0210
Obs.	204,332	204,332	204,332

- U.S. public firms gain firm value, following open-source releases
- On the contrary, effect from proprietary releases is rather muted

Validation Tests for Accessibility and Affordability

	$CAR[-1,+5]_{ig}$	
	(1)	(2)
$Accessibility_g$	0.358***	
	(0.0150)	
$User\ Costs_g$		-0.313***
		(0.0178)
Controls	log(Assets), D/E, ROA	
Firm FE	Y	Y
Adj. R^2	0.0200	0.0204
Obs.	204,332	164,248

Explanatory variables are standardized:

(1) $Accessibility_g$, measured as the standardized average of three binary indicators - whether the model provides open-source weights, provides open-source training code, and has 3 billion or fewer parameters

(2) User $Cost_g$, measured as the standardized dollar cost per million tokens (collected from the WayBack Machine)

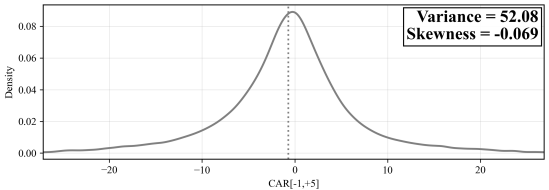
- Accessibility and affordability are key factors of open-source AI models.

Empirical Results

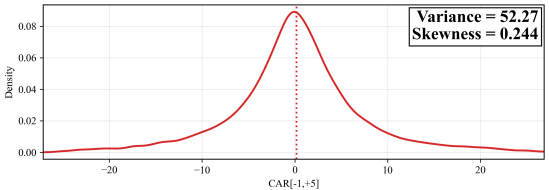
Distributional Effects of Open-Source Models

Heterogeneity in Distribution of CAR[-1,+5] by Model Types: Open-Source vs. Proprietary

Panel A: Proprietary Models

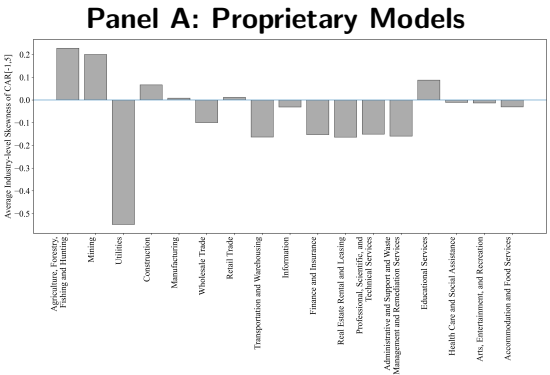


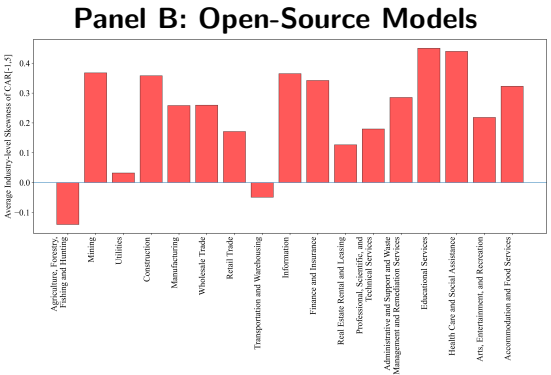
Panel B: Open-Source Models



- Positive skewness of firm-level return distribution following open-source models
  - Proprietary: Share of CARs above 0+1.65σ is 4.42%, and share below 0-1.65σ is 6.11%.
    - ⋆ thinner right, thicker left
  - Open-Source: Share of CARs above 0+1.65σ is 5.45%, and share below 0-1.65σ is 4.90%.
    - ⋆ thicker right, thinner left

Heterogeneity in Skewness by Industry: Proprietary vs. Open-Source





- Positive skewness is persistent at industry-level, following open-source releases

Heterogeneity in Skewness Across Industries

Panel A: The Effects of Open-Source AI Models		
	(1) CAR Skewness$_{jg}$	(2) CRR Skewness$_{jg}$
OS_g	0.156***	0.129***
	(0.031)	(0.031)
Constant	0.010	0.053*
	(0.024)	(0.025)
Adj. R^2	0.006	0.004
Obs.	1,139	1,139

- Open-source releases increase skewness of returns at the industry level

Quantile Regressions of Stock-Market Reactions

Panel A: Heterogeneity Across Gen-AI Exposure (Unconditional)

	(1)	(2)	(3)	(4)	(5)
			$CAR[-1,+5]_{ig}$		
Quantile:	0.05	0.20	0.50	0.80	0.95
OS_g	-0.406	0.289	0.710***	0.480*	0.788
	(0.555)	(0.229)	(0.155)	(0.258)	(0.589)
$OS_g \times AI\ Exposure_{ig}^{raw}$	4.518***	1.320**	-0.381	0.910	1.504
	(1.515)	(0.561)	(0.405)	(0.669)	(1.478)
$AI\ Exposure_{ig}^{raw}$	-2.249**	0.374	0.304	-0.741	1.265
	(1.117)	(0.477)	(0.314)	(0.462)	(1.290)
Constant	-20.194***	-9.531***	-1.117***	7.059***	19.591***
	(0.518)	(0.214)	(0.137)	(0.213)	(0.612)
Controls			log(Assets), D/E, ROA		
Industry FE			Y		
Adj. R^2	0.051	0.065	0.010	0.021	0.028
Obs.	204,332	204,332	204,332	204,332	204,332

- Open-source releases increase skewness of returns for industries with AI exposure

Quantile Regressions of Stock-Market Reactions

Panel B: Heterogeneity Across Gen-AI Exposure (Conditional)

	(1)	(2)	(3)	(4)	(5)
			$CAR[-1,+5]_{ig}$		
Quantile:	0.05	0.20	0.50	0.80	0.95
OS_g	0.172	0.337*	0.460**	0.585**	0.766**
	(0.303)	(0.192)	(0.203)	(0.263)	(0.372)
$OS_g \times AI\ Exposure_{ig}^{raw}$	1.585**	1.282***	1.056*	0.825	0.492
	(0.774)	(0.486)	(0.542)	(0.665)	(0.978)
$AI\ Exposure_{ig}^{raw}$	-0.313	-0.125	0.014	0.157	0.363
	(0.563)	(0.334)	(0.332)	(0.431)	(0.573)
Constant	-19.280***	-8.931***	-1.247***	6.640***	17.998***
	(0.277)	(0.143)	(0.153)	(0.190)	(0.241)
Controls			log(Assets), D/E, ROA		
Industry FE			Y		
R^2	0.134	0.068	0.006	0.021	0.093
Obs.	204,332	204,332	204,332	204,332	204,332

- Open-source releases increase skewness of returns for industries with AI exposure

Heterogeneity Across Firms: Size and Resource Constraints

Panel A: Heterogeneity by Total Assets		
	$CAR[-1,+5]_{ig}$	
	Bottom tercile	Top tercile
	(1)	(2)
OS_g	1.241***	0.467***
	(0.070)	(0.035)
Constant	-2.035**	2.681
	(0.882)	(3.144)
Diff. in Coef. on OS_g: (1) - (2)	0.774***	
	(0.078)	
Coefficient on OS_g + Constant	-0.794	3.148
	(0.884)	(3.146)
Controls	log(Assets), D/E, ROA	
Firm FE	Y	Y
Adj. R^2	0.023	0.015
Obs.	65,442	71,632

- Smaller firms lost following proprietary releases but did not, following open-source releases. Open-source model releases provide broad-based positive reactions.

Heterogeneity Across Firms: Size and Resource Constraints

Panel B: Heterogeneity by Size-Age (SA) Index (Hadlock and Pierce, 2010)

	$CAR[-1,+5]_{ig}$	
	Bottom tercile	Top tercile (higher constraints)
	(1)	(2)
OS_g	0.565***	1.153***
	(0.037)	(0.068)
Constant	7.314***	-2.187**
	(2.706)	(0.990)
Diff. in Coef. on OS_g: (1) - (2)	-0.588***	
	(0.077)	
Coefficient on OS_g + Constant	7.879***	-1.034
	(2.706)	(0.992)
Controls	log(Assets), D/E, ROA	
Firm FE	Y	Y
Adj. R^2	0.015	0.022
Obs.	68,323	68,066

- Firms with higher financial constraints lost following proprietary releases but did not, following open-source releases

Heterogeneity Across Firms: Size and Resource Constraints

Panel C: Heterogeneity by Whited-Wu (2006) Index

	$CAR[-1,+5]_{ig}$	
	Bottom tercile (1)	Top tercile (higher constraints) (2)
OS_g	0.458***	1.135***
	(0.036)	(0.058)
Constant	3.573	-0.707
	(3.931)	(1.074)
Diff. in Coef. on OS_g: (1) - (2)	-0.677***	
	(0.068)	
Coefficient on OS_g + Constant	4.031	0.428
	(3.934)	(1.075)
Controls	log(Assets), D/E, ROA	
Firm FE	Y	Y
Adj. R^2	0.015	0.022
Obs.	60,918	84,974

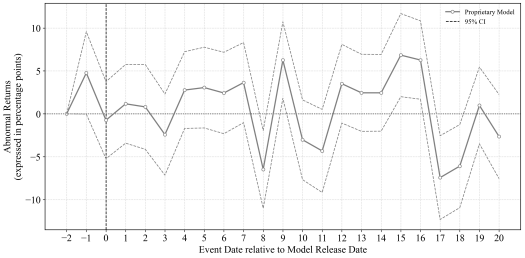
- Firms with higher financial constraints to adopt Gen-AI benefit more, following open-source releases

Empirical Results

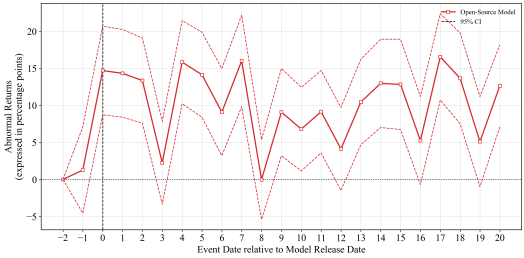
Heterogeneity in Firm-level AI Exposure

Heterogeneity in Abnormal Returns by AI Exposure

Panel A: Proprietary



Panel B: Open-source



- Firms with higher AI exposure do not benefit, following proprietary releases
- Firms with higher AI exposure benefit, following open-source releases

Heterogeneity Across Firms: Gen-AI Exposure

Panel A: Workforce with Potential AI Exposure			
	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$
$OS_g \times AI\ Exposure_{ig}^{raw}$	0.387**	0.822***	1.050**
	(0.190)	(0.266)	(0.458)
$AI\ Exposure_{ig}^{raw}$	-0.170	-0.236	0.018
	(0.146)	(0.201)	(0.396)
OS_g	0.081	0.034	0.463***
	(0.072)	(0.103)	(0.177)
Constant	-0.077	-0.140	-1.045***
	(0.066)	(0.092)	(0.183)
Controls		log(Assets), D/E, ROA	
Industry FE	Y	Y	Y
Adj. R^2	0.00300	0.00600	0.0210
Obs.	204,332	204,332	204,332

- Firms with high AI exposure benefit, following open-source releases

Heterogeneity Across Firms: Gen-AI Exposure

Panel B: Workforce with Realized AI Usage Intensity

	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$
$OS_g \times AI\ Exposure_{ig}^{claude}$	0.447	0.751*	2.084***
	(0.363)	(0.440)	(0.797)
$AI\ Exposure_{ig}^{claude}$	0.199	0.036	0.414
	(0.245)	(0.323)	(0.633)
OS_g	0.116	0.158	0.337*
	(0.091)	(0.112)	(0.201)
Constant	-0.190***	-0.235**	-1.115***
	(0.072)	(0.095)	(0.189)
Controls		log(Assets), D/E, ROA	
Industry FE	Y	Y	Y
Adj. R^2	0.00200	0.00400	0.0110
Obs.	200,788	200,788	200,788

- Firms with high realized AI usage intensity benefit, following open-source releases

Heterogeneity Across Firms: AI Augmentation vs. Displacement

Panel A: Exposure to AI Augmentation			
	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$
$OS_g \times AI\ Augmentation_{ig}$	0.417	0.744**	2.175***
	(0.302)	(0.372)	(0.663)
$AI\ Augmentation_{ig}$	0.133	0.072	0.350
	(0.213)	(0.278)	(0.554)
OS_g	0.097	0.113	0.178
	(0.094)	(0.118)	(0.209)
Constant	-0.182**	-0.251**	-1.127***
	(0.077)	(0.100)	(0.202)
Controls		log(Assets), D/E, ROA	
Industry FE	Y	Y	Y
Adj. R^2	0.002	0.004	0.011
Obs.	200,788	200,788	200,788

- Augmentation-oriented firms benefit, following open-source releases

Heterogeneity Across Firms: AI Augmentation vs. Displacement

Panel B: Exposure to AI Displacement

	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$
$OS_g \times AI\ Displacement_{ig}$	-0.233	-0.345	-0.556
	(0.354)	(0.454)	(0.793)
$AI\ Displacement_{ig}$	0.148	-0.303	0.701
	(0.235)	(0.313)	(0.616)
OS_g	0.284***	0.429***	0.997***
	(0.084)	(0.108)	(0.189)
Constant	-0.169**	-0.145	-1.160***
	(0.066)	(0.089)	(0.175)
Controls		log(Assets), D/E, ROA	
Industry FE	Y	Y	Y
Adj. R^2	0.002	0.004	0.011
Obs.	200,788	200,788	200,788

- Displacement-oriented firms do not benefit, following open-source releases

Heterogeneity Across Firms: AI Augmentation vs. Displacement

Panel C: Exposure to AI Augmentation and Displacement: Small vs. Large

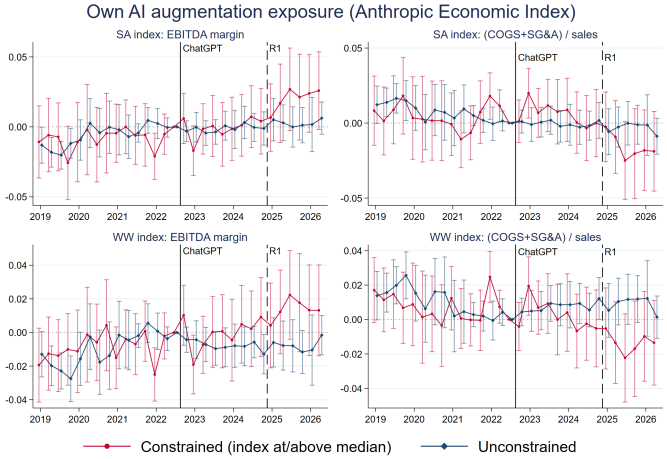
	(1)	(2)	(3)	(4)
	$CAR[-1,+5]_{ig}$			
	Small (Bottom tercile)	Large (Top tercile)	Small (Bottom tercile)	Large (Top tercile)
$OS_g \times AI\ Augmentation_{ig}$	2.783*	-0.124		
	(1.498)	(0.756)		
$AI\ Augmentation_{ig}$	1.095	0.985*		
	(1.255)	(0.547)		
$OS_g \times AI\ Displacement_{ig}$			-1.248	0.407
			(0.987)	(0.527)
$AI\ Displacement_{ig}$			-0.098	0.392
			(0.681)	(0.349)
OS_g	0.340	0.487**	0.685***	0.124
	(0.487)	(0.238)	(0.238)	(0.124)
Constant	-2.704***	-0.337	-0.892***	0.345***
	(0.472)	(0.254)	(0.214)	(0.124)
Controls		log(Assets), D/E, ROA		
Industry FE	Y	Y	Y	Y
Adj. R^2	0.009	0.010	0.003	0.005
Obs.	66,262	68,181	66,262	68,181

- Augmentation-oriented firms with constraints benefit, following open-source releases

Early Observations on Real Impacts: Cost Savings and Augmentation

- Comparison between constrained and unconstrained firms
  - Differential effect of the focal firm's own realized AI intensity

Early Observations on Real Impacts: Cost Savings and Augmentation



$$Y_{i,t} = FirmFE + NAICS2xQuarterFE + \sum_{k \neq 2022Q3} \beta_k \times \mathbf{1}[t = k] \times Usage_i + \varepsilon_{i,t}$$

Early Observations on Real Impacts: Cost Savings and Augmentation

	(1)	(2)	(3)	(4)
Constraint index	SA	SA	WW	WW
Dependent variable	EBITDA margin	Costs / sales	EBITDA margin	Costs / sales
Post-R1 × Usage × Constrained	0.0123*	−0.0121*	0.0149**	−0.0178***
	(0.0067)	(0.0064)	(0.0066)	(0.0066)
Post-R1 × Usage	0.0063*	−0.0057*	0.0009	0.0009
	(0.0032)	(0.0034)	(0.0034)	(0.0035)
Post-R1 × Constrained	−0.0013	−0.0019	0.0096	−0.0098
	(0.0066)	(0.0062)	(0.0065)	(0.0065)
Obs.	27,847	26,625	25,606	24,472
Firms	2,038	1,937	1,859	1,769
Adjusted R^2	0.825	0.821	0.816	0.809
Firm FE	Y	Y	Y	Y
NAICS2 × quarter FE	Y	Y	Y	Y

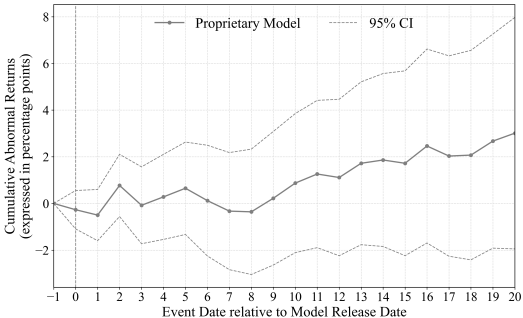
- Gains appear to operate through lower operating costs rather than expansion or new investment, consistent with AI adoption that requires little upfront capital

Empirical Results

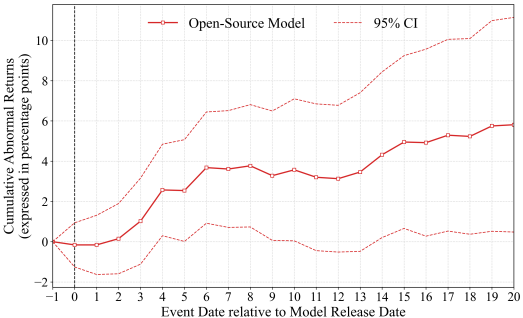
Spillovers Along Supply Chains

High Upstream AI Exposure and Model Releases (Sample = Low AI Exposure Industries)

Panel A: Proprietary



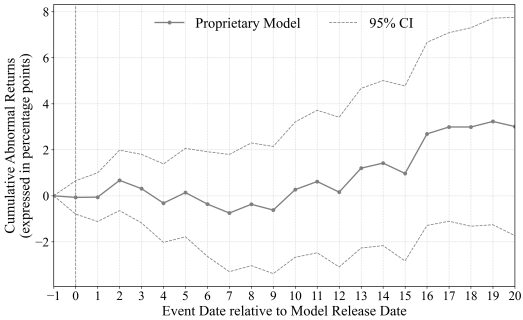
Panel B: Open-source



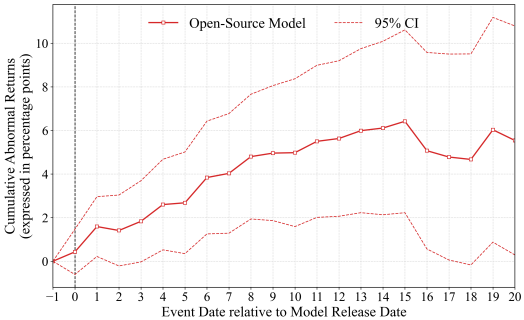
- Firms with higher upstream AI exposure do not benefit, following proprietary releases
- Firms with higher upstream AI exposure benefit, following open-source releases

High Downstream AI Exposure and Model Releases (Sample = Low AI Exposure Industries)

Panel A: Proprietary



Panel B: Open-source



- Firms with higher downstream AI exposure do not benefit, following proprietary releases
- Firms with higher downstream AI exposure benefit, following open-source releases

Spillover Effects of Open-Source Releases (Sample = Low AI Exposure Industries)

	(1)	(2)	(3)	(4)
	Proprietary AI Models		Open-Source AI Models	
	$CAR[-1,+5]_{ig}$			
$AI\ Exposure_{jg}^{upstream}$	0.652		2.586**	
	(1.006)		(1.289)	
$AI\ Exposure_{jg}^{downstream}$		0.138		2.708**
		(0.983)		(1.190)
$AI\ Exposure_{jg}^{ind}$	1.154	-0.660	2.797	1.041
	(4.105)	(3.494)	(4.718)	(3.928)
Constant	-2.810*	-2.081	-1.490	-0.772
	(1.584)	(1.346)	(1.848)	(1.496)
Controls	log(Assets), D/E, ROA			
Industry FE	Y	Y	Y	Y
Adj. R^2	0.018	0.018	0.011	0.011
Obs.	32,173	32,173	20,400	20,400

- Positive spillovers from both upstream and downstream
  - From upstream: Focal firms gain access to more efficient production inputs and face reduced integration frictions
  - From downstream: Focal firms may reduce inventory-holding costs as high-AI-exposure customers drive market expansion and enable superior demand forecasting

Robustness Checks

Issues – Two-way Fixed Effects with 2 Event Groups (n and m)

1. **Contamination of control**
   - $\Rightarrow$ Not an issue for us for each event is new
2. **Model type imbalance and timing bias**
   - $\Rightarrow$ Solved via PSM and Entropy Balancing
3. **Event windows overlapped**
   - $\Rightarrow$ Isolated cleanly by constructing event-specific cohorts (stacks)
4. **Heterogeneity across firms**
   - $\Rightarrow$ Solved via stacked Triple DID (Bilicka (2019) and and McMullin and Schonberger (2020))
5. **Heterogeneity across event types within groups**
   - Can introduce negative weighting biases & artificial pre-event trends
   - $\Rightarrow$ Robustness cleared via counterfactual estimators

Robustness: Propensity Score Matching and Entropy Balancing (Balancing OS with PM)

Panel A: Propensity Score Matching

	(1)	(2)
	$CAR[-1,+5]_{ig}$	
	Proprietary AI Models	Open-Source AI Models
Constant	-0.674***	0.173***
	(0.000)	(0.000)
Controls	N	N
Firm FE	Y	Y
Adj. R^2	0.020	0.010
Obs.	79,227	79,245

Panel B: Entropy Balancing

	(1)	(2)
	$CAR[-1,+5]_{ig}$	
	Proprietary AI Models	Open-Source AI Models
Constant	-0.687***	0.173***
	(0.000)	(0.000)
Controls	N	N
Firm FE	Y	Y
Adj. R^2	0.022	0.010
Obs.	125,485	79,245

- Robust results using PSM and Entropy Balancing

Robustness: Propensity Score Matching and Entropy Balancing (Balancing OS with PM)

- **Objective:** Address the numerical imbalance (26 OS vs. 41 Proprietary) and ensure results are not driven by the macro-timing of announcements.
- **Propensity Score Matching (PSM):** Matches each OS firm-event observation to a Proprietary firm-event observation exhibiting a similar pre-release market trajectory.:
  - **LHS (Treatment):** The open-source model release dummy, $1(OS_g)$
  - **RHS (Covariates):** Pre-event firm-level abnormal returns from $t = -5$ to $t = -2$.
    - A Propensity Score equation
  - Uses 0.01 standard deviation caliper to ensure both groups share the same "market state" before model releases. (use the PS estimates to match firms from the P with those in OS crowds)
- ⋆ **Entropy Balancing:** Reweights the entire proprietary sample so its pre-event return moments (mean, variance) exactly replicate the open-source sample.

Robustness: Stacked DID Specification (Avoiding Overlapping Windows)

Panel A: DID Estimation

	AR_{itg}	
	(1)	(2)
$OS_g \times Post_{tg} \times AI\ Exposure_{ig}$	0.150***	0.119***
	(0.0300)	(0.0340)
$Post_{tg} \times AI\ Exposure_{ig}$	-0.0410**	-0.0720***
	(0.0190)	(0.0210)
Constant	0.222	0.040
	(1.557)	(1.514)
Controls	log(Assets), D/E, ROA	
Event-Firm FE	Y	Y
Event-Date FE	Y	Y
Event-Industry-Date FE	N	Y
Adj. R^2	0.0270	0.0490
Obs.	1,429,842	1,429,842

- Robust results using stacked DID specification

Robustness: Stacked DID Specification with PSM (Avoiding Overlapping Windows & Matching High and Low AI Firms in Pre-event Run-ups)

Panel B: DID Estimation with Propensity Score Matching

	AR_{itg}	
	(1)	(2)
$OS_g \times Post_{tg} \times AI\ Exposure_{ig}$	0.114***	0.0970***
	(0.0310)	(0.0360)
$Post_{tg} \times AI\ Exposure_{ig}$	-0.0180	-0.0500**
	(0.0200)	(0.0230)
Constant	0.610	0.208
	(1.851)	(1.768)
Controls	log(Assets), D/E, ROA	
Event-Firm FE	Y	Y
Event-Date FE	Y	Y
Event-Industry-Date FE	N	Y
Adj. R^2	0.0260	0.0480
Obs.	1,222,844	1,222,830

- Robust results using stacked DID specification with PSM

Robustness: Stacked DID Specification with Entropy Balancing

Panel C: DID Estimation with Entropy Balancing

	AR_{itg}	
	(1)	(2)
$OS_g \times Post_{tg} \times AI\ Exposure_{ig}$	0.150***	0.130***
	(0.031)	(0.036)
$Post_{tg} \times AI\ Exposure_{ig}$	-0.044**	-0.077***
	(0.020)	(0.022)
Constant	0.546	0.352
	(1.576)	(1.551)
Controls	log(Assets), D/E, ROA	
Event-Firm FE	Y	Y
Event-Date FE	Y	Y
Event-Industry-Date FE	N	Y
Adj. R^2	0.027	0.049
Obs.	1,429,602	1,429,602

- Robust results using stacked DID specification with Entropy Balancing

Robustness: Stacked DID Specification

- **Objective:** Address potential selection bias into firm types and satisfy the parallel trends assumption by matching on pre-event trajectories.
- Following Bilicka (2019), we implement **Propensity Score Matching (PSM)** within each event-specific stack:
  - **Step 1:** Estimate a logit model for each event where the dependent variable is the High AI Exposure dummy, $1(AI\ Exposure > \text{median})$.
  - **Step 2:** Use abnormal returns (AR) from trading days $t = -5$ to $t = -2$ as predictors to calculate propensity scores.
  - **Step 3:** Pair each high-exposure firm with a low-exposure firm using one-to-one nearest-neighbor matching (0.01 std. dev. caliper, no replacement).
- **Estimation:** Pool the matched stacks and estimate a **Triple-DID** using **firm-date** observations over the $[-1, +5]$ event window.

⋆ **Entropy Balancing:** As an alternative, we reweight control firms to match the run-up return moments (mean, variance) of the high-exposure group.

Robustness: Counterfactual Estimators

	AR_{itg} − Counterfactual AR_{itg}			
	FEct	iFEct	MC	Synthetic Matching
	(1)	(2)	(3)	(4)
OS_g	0.0900***	0.0560***	0.0570***	0.0230***
	(0.0100)	(0.00900)	(0.00900)	(0.00700)
Controls	log(Assets), D/E, ROA			
Firm FE	Y	Y	Y	Y
Date FE	Y	Y	Y	Y
Adj. R^2	0.00300	0.00200	0.00200	0.00300
Obs.	672,223	783,973	783,973	656,561

- Robust results using counterfactual estimators

Robustness: Counterfactual Estimators

- Following Liu et al. (2024b), we complement our baseline estimates with three counterfactual estimators accommodating richer unobserved time-varying heterogeneity:
  (i) the fixed-effects counterfactual estimator (FEct)
  (ii) the interactive fixed-effects counterfactual estimator (iFEct)
  (iii) the matrix completion estimator (MC)
- Specifically, for each event, we define treated firms as those with *AIExposure* above the sample median for the corresponding year,
- For each treated observation, we predict "counterfactual" outcome* using
- and
- This procedure yields a panel of individual treatment effects for all firms in the treatment groups over the release periods.
- Using this panel, we regress the individual treatment effects on a single dummy variable, OS_g, and report the estimation results in Table XIV
- Additionally, we implement a synthetic matching methodology, following Goldsmith-Pinkham and Lyu (2025) which improves upon standard propensity score matching by constructing control firms with pre-event abnormal return trajectories that closely resemble those of treated firms

Robustness: Alternative Measures

- AI Exposure measures
  - Constructed from 10-K filings
  - Constructed from earnings call transcripts
- Return measures
  - Cumulative Raw Returns
  - Cumulative Abnormal Returns adjusted by Fama-French 3 Factors

Robustness: Alternative Measures of AI Exposure Based on 10-K Filings

Panel A: AI Exposure based on 10-K Filings

	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$	(4) $CAR[-1,+20]_{ig}$
$OS_g \times AI\ Exposure_{ig}^{10k}$	0.0780**	0.0550	0.192**	0.443***
	(0.0360)	(0.0460)	(0.0790)	(0.118)
OS_g	0.215***	0.338***	0.831***	0.639***
	(0.0150)	(0.0190)	(0.0350)	(0.0510)
Constant	-0.510	-0.985*	1.065	5.892**
	(0.393)	(0.582)	(1.119)	(2.897)
Controls		log(Assets), D/E, ROA		
Firm FE	Y	Y	Y	Y
Adj. R^2	0.00300	0.00600	0.0210	0.0610
Obs.	204,332	204,332	204,332	204,332

- Robust results using alternative measurements of AI exposure constructed from 10-K filings

Robustness: Alternative Measures of AI Exposure Based on Earnings Calls

Panel B: AI Exposure based on Earnings Calls

	(1) $CAR[-1,+0]_{ig}$	(2) $CAR[-1,+1]_{ig}$	(3) $CAR[-1,+5]_{ig}$	(4) $CAR[-1,+20]_{ig}$
$OS_g \times AI\ Exposure^{ec}_{ig}$	0.0590*	0.0170	-0.0530	0.299***
	(0.0350)	(0.0430)	(0.0790)	(0.115)
OS_g	0.220***	0.346***	0.876***	0.675***
	(0.0150)	(0.0200)	(0.0340)	(0.0510)
Constant	-0.508	-0.984*	1.069	5.902**
	(0.393)	(0.582)	(1.119)	(2.897)
Controls		log(Assets), D/E, ROA		
Firm FE	Y	Y	Y	Y
Adj. R^2	0.00300	0.00600	0.0210	0.0610
Obs.	204,332	204,332	204,332	204,332

- Robust results using alternative measurements of AI exposure constructed from earnings call transcripts

Robustness: Alternative Outcome Variables

Panel A: Cumulative Raw Returns				
	(1) $CRR[-1,+0]_{ig}$	(2) $CRR[-1,+1]_{ig}$	(3) $CRR[-1,+5]_{ig}$	(4) $CRR[-1,+20]_{ig}$
OS_g	0.0820***	0.411***	1.08***	1.26***
	(0.0140)	(0.0180)	(0.0300)	(0.0470)
Constant	-0.266	-1.089*	1.075	6.070**
	(0.434)	(0.636)	(1.221)	(3.033)
Controls		log(Assets), D/E, ROA		
Firm FE	Y	Y	Y	Y
Adj. R^2	0.00100	0.00500	0.0190	0.0530
Obs.	204,332	204,332	204,332	204,332

- Robust results using cumulative raw returns as outcome variable

Robustness: Alternative Outcome Variables

Panel B: Cumulative Abnormal Returns Adjusted by Fama-French 3 Factors

	(1) $CAR[-1,+0]_{ig}^{FF3}$	(2) $CAR[-1,+1]_{ig}^{FF3}$	(3) $CAR[-1,+5]_{ig}^{FF3}$	(4) $CAR[-1,+20]_{ig}^{FF3}$
OS_g	-0.00500	0.00200	0.0800***	0.453***
	(0.0130)	(0.0170)	(0.0290)	(0.0440)
Constant	-0.715*	-0.940*	1.015	5.497*
	(0.386)	(0.555)	(1.112)	(2.812)
Controls		log(Assets), D/E, ROA		
Firm FE	Y	Y	Y	Y
Adj. R^2	0.00200	0.00500	0.0190	0.0610
Obs.	204,332	204,332	204,332	204,332

- Robust results using cumulative abnormal returns adjusted by Fama-French 3 Factors as outcome variable

Conclusion

- Open-source releases elicit higher & broader stock market reactions than proprietary releases; the latter, on the contrary, are mostly negative with a positive surge on the other end
  - Open-source models provide downside protection for firms with high AI exposure, leading to slightly positive skewness in CARs
- Benefited firms:
  (i) greater financial constraints (ii) higher AI exposure (iii) mostly augmentation-oriented firms
- Positive valuation spillovers extend from high-AI exposure firms to low-AI exposure firms in both upstream and downstream markets

Unresolved:

- Conjecture: Open-source AI shifts value creation from the foundation-model layer to the application layer, where dynamic competition allows smaller firms to exploit newly accessible model capabilities
- Governance and firm level changes

Appendix

Gen-AI Model Release Dates

Event ID	Date	Company and Model Name	Open-Source
1	2022-11-30	OpenAI/GPT-3.5	No
2	2023-01-26	Facebook/opt-iml-max-1.3b	Yes
3	2023-03-14	OpenAI/GPT-4	No
4	2023-03-16	Baidu/ERNIE bot	No
5	2023-06-24	Zhipu/chatglm2-6b	Yes
6	2023-06-27	Baidu/ERNIE 3.5	No
7	2023-07-07	Huawei/Pangu Models 3.0	No
8	2023-07-11	Anthropic/Claude 2.0	No
9	2023-07-13	Meta/Llama-2-7b-chat-hf	Yes
10	2023-08-03	Alibaba/Qwen-7B-Chat	Yes
11	2023-09-06	Tencent/Hunyuan	No
12	2023-09-28	Amazon/Titan v1	No
13	2023-10-09	Moonshot AI/Kimi	No
14	2023-10-17	Baidu/ERNIE 4.0	No
15	2023-10-25	Zhipu/chatglm3-6b	Yes
16	2023-11-04	xAI/Grok-1	No
17	2023-11-14	Intel/neural-chat-7b-v3-1	Yes
18	2023-11-21	Anthropic/Claude 2.1	No
19	2023-11-29	Bigcode/Starcoder2-3b	Yes
20	2023-12-06	Google/Gemini 1.0 Nano	No

Gen-AI Model Release Dates

Event ID	Date	Company and Model Name	Open-Source
21	2023-12-13	Google/Gemini 1.0 Pro	No
22	2024-01-22	Alibaba/Qwen1.5-0.5B	Yes
23	2024-02-08	Google/gemma-2b	Yes
24	2024-03-04	Anthropic/Claude 3 Sonnet	No
25	2024-03-14	Anthropic/Claude 3 Haiku	No
26	2024-04-09	Google/Gemini 1.5 Pro	No
27	2024-04-12	Apple/OpenELM-1_1B-Instruct	Yes
28	2024-04-17	Meta/Meta-Llama-3-8B-Instruct	Yes
29	2024-04-30	Amazon/Titan Text Embeddings v2	No
30	2024-05-13	OpenAI/GPT-4o	No
31	2024-05-15	xAI/Grok-1.5	No
32	2024-05-31	Alibaba/Qwen2-0.5B	Yes
33	2024-06-20	Anthropic/Claude 3.5 Sonnet	No
34	2024-06-21	Huawei/Pangu Models 5.0	No
35	2024-07-14	Meta/Llama-3.1-8B	Yes
36	2024-08-06	Amazon/Titan Image Generator G1 v2	No
37	2024-08-13	xAI/Grok-2	No
38	2024-08-16	Microsoft/Phi-3.5-mini-instruct	Yes
39	2024-09-12	OpenAI/o1	No
40	2024-09-18	Meta/Llama-3.2-1B	Yes

Gen-AI Model Release Dates

Event ID	Date	Company and Model Name	Open-Source
41	2024-10-22	Anthropic/Claude 3.5 Haiku	No
42	2024-11-26	Meta/Llama-3.3-70B-Instruct	Yes
43	2024-12-11	Google/Gemini 2.0 Flash	No
44	2024-12-25	Deepseek/DeepSeek-V3	Yes
45	2025-01-20	Deepseek/DeepSeek-R1	Yes
46	2025-01-22	ByteDance/Doubao 1.5	No
47	2025-01-31	OpenAI/o3	No
48	2025-02-17	xAI/Grok-3	No
49	2025-02-24	Anthropic/Claude 3.7 Sonnet	No
50	2025-02-27	OpenAI/GPT-4.5	No
51	2025-03-10	Google/gemma-3-1b-it	Yes
52	2025-03-16	Baidu/ERNIE 4.5	No
53	2025-03-24	Deepseek/DeepSeek-V3-0324	Yes
54	2025-03-25	Google/Gemini 2.5	No
55	2025-04-14	OpenAI/GPT-4.1	No
56	2025-04-16	OpenAI/o4-mini	No
57	2025-04-27	Qwen/Qwen3-8B	Yes
58	2025-05-22	Anthropic/Claude Sonnet 4	No
59	2025-05-28	Deepseek/DeepSeek-R1-0528	Yes
60	2025-06-11	ByteDance/Doubao 1.6	No

Gen-AI Model Release Dates

Event ID	Date	Company and Model Name	Open-Source
61	2025-06-19	Google/t5gemma-b-b-prefixlm	Yes
62	2025-06-28	Baidu/ERNIE-4.5-300B-A47B-Base-Paddle	Yes
63	2025-07-09	xAI/Grok-4	No
64	2025-07-28	Qwen/Qwen3-30B-A3B-Instruct-2507	Yes
65	2025-08-04	OpenAI/gpt-oss-120b	Yes
66	2025-08-05	Anthropic/Claude Opus 4.1	No
67	2025-08-07	OpenAI/GPT-5	No

back

Labor Market Approach - Details

- Gen-AI exposure is measured at the firm level by aggregating firm-occupational employment data from Revelio Labs, which provides the number of employees in each job position as of December 2024.
- Each job position's Gen-AI exposure is sourced from (Eisfeldt et al., 2023), quantifying the share of tasks within each position that can benefit from Gen-AI to boost productivity; firm-level exposure is calculated as a weighted average of job role exposures using employee counts as weights.

Labor Market Approach - Details

Telemarketers

41-9041.00

Updated 2025

Solicit donations or orders for goods or services over the telephone.

Sample of reported job titles: Call Agent, Inside Sales Representative (Inside Sales Rep), Telemarketer, Telemarketing Sales Representative (Telemarketing Sales Rep), Telephone Sales Representative (TSR), Telephone Service Representative (TSR), Telesales Representative (Telesales Rep), Telesales Specialist

Summary | Details | Custom | Easy Read | Veterans | Español

Contents

Occupation-Specific Information

Tasks

All 12 displayed

- Contact businesses or private individuals by telephone to solicit sales for goods or services, or to request donations for charitable causes.
- Obtain customer information such as name, address, and payment method, and enter orders into computers.
- Explain products or services and prices, and answer questions from customers.
- Record names, addresses, purchases, and reactions of prospects contacted.
- Maintain records of contacts, accounts, and orders.
- Answer telephone calls from potential customers who have been solicited through advertisements.
- Deliver prepared sales talks, reading from scripts that describe products or services, to persuade potential customers to purchase a product or service or to make a donation.
- Telephone or write letters to respond to correspondence from customers or to follow up initial sales contacts.
- Adjust sales scripts to better target the needs and interests of specific individuals.
- Obtain names and telephone numbers of potential customers from sources such as telephone directories, magazine reply cards, and lists purchased from other organizations.
- Schedule appointments for sales representatives to meet with prospective customers or for customers to attend sales presentations.
- Conduct client or market surveys to obtain information about potential customers.

Find occupations related to multiple tasks

Labor Market Approach - Details

- We follow (Eisfeldt et al., 2023)

SOC Code	Occupation Title	Exposure Score
39-5093	Shampooers	0
51-6041	Shoe and leather workers and repairers	0
51-6042	Shoe machine operators and tenders	0
51-3023	Slaughterers and meat packers	0
47-2022	Stonemasons	0
47-2221	Structural iron and steel workers	0
51-2041	Structural metal fabricators and fitters	0
29-9093	Surgical assistants	0
51-6052	Tailors, dressmakers, and custom sewers	0
47-2082	Tapers	0
49-9052	Telecommunications line installers and repairers	0
47-2053	Terrazzo workers and finishers	0
51-6064	Textile winding, twisting, and drawing out machine setters, operators, and tenders	0
47-2044	Tile and stone setters	0
51-9197	Tire builders	0
49-3093	Tire repairers and changers	0
51-4194	Tool grinders, filers, and sharpeners	0
39-3031	Ushers, lobby attendants, and ticket takers	0
49-9064	Watch and clock repairers	0
53-7073	Wellhead pumpers	0

Labor Market Approach - Details

- We follow (Eisfeldt et al., 2023)

SOC Code	Occupation Title	Exposure Score
41-9041	Telemarketers	.96
43-9081	Proofreaders and copy markers	.95
43-3031	Bookkeeping, accounting, and auditing clerks	.87
15-2021	Mathematicians	.86
15-1251	Computer programmers	.85
43-9022	Word processors and typists	.85
43-3011	Bill and account collectors	.83
27-3091	Interpreters and translators	.82
43-9111	Statistical assistants	.82
15-1254	Web developers	.81
43-6011	Executive secretaries and executive administrative assistants	.77
43-3051	Payroll and timekeeping clerks	.77
43-6014	Secretaries and administrative assistants, except legal, medical, and executive	.77
43-5061	Production, planning, and expediting clerks	.76
15-1212	Information security analysts	.75
43-6013	Medical secretaries and administrative assistants	.75
27-3043	Writers and authors	.75
43-4021	Correspondence clerks	.74
43-9061	Office clerks, general	.74
41-3091	Sales representatives of services, except advertising, insurance, financial services, and travel	.73

back

References I

Bertomeu, Jeremy, Yupeng Lin, Yibin Liu, and Zhenghui Ni, 2025, The impact of generative AI on information processing: Evidence from the ban of ChatGPT in Italy, *Journal of Accounting and Economics* 1, 101782.

Bilicka, K. A., 2019, Comparing UK tax returns of foreign multinationals to matched domestic firms., *American Economic Review* 109, 2921–2953.

Eisfeldt, Andrea L., Gregor Schubert, Bledi Taska, and Miao Ben Zhang, 2023, Generative AI and firm values, NBER Working Paper.

Goldsmith-Pinkham, Paul, and Tianshu Lyu, 2025, Causal inference in financial event studies, Working Paper.

McMullin, Jeff L., and Bryce Schonberger, 2020, Entropy-balanced accruals, *Review of Accounting Studies* 25, 84–119.