AI and Scale: A Quantitative Task-Based Theory of Automation

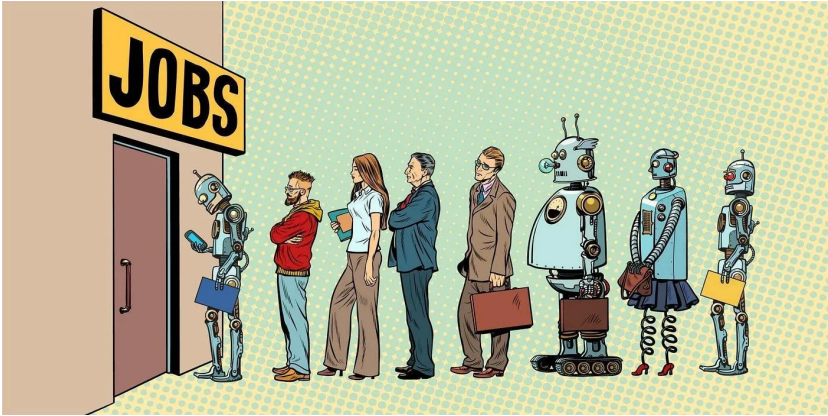
Danial Lashkari[*+], Wensu Li[+], Christina Qiu[†+], Neil Thompson[+]

[*]NY Fed, [+] MIT Future Tech, [†]Yale

October 10, 2025

The views expressed here are my own and do not represent those of the Federal Reserve System.

The Automation Potential of AI?



AI Automation and Scale?

Task-based Theories of Automation focus on relative marginal costs of production

Zeira (1998); Autor et al. (2003); Acemoglu & Autor (2011); Acemoglu & Restropo (2018, 2020, 2022)

AI Automation and Scale?

Task-based Theories of Automation focus on relative marginal costs of production

Zeira (1998); Autor et al. (2003); Acemoglu & Autor (2011); Acemoglu & Restropo (2018, 2020, 2022)

Recent progress in AI driven by scaling training (fixed) costs:

	Model Size	Training Data	Training Compute
AlexNet (2012)	60M pars.	1.2M images	$\sim 10^{18}$ FLOPs
ViT-22B (2022)	22B pars.	300M images	$\sim 10^{24}$ FLOPs

AI Automation and Scale?

Task-based Theories of Automation focus on relative marginal costs of production

Zeira (1998); Autor et al. (2003); Acemoglu & Autor (2011); Acemoglu & Restropo (2018, 2020, 2022)

Recent progress in AI driven by scaling training (fixed) costs:

	Model Size	Training Data	Training Compute
AlexNet (2012)	60M pars.	1.2M images	$\sim 10^{18}$ FLOPs
ViT-22B (2022)	22B pars.	300M images	$\sim 10^{24}$ FLOPs

Businesses incur similar "last mile" fine-tuning costs to specialize AI for production tasks

Fleming et al. (2024); Svanberg et al. (2024)

AI Automation and Scale?

Task-based Theories of Automation focus on relative marginal costs of production

Zeira (1998); Autor et al. (2003); Acemoglu & Autor (2011); Acemoglu & Restropo (2018, 2020, 2022)

Recent progress in AI driven by scaling training (fixed) costs:

	Model Size	Training Data	Training Compute
AlexNet (2012)	60M pars.	1.2M images	$\sim 10^{18}$ FLOPs
ViT-22B (2022)	22B pars.	300M images	$\sim 10^{24}$ FLOPs

Businesses incur similar "last mile" fine-tuning costs to specialize AI for production tasks

Fleming et al. (2024); Svanberg et al. (2024)

Economics of AI automation to account for task-level fine-tuning costs and relevance of scale?

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

- AI Scaling Laws: CV experiment $\Rightarrow$ AI inference/training cost structure

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

- AI Scaling Laws: CV experiment $\Rightarrow$ AI inference/training cost structure
- Data: survey data $\Rightarrow$ task-level required performance + LLM data $\Rightarrow$ labor costs

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

- AI Scaling Laws: CV experiment $\Rightarrow$ AI inference/training cost structure
- Data: survey data $\Rightarrow$ task-level required performance + LLM data $\Rightarrow$ labor costs

Results: predictions for future path of CV AI adoption

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

- AI Scaling Laws: CV experiment $\Rightarrow$ AI inference/training cost structure
- Data: survey data $\Rightarrow$ task-level required performance + LLM data $\Rightarrow$ labor costs

Results: predictions for future path of CV AI adoption

- S-curve of adoption of CV over time with declining computing prices

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

- AI Scaling Laws: CV experiment $\Rightarrow$ AI inference/training cost structure
- Data: survey data $\Rightarrow$ task-level required performance + LLM data $\Rightarrow$ labor costs

Results: predictions for future path of CV AI adoption

- S-curve of adoption of CV over time with declining computing prices
- Scale advantage explains lion's share of cross-task patterns of comparative advantage

Quantitative Task-Based Model of Automation

Model: het. firms in different industries produce tasks with machines (AI) or labor

- AI Marginal Cost Advantage: machine marginal (inference) cost relative to labor marginal costs
- AI Scale Advantage: function of machine fixed (training) costs relative to labor total costs

Application: automation potential of Computer Vision (CV) AI?

- AI Scaling Laws: CV experiment $\Rightarrow$ AI inference/training cost structure
- Data: survey data $\Rightarrow$ task-level required performance + LLM data $\Rightarrow$ labor costs

Results: predictions for future path of CV AI adoption

- S-curve of adoption of CV over time with declining computing prices
- Scale advantage explains lion's share of cross-task patterns of comparative advantage
- CV adoption predicted to raise labor productivity by 1.5% over 10-50 years

Roadmap

Theory

Application to AI Scaling Laws

Quantitative Results

Conclusion

Firms and Production

Production Function:

- Firm needs to produce unit continuum varieties ν of task types $\tau \in \mathcal{T}$:

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} \, d\nu \right)^{\frac{\lambda}{\lambda-1}}$$

Firms and Production

Production Function:

- Firm needs to produce unit continuum varieties ν of task types $\tau \in \mathcal{T}$:

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} \, d\nu \right)^{\frac{\lambda}{\lambda-1}} \qquad \lambda, \xi_\tau \in [0,1], \ \sum_\tau \xi_\tau = 1$$

Firms and Production

Production Function:

- Firm needs to produce unit continuum varieties ν of task types $\tau \in \mathcal{T}$:

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} \, d\nu \right)^{\frac{\lambda}{\lambda-1}} \qquad \lambda, \xi_\tau \in [0,1], \ \sum_\tau \xi_\tau = 1$$

To produce task output $y_\tau(\nu)$ firm can deploy:

Firms and Production

Production Function:

- Firm needs to produce unit continuum varieties ν of task types $\tau \in \mathcal{T}$:

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} \, d\nu \right)^{\frac{\lambda}{\lambda-1}} \qquad \lambda, \xi_\tau \in [0,1], \ \sum_\tau \xi_\tau = 1$$

To produce task output $y_\tau(\nu)$ firm can deploy:

- Workers: at wage rate w with productivity $z_\tau(\nu)$ across task-τ varieties distributed as

$$\mathbb{P}(z_\tau(\nu) \leq z) = G\left(\frac{z}{Z_\tau}\right)$$

Firms and Production

Production Function:

- Firm needs to produce unit continuum varieties ν of task types $\tau \in \mathcal{T}$:

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} \, d\nu \right)^{\frac{\lambda}{\lambda-1}} \qquad \lambda, \xi_\tau \in [0,1], \ \sum_\tau \xi_\tau = 1$$

To produce task output $y_\tau(\nu)$ firm can deploy:

- Workers: at wage rate w with productivity $z_\tau(\nu)$ across task-τ varieties distributed as

$$\mathbb{P}(z_\tau(\nu) \leq z) = G\left(\frac{z}{Z_\tau}\right) = \exp\left[-\left(\frac{z}{Z_\tau}\right)^{-\theta}\right]$$

Firms and Production

Production Function:

- Firm needs to produce unit continuum varieties ν of task types $\tau \in \mathcal{T}$:

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} d\nu \right)^{\frac{\lambda}{\lambda-1}} \qquad \lambda, \xi_\tau \in [0,1], \ \sum_\tau \xi_\tau = 1$$

To produce task output $y_\tau(\nu)$ firm can deploy:

- Workers: at wage rate w with productivity $z_\tau(\nu)$ across task-τ varieties distributed as

$$\mathbb{P}(z_\tau(\nu) \leq z) = G\left(\frac{z}{Z_\tau}\right) = \exp\left[-\left(\frac{z}{Z_\tau}\right)^{-\theta}\right]$$

- Machines: at price r with marginal productivity $z_{K,\tau}$ and fixed productivity $z_{K,\tau}^f$ across task-τ varieties

Cost Minimization Problem

Firm problem given scale y and factor prices (w, r):

- Choose variety output $\boldsymbol{y} \equiv \left([y_\tau(\nu)]_0^1\right)_{\tau \in \mathcal{T}}$ and input choice $\boldsymbol{x} \equiv \left([x_\tau(\nu)]_0^1\right)_{\tau \in \mathcal{T}}$ with

$$y = a \left(\sum_{\tau \in \mathcal{T}} \xi_\tau^{\frac{1}{\lambda}} \int_0^1 y_\tau(\nu)^{\frac{\lambda-1}{\lambda}} \, d\nu \right)^{\frac{\lambda}{\lambda-1}}$$

such that

$$C(w, r; y) = \min_{\boldsymbol{y}, \boldsymbol{x}} \sum_{\tau \in \mathcal{T}} \int_0^1 \left(c_\tau(\nu) \, y_\tau(\nu) + c_\tau^f(\nu) \right) d\nu$$

and that marginal and fixed costs of production $c_\tau(\nu)$ and $c_\tau^f(\nu)$ satisfy

$$c_\tau(\nu) = \begin{cases} \frac{w}{z_\tau(\nu)}, & \text{if } x_\tau(\nu) = \ell_\tau(\nu) \\ \frac{r}{z_{K,\tau}}, & \text{if } x_\tau(\nu) = k_\tau(\nu) \end{cases} \qquad c_\tau^f(\nu) = \begin{cases} 0, & \text{if } x_\tau(\nu) = \ell_\tau(\nu) \\ \frac{r}{z_{K,\tau}^f}, & \text{if } x_\tau(\nu) = k_\tau(\nu) \end{cases}$$

Automation Decisions: Variety Level

Task type τ: given firm-level marginal costs c variety-level relative labor productivity $z_\tau(\nu)/Z_\tau$

$$x_\tau(\nu) = \begin{cases} \ell_\tau(\nu), & \text{if } \widehat{z}_\tau \leq \frac{z_\tau(\nu)}{Z_\tau} \\ k_\tau(\nu), & \text{if } \frac{z_\tau(\nu)}{Z_\tau} \leq \widehat{z}_\tau \end{cases}$$

where machine comparative advantage $\widehat{z}_\tau$ defined by

$$\widehat{z}_\tau = \left[\left(\frac{z_{K,\tau}/Z_\tau}{r/w} \right)^{\lambda-1} + SA_\tau(w, r; y)^{\lambda-1} \right]^{\frac{1}{\lambda-1}}$$

and where scale advantage $SA_{j\tau}$ depends on the ratio of (variable) labor costs to fixed costs

$$SA_{j\tau}(w, r; y) \equiv \left(\frac{1}{1-\lambda} \frac{z^f_{K,\tau}/Z_\tau \times y_{L,\tau}}{r/w} \right)^{\frac{1}{1-\lambda}} \qquad y_{L,\tau} \equiv \xi_\tau \left(\frac{w/Z_\tau}{c} \right)^{-\lambda} \frac{y}{a}$$

Automation Decisions: Comparative and Scale Advantages

Machines with No Fixed Costs: automation determined by relative marginal costs

$$\lim_{z^f_{K,\tau} \to \infty} \widehat{z}_\tau = \frac{z_{K,\tau}/Z_\tau}{r/w}$$

Automation Decisions: Comparative and Scale Advantages

Machines with No Fixed Costs: automation determined by relative marginal costs

$$\lim_{z^f_{K,\tau} \to \infty} \widehat{z}_\tau = \frac{z_{K,\tau}/Z_\tau}{r/w}$$

Machines with No Marginal Costs: automation determined by scale advantage

$$\lim_{z_{K,\tau} \to \infty} \widehat{z}_\tau = SA_\tau = \left(\frac{1}{1-\lambda} \frac{z^f_{K,\tau}/Z_\tau \times y_{L,\tau}}{r/w} \right)^{\frac{1}{1-\lambda}}$$

Automation Decisions: Comparative and Scale Advantages

Machines with No Fixed Costs: automation determined by relative marginal costs

$$\lim_{z^f_{K,\tau}\to\infty} \widehat{z}_\tau = \frac{z_{K,\tau}/Z_\tau}{r/w}$$

Machines with No Marginal Costs: automation determined by scale advantage

$$\lim_{z_{K,\tau}\to\infty} \widehat{z}_\tau = SA_\tau = \left(\frac{1}{1-\lambda}\frac{z^f_{K,\tau}/Z_\tau \times y_{L,\tau}}{r/w}\right)^{\frac{1}{1-\lambda}}$$

Automation share: task-τ's share of automated varieties for firm j

$$\varphi_{j\tau} \equiv G\left(\widehat{z}_\tau\right)$$

Automation Decisions: Comparative and Scale Advantages

Machines with No Fixed Costs: automation determined by relative marginal costs

$$\lim_{z^f_{K,\tau}\to\infty} \widehat{z}_\tau = \frac{z_{K,\tau}/Z_\tau}{r/w}$$

Machines with No Marginal Costs: automation determined by scale advantage

$$\lim_{z_{K,\tau}\to\infty} \widehat{z}_\tau = SA_\tau = \left(\frac{1}{1-\lambda}\frac{z^f_{K,\tau}/Z_\tau \times y_{L,\tau}}{r/w}\right)^{\frac{1}{1-\lambda}}$$

Automation share: task-τ's share of automated varieties for firm j

$$\varphi_{j\tau} \equiv G\left(\widehat{z}_\tau\right) = \exp\left[-\widehat{z}_{j\tau}{}^{-\theta}\right]$$

Automation Decisions: Task Level

Task-τ marginal cost for the firm: given by c_τ / a where

$$c_\tau^{1-\lambda} = \overbrace{\underbrace{\left(\frac{w}{Z_\tau}\right)^{1-\lambda}}_{\text{intensive}} \underbrace{G_{\lambda-1}\left(\widehat{z}_\tau\right)}_{\text{extensive}}}^{\text{labor}} + \overbrace{\underbrace{\left(\frac{r}{z_{K,\tau}}\right)^{1-\lambda}}_{\text{intensive}} \underbrace{G\left(\widehat{z}_\tau\right)}_{\varphi_\tau \equiv \text{extensive}}}^{\text{machines}}$$

with $G_{\lambda-1}(x) \equiv \int_x^\infty u^{\lambda-1}\, dG(u)$

Automation Decisions: Task Level

Task-τ marginal cost for the firm: given by c_τ / a where

$$c_\tau^{1-\lambda} = \overbrace{\underbrace{\left(\frac{w}{Z_\tau}\right)^{1-\lambda}}_{\text{intensive}} \underbrace{G_{\lambda-1}\left(\widehat{z}_\tau\right)}_{\text{extensive}}}^{\text{labor}} + \overbrace{\underbrace{\left(\frac{r}{z_{K,\tau}}\right)^{1-\lambda}}_{\text{intensive}} \underbrace{G\left(\widehat{z}_\tau\right)}_{\varphi_\tau \equiv \text{extensive}}}^{\text{machines}}$$

with $G_{\lambda-1}(x) \equiv \int_x^\infty u^{\lambda-1}\, dG(u)$

Firm-j machine and labor shares in variable costs + fixed/variable ratio:

$$\omega_{K,\tau} \equiv \frac{r\, k_{j\tau}}{c_\tau y_\tau} = \left(\frac{r / z_{K,\tau}}{c_\tau}\right)^{1-\lambda} G\left(\widehat{z}_\tau\right)$$

$$\omega_\tau^f \equiv \frac{c_\tau^f}{c_\tau y_\tau} = \frac{r / z_{K,\tau}^f}{c_\tau y_\tau}\, G\left(\widehat{z}_\tau\right)$$

Graphical Depiction

Automation Decisions: Firm Level

Firm-level costs: characterized by fixed point problem

$$C(w, r; y) = c\frac{y}{a} + c^f \qquad c^{1-\lambda} = \sum_{\tau} \xi_\tau c_\tau^{1-\lambda} \qquad c^f \equiv \sum_{\tau} \frac{r}{z_{K,\tau}^f} G(\widehat{z}_\tau)$$

where task-level MC c_τ and AI comparative advantage $\widehat{z}_\tau$ compatible with firm-level MC c

Automation Decisions: Firm Level

Firm-level costs: characterized by fixed point problem

$$C(w, r; y) = c\,\frac{y}{a} + c^f \qquad c^{1-\lambda} = \sum_{\tau} \xi_\tau c_\tau^{1-\lambda} \qquad c^f \equiv \sum_{\tau} \frac{r}{z_{K,\tau}^f} G(\widehat{z}_\tau)$$

where task-level MC c_τ and AI comparative advantage $\widehat{z}_\tau$ compatible with firm-level MC c

Share of machine in firm-level variable costs and fixed-to-variable cost ratios

$$\omega_K \equiv \frac{r\,k}{c\,(y/a)} = \mathbb{E}_\tau\left[\omega_{K,\tau}\right] \qquad \omega^f \equiv \frac{c^f}{c\,(y/a)} = \mathbb{E}_\tau^f\left[\omega_\tau^f\right]$$

where $\mathbb{E}_\tau[\cdot]$ and $\mathbb{E}_\tau^f[\cdot]$ expectation w.r.t task-τ share in variable ψ_τ and fixed $\psi_\tau \frac{\omega_\tau^f}{\omega^f}$ costs

$$\psi_\tau \equiv \frac{c_\tau y_\tau}{c\,(y/a)} \equiv \xi_\tau \left(\frac{c_\tau}{c}\right)^{1-\lambda}$$

Special Case I: No Fixed Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = 1 \qquad \Rightarrow \qquad CRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \omega_K \qquad \Leftarrow \qquad \omega^f = 0$$

Special Case I: No Fixed Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = 1 \quad \Rightarrow \quad CRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \omega_K \quad \Leftarrow \quad \omega^f = 0$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = 0$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \Omega_K = \omega_K$$

Special Case I: No Fixed Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = 1 \quad \Rightarrow \quad CRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \omega_K \quad \Leftarrow \quad \omega^f = 0$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = 0$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \Omega_K = \omega_K$$

Firm and Task-level Substitution Elasticities:

$$\sigma = \mathbb{E}_\tau^L \left[\sigma_\tau \right] \qquad \sigma_\tau \equiv \frac{1}{\Omega_K} \frac{\partial \log \ell_\tau}{\partial \log r}\bigg|_{y,w} = \lambda + \underbrace{\frac{1}{\omega_K} \widetilde{G}_{\lambda-1}(\widehat{z}_\tau)}_{\text{automation}}$$

Special Case I: No Fixed Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = 1 \quad \Rightarrow \quad CRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \omega_K \quad \Leftarrow \quad \omega^f = 0$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = 0$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \Omega_K = \omega_K$$

Firm and Task-level Substitution Elasticities:

$$\sigma = \mathbb{E}_\tau^L [\sigma_\tau] \qquad \sigma_\tau \equiv \frac{1}{\Omega_K} \frac{\partial \log \ell_\tau}{\partial \log r}\bigg|_{y,w} = \lambda + \underbrace{\frac{1}{\omega_K} \widetilde{G}_{\lambda-1}(\widehat{z}_\tau)}_{\text{automation}}$$

Tasks with higher comparative advantage are more substitutable:

$$\widehat{z}_{\tau'} < \widehat{z}_\tau \quad \Leftrightarrow \quad \sigma_{\tau'} < \sigma_\tau$$

Special Case I: Illustration of Case Without Fixed Costs

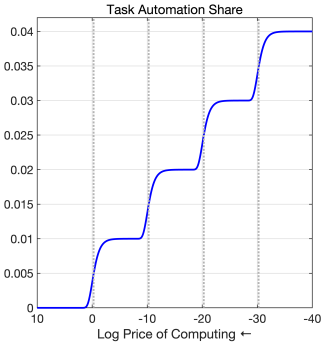
Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right), \qquad \boldsymbol{z}_K^f \equiv (\infty,\ \infty,\ \infty,\ \infty,\ 0)$$

Special Case I: Illustration of Case Without Fixed Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

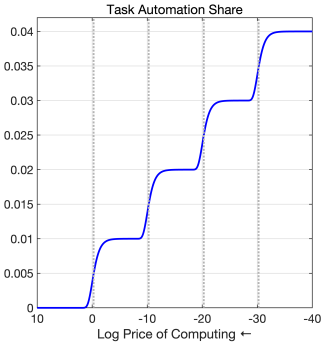
$$Z_\tau = 1, \qquad \boldsymbol{z}_K = \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right), \qquad \boldsymbol{z}_K^f \equiv (\infty,\ \infty,\ \infty,\ \infty,\ 0)$$

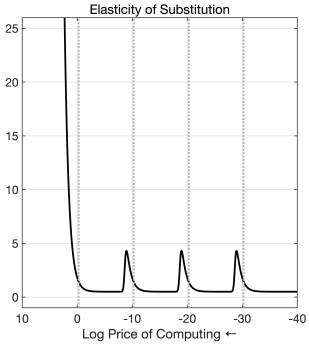


Special Case I: Illustration of Case Without Fixed Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right), \qquad \boldsymbol{z}_K^f \equiv (\infty,\ \infty,\ \infty,\ \infty,\ 0)$$

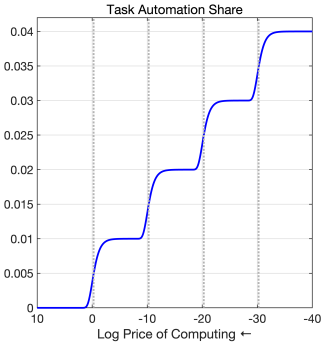


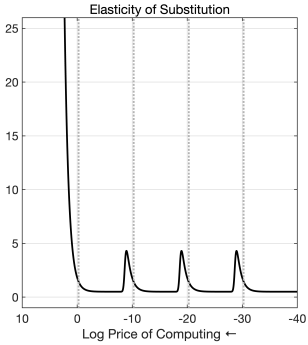


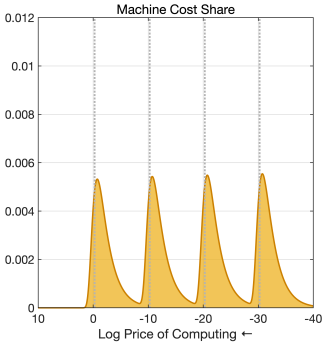
Special Case I: Illustration of Case Without Fixed Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right), \qquad \boldsymbol{z}_K^f \equiv (\infty,\ \infty,\ \infty,\ \infty,\ 0)$$







Special Case II: No Marginal Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{1}{1+\omega^f} < 1 \qquad \Rightarrow \qquad IRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{\omega^f}{1+\omega^f}$$

Special Case II: No Marginal Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{1}{1+\omega^f} < 1 \quad \Rightarrow \quad IRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{\omega^f}{1+\omega^f}$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{\Upsilon}{\lambda} \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\widetilde{G}(\widehat{z}_\tau)\right]}$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \frac{\Upsilon}{\lambda} \qquad \zeta_\tau \equiv 1$$

Special Case II: No Marginal Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{1}{1+\omega^f} < 1 \quad \Rightarrow \quad IRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{\omega^f}{1+\omega^f}$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{\Upsilon}{\lambda} \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\widetilde{G}(\widehat{z}_\tau)\right]}$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \frac{\Upsilon}{\lambda} \qquad \zeta_\tau \equiv 1$$

Demand Elasticities (Firm and Task-level):

$$\eta_K = \frac{\Upsilon}{\lambda\,\omega^f}$$

$$\sigma = \mathbb{E}_\tau^L\left[\sigma_\tau\right] \qquad \sigma_\tau \equiv \frac{1}{\Omega_K}\frac{\partial \log \ell_\tau}{\partial \log r}\bigg|_{y,w} = \underbrace{\Upsilon\,\frac{1}{\Omega_K} + (1-\Upsilon)\,\frac{1}{1-\lambda}\,\frac{1}{\Omega_K}\,\widetilde{G}_{\lambda-1}(\widehat{z}_\tau)}_{\text{automation}}$$

Special Case II: No Marginal Costs

Elasticities of cost:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{1}{1+\omega^f} < 1 \quad \Rightarrow \quad IRS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{\omega^f}{1+\omega^f}$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{\Upsilon}{\lambda} \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\widetilde{G}(\widehat{z}_\tau)\right]}$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \frac{\Upsilon}{\lambda} \qquad \zeta_\tau \equiv 1$$

Demand Elasticities (Firm and Task-level):

$$\eta_K = \frac{\Upsilon}{\lambda\,\omega^f}$$

$$\sigma = \mathbb{E}_\tau^L\left[\sigma_\tau\right] \qquad \sigma_\tau \equiv \frac{1}{\Omega_K}\frac{\partial \log \ell_\tau}{\partial \log r}\bigg|_{y,w} = \underbrace{\Upsilon\,\frac{1}{\Omega_K} + (1-\Upsilon)\,\frac{1}{1-\lambda}\,\frac{1}{\Omega_K}\,\widetilde{G}_{\lambda-1}(\widehat{z}_\tau)}_{\text{automation}}$$

Tasks with higher comparative advantage are more substitutable:

$$\widehat{z}_{\tau'} < \widehat{z}_\tau \qquad \Leftrightarrow \qquad \sigma_{\tau'} < \sigma_\tau$$

Special Case II: Illustration of Case Without Marginal Costs

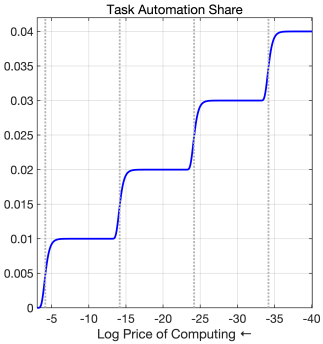
Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty,\ \infty,\ \infty,\ \infty,\ 0)\,, \qquad \boldsymbol{z}_K^f \equiv \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right)$$

Special Case II: Illustration of Case Without Marginal Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

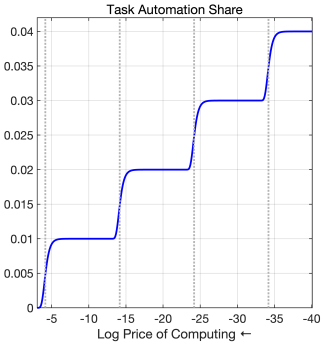
$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty,\ \infty,\ \infty,\ \infty,\ 0)\,, \qquad \boldsymbol{z}_K^f \equiv \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right)$$

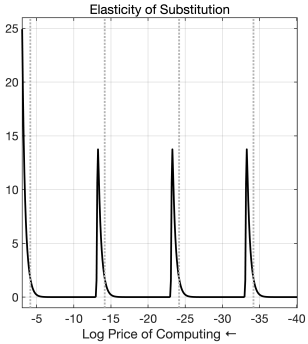


Special Case II: Illustration of Case Without Marginal Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty,\ \infty,\ \infty,\ \infty,\ 0)\,, \qquad \boldsymbol{z}_K^f \equiv \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right)$$

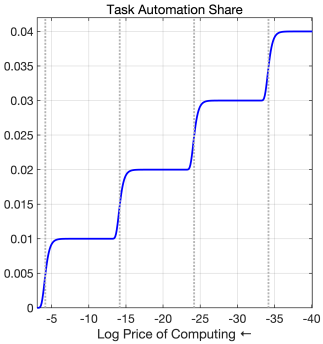


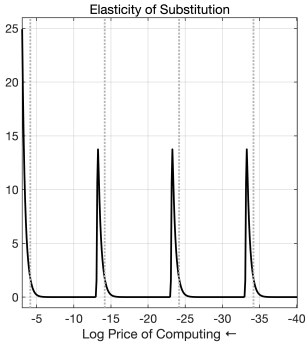


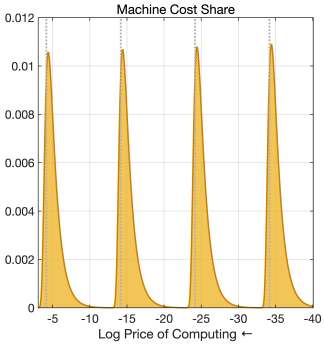
Special Case II: Illustration of Case Without Marginal Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty,\ \infty,\ \infty,\ \infty,\ 0)\,, \qquad \boldsymbol{z}_K^f \equiv \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right)$$



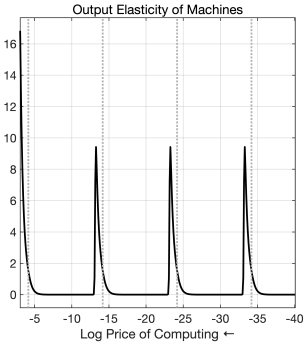




Special Case II: Illustration of Case Without Marginal Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

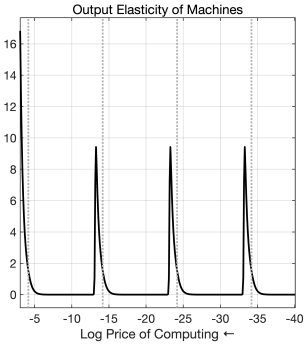
$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty,\ \infty,\ \infty,\ \infty,\ 0)\,, \qquad \boldsymbol{z}_K^f \equiv \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right)$$

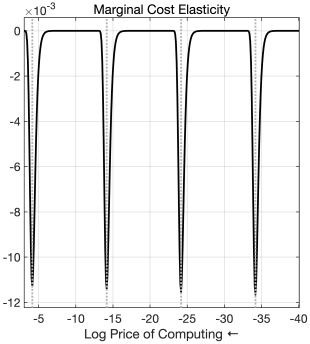


Special Case II: Illustration of Case Without Marginal Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty,\ \infty,\ \infty,\ \infty,\ 0)\,, \qquad \boldsymbol{z}_K^f \equiv \left(1,\ e^{-10},\ e^{-20},\ e^{-30},\ 0\right)$$

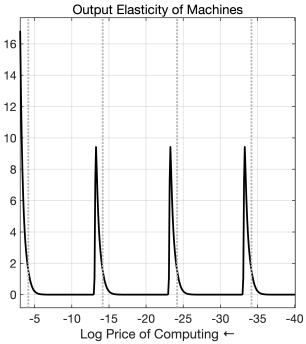


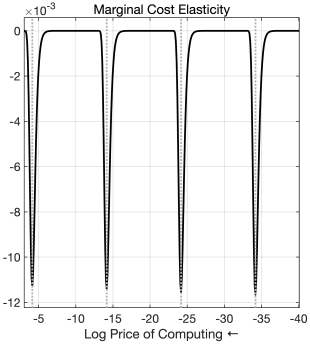


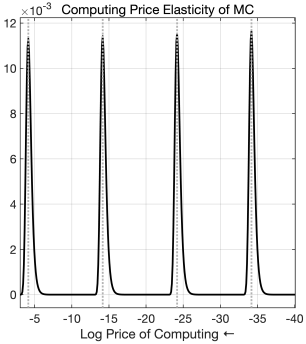
Special Case II: Illustration of Case Without Marginal Costs

Toy Example: consider $\lambda = 0.5$, $\theta = 1.3$, and $|\mathcal{T}| = 5$ with $\xi_\tau = 0.01$ for $\tau \leq 4$.

$$Z_\tau = 1, \qquad \boldsymbol{z}_K = (\infty, \infty, \infty, \infty, 0), \qquad \boldsymbol{z}_K^f \equiv \left(1, e^{-10}, e^{-20}, e^{-30}, 0\right)$$







General Equilibrium Environment

Production: I industries each w. many monopol. comp. single-product firms $j \in \mathcal{I}_i$ with a_j

$$Y^{\frac{\rho-1}{\rho}} = \sum_{1 \leq i \leq I} Y_i^{\frac{\rho-1}{\rho}} \qquad Y_i^{\frac{\eta-1}{\eta}} = \sum_{\mathcal{J}_i} y_j^{\frac{\eta-1}{\eta}}$$

General Equilibrium Environment

Production: I industries each w. many monopol. comp. single-product firms $j \in \mathcal{I}_i$ with a_j

$$Y^{\frac{\rho-1}{\rho}} = \sum_{1 \leq i \leq I} Y_i^{\frac{\rho-1}{\rho}} \qquad Y_i^{\frac{\eta-1}{\eta}} = \sum_{\mathcal{J}_i} y_j^{\frac{\eta-1}{\eta}}$$

Factor Supplies: exogenous supplies $\left(\overline{K}, \overline{L}\right) \Rightarrow$ factor prices (r, w) and output Y

General Equilibrium Environment

Production: I industries each w. many monopol. comp. single-product firms $j \in \mathcal{I}_i$ with a_j

$$Y^{\frac{\rho-1}{\rho}} = \sum_{1 \leq i \leq I} Y_i^{\frac{\rho-1}{\rho}} \qquad\qquad Y_i^{\frac{\eta-1}{\eta}} = \sum_{\mathcal{J}_i} y_j^{\frac{\eta-1}{\eta}}$$

Factor Supplies: exogenous supplies $\left(\overline{K}, \overline{L}\right) \Rightarrow$ factor prices (r, w) and output Y

Aggregate Production: under monopolistic competition, market equilibrium efficient

$$Y = F(K, L) \qquad \Leftrightarrow \qquad \overline{C}(w, r; Y) = \overline{c}\, Y + \overline{c}^f$$

General Equilibrium Environment

Production: I industries each w. many monopol. comp. single-product firms $j \in \mathcal{I}_i$ with a_j

$$Y^{\frac{\rho-1}{\rho}} = \sum_{1 \leq i \leq I} Y_i^{\frac{\rho-1}{\rho}} \qquad\qquad Y_i^{\frac{\eta-1}{\eta}} = \sum_{\mathcal{J}_i} y_j^{\frac{\eta-1}{\eta}}$$

Factor Supplies: exogenous supplies $\left(\overline{K}, \overline{L}\right) \Rightarrow$ factor prices (r, w) and output Y

Aggregate Production: under monopolistic competition, market equilibrium efficient

$$Y = F(K, L) \qquad \Leftrightarrow \qquad \overline{C}(w, r; Y) = \overline{c}\, Y + \overline{c}^f$$

- Elasticities of aggregate (total/marginal) cost and demand:
  Lashkari et al. (2024)

Aggregation

$$\overline{\varepsilon} \equiv \frac{\partial \log \overline{C}}{\partial \log r} \qquad \overline{\varepsilon}^m \equiv \frac{\partial \log \overline{c}}{\partial \log r} \qquad \overline{\eta}_K \equiv \frac{\partial \log K}{\partial \log Y}$$

$$\overline{\Omega}_K \equiv \frac{\partial \log \overline{C}}{\partial \log r} \qquad \overline{\Omega}_K^m \equiv \frac{\partial \log \overline{c}}{\partial \log r} \qquad \overline{\sigma} \equiv \frac{1}{\overline{\Omega}_K} \frac{\partial \log L}{\partial \log r}$$

GE Response to Shock In Machine Prices

Consider decline $\Delta \log r$ in price of machines

GE Response to Shock In Machine Prices

Consider decline $\Delta \log r$ in price of machines

$$\overline{L} = L\left(w, r; Y\right) \qquad \Rightarrow \qquad 0 \approx \overline{\Omega}_K \, \overline{\sigma} \, \left(\Delta \log r - \Delta \log w\right) + \overline{\eta}_L \, \Delta \log Y$$

$$1 = \mu \, \overline{c}\left(w, r; Y\right) \qquad \Rightarrow \qquad 0 \approx \overline{\Omega}_K^m \, \Delta \log r + \overline{\Omega}_L^m \, \Delta \log w + \varepsilon^m \, \Delta \log Y$$

$\overline{\Omega}_L^m$ and $\overline{\eta}_L$

GE Response to Shock In Machine Prices

Consider decline $\Delta \log r$ in price of machines

$$\overline{L} = L\left(w, r; Y\right) \qquad \Rightarrow \qquad 0 \approx \overline{\Omega}_K\, \overline{\sigma}\, \left(\Delta \log r - \Delta \log w\right) + \overline{\eta}_L\, \Delta \log Y$$

$$1 = \mu\, \overline{c}\left(w, r; Y\right) \qquad \Rightarrow \qquad 0 \approx \overline{\Omega}_K^m\, \Delta \log r + \overline{\Omega}_L^m\, \Delta \log w + \varepsilon^m\, \Delta \log Y$$

Response of wage, output, and labor share:

$\overline{\Omega}_L^m$ and $\overline{\eta}_L$

$$\Delta \log w \approx -\frac{\overline{\eta}_L \overline{\Omega}_K^m - \varepsilon^m\, \overline{\Omega}_K \overline{\sigma}}{\overline{\eta}_L \overline{\Omega}_L^m + \varepsilon^m\, \overline{\Omega}_K \overline{\sigma}}\, \Delta \log r$$

$$\Delta \log Y \approx -\frac{\left(\overline{\Omega}_K^m + \overline{\Omega}_L^m\right)\overline{\Omega}_K}{\overline{\eta}_L \overline{\Omega}_L^m + \varepsilon^m\, \overline{\Omega}_K \overline{\sigma}}\, \overline{\sigma}\, \Delta \log r$$

$$\Delta \log \overline{\Omega}_L \approx \frac{\left(\overline{\Omega}_K^m + \overline{\Omega}_L^m\right)\overline{\Omega}_K}{\overline{\eta}_L \overline{\Omega}_L^m + \varepsilon^m\, \overline{\Omega}_K \overline{\sigma}} \left(\frac{\overline{\varepsilon}}{\overline{\eta}_L}\, \overline{\sigma} - 1\right) \Delta \log r$$

GE Response to Shock In Machine Prices

Consider decline $\Delta \log r$ in price of machines

$$\overline{L} = L(w, r; Y) \quad \Rightarrow \quad 0 \approx \overline{\Omega}_K \, \overline{\sigma} \, (\Delta \log r - \Delta \log w) + \overline{\eta}_L \, \Delta \log Y$$

$$1 = \mu \, \overline{c}(w, r; Y) \quad \Rightarrow \quad 0 \approx \overline{\Omega}_K^m \, \Delta \log r + \overline{\Omega}_L^m \, \Delta \log w + \varepsilon^m \, \Delta \log Y$$

$\overline{\Omega}_L^m$ and $\overline{\eta}_L$

Response of wage, output, and labor share:

$$\Delta \log w \approx -\frac{\overline{\eta}_L \overline{\Omega}_K^m - \varepsilon^m \overline{\Omega}_K \overline{\sigma}}{\overline{\eta}_L \overline{\Omega}_L^m + \varepsilon^m \overline{\Omega}_K \overline{\sigma}} \, \Delta \log r = -\frac{\overline{\Omega}_K}{\overline{\Omega}_L} \, \Delta \log r \qquad z_{K,\tau}^f = \infty$$

$$\Delta \log Y \approx -\frac{(\overline{\Omega}_K^m + \overline{\Omega}_L^m)\overline{\Omega}_K}{\overline{\eta}_L \overline{\Omega}_L^m + \varepsilon^m \overline{\Omega}_K \overline{\sigma}} \, \overline{\sigma} \, \Delta \log r = -\overline{\sigma} \, \frac{\overline{\Omega}_K}{\overline{\Omega}_L} \, \Delta \log r$$

$$\Delta \log \overline{\Omega}_L \approx \frac{(\overline{\Omega}_K^m + \overline{\Omega}_L^m)\overline{\Omega}_K}{\overline{\eta}_L \overline{\Omega}_L^m + \varepsilon^m \overline{\Omega}_K \overline{\sigma}} \left(\frac{\overline{\varepsilon}}{\overline{\eta}_L} \overline{\sigma} - 1 \right) \Delta \log r = (\overline{\sigma} - 1) \, \frac{\overline{\Omega}_K}{\overline{\Omega}_L} \, \Delta \log r$$

Wage Output Labor Share

Roadmap

Theory

Application to AI Scaling Laws

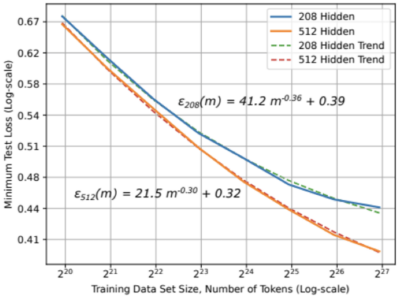
Quantitative Results

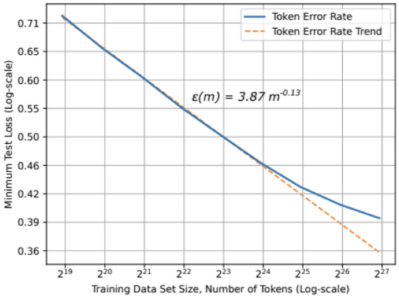
Conclusion

AI Scaling Laws

- Empirical relationship bet. AI performance and computational inputs (model size & data)

  Hestness et al. (2017); Rosenfeld (2020); Kaplan et al. (2020); Zhai et al. (2022)

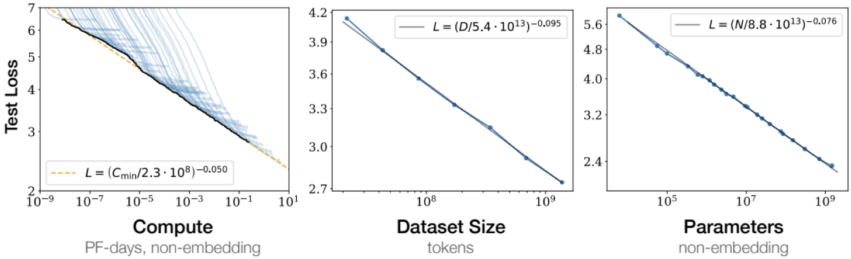




AI Scaling Laws

- Empirical relationship bet. AI performance and computational inputs (model size & data)

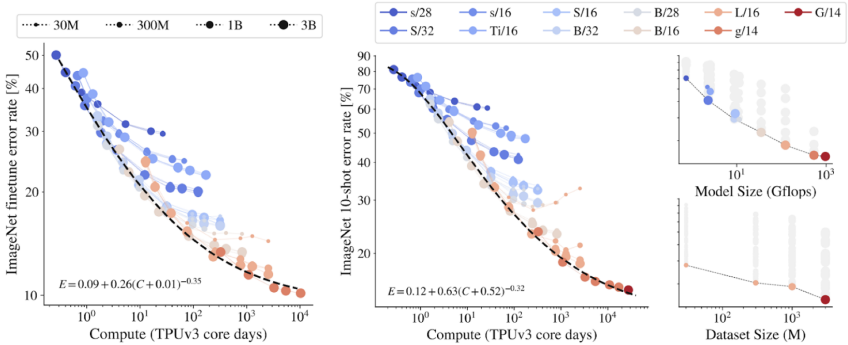
Hestness et al. (2017); Rosenfeld (2020); Kaplan et al. (2020); Zhai et al. (2022)



AI Scaling Laws

- Empirical relationship bet. AI performance and computational inputs (model size & data)

Hestness et al. (2017); Rosenfeld (2020); Kaplan et al. (2020); Zhai et al. (2022)



Scaling Laws Determine Comparative Advantage of AI

Fine-tuning: process of tailoring a pre-trained AI system to specific applications

- Requires inputs model $m_{i\tau}$ and data $d_{i\tau}$ size to lower cross-entropy loss $h_{i\tau}$:

  Li, et al. (2024)

$$h_{i\tau} = \mathcal{H}\left(m_{i\tau}, d_{i\tau}; \chi_{i\tau}\right) = \left[\psi_m\left(\chi_{i\tau}\right)\ m_{i\tau}\right]^{-\alpha_m(\chi_{i\tau})} + \left[\psi_d\left(\chi_{i\tau}\right)\ d_{i\tau}\right]^{-\alpha_d(\chi_{i\tau})} + \underline{h}$$

- Task complexity $\chi_{i\tau}$: number of outcomes

Scaling Laws Determine Comparative Advantage of AI

Fine-tuning: process of tailoring a pre-trained AI system to specific applications

- Requires inputs model $m_{i\tau}$ and data $d_{i\tau}$ size to lower cross-entropy loss $h_{i\tau}$:

  Li, et al. (2024)

$$h_{i\tau} = \mathcal{H}\left(m_{i\tau}, d_{i\tau}; \chi_{i\tau}\right) = \left[\psi_m\left(\chi_{i\tau}\right) m_{i\tau}\right]^{-\alpha_m(\chi_{i\tau})} + \left[\psi_d\left(\chi_{i\tau}\right) d_{i\tau}\right]^{-\alpha_d(\chi_{i\tau})} + \underline{h}$$

- Task complexity $\chi_{i\tau}$: number of outcomes
- Task error rate: required $\epsilon_{i\tau}$ translates to cross-entropy loss $h_{i\tau} = h\left(\epsilon_{i\tau}; \chi_{i\tau}\right)$

Scaling Laws Determine Comparative Advantage of AI

Fine-tuning: process of tailoring a pre-trained AI system to specific applications

- Requires inputs model $m_{i\tau}$ and data $d_{i\tau}$ size to lower cross-entropy loss $h_{i\tau}$:

  Li, et al. (2024)

$$h_{i\tau} = \mathcal{H}\left(m_{i\tau}, d_{i\tau}; \chi_{i\tau}\right) = \left[\psi_m\left(\chi_{i\tau}\right)\ m_{i\tau}\right]^{-\alpha_m(\chi_{i\tau})} + \left[\psi_d\left(\chi_{i\tau}\right)\ d_{i\tau}\right]^{-\alpha_d(\chi_{i\tau})} + \underline{h}$$

- Task complexity $\chi_{i\tau}$: number of outcomes
- Task error rate: required $\epsilon_{i\tau}$ translates to cross-entropy loss $h_{i\tau} = h\left(\epsilon_{i\tau}; \chi_{i\tau}\right)$

Training Cost Minimization: firms choose $d_{i\tau}$, $m_{i\tau}$ to minimize training costs:

$$\frac{1}{z^f_{K\tau}} = \min_{m_{i\tau}, d_{i\tau}} 6 \cdot B \cdot m_{i\tau} d_{i\tau} \quad s.t. \quad h\left(\epsilon_{i\tau}; \chi_{i\tau}\right) = \mathcal{H}\left(m_{i\tau}, d_{i\tau}; \chi_{i\tau}\right)$$

Inference (AI marginal) costs scale with model size $m_{i\tau}$ implying $z_{K,\tau} = 6\, d_{i\tau}\, z^f_{K,\tau}$

Details

Estimating the Fine-Tuning Scaling Law

Computer Vision Fine-Tuning Experiment: estimating scaling-law parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\underline{h}$

- Fine-tune SwinTransformer model varying task complexity (class number), model size, data size Details

Estimating the Fine-Tuning Scaling Law

Computer Vision Fine-Tuning Experiment: estimating scaling-law parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\underline{h}$

- Fine-tune SwinTransformer model varying task complexity (class number), model size, data size Details

Parametrization:

$$\log \alpha_x(\chi_{i\tau}) \equiv \beta^{\alpha}_{x,0} + \beta^{\alpha}_{x,1} \log \chi_{i\tau} \qquad \log \psi_x(\chi_{i\tau}) \equiv \beta^{\psi}_{x,0} + \beta^{\psi}_{x,1} \log \chi_{i\tau} \qquad x \in \{m, d\}$$

Parameter	$\beta^{\psi}_{m,0}$	$\beta^{\psi}_{m,1}$	$\beta^{\alpha}_{m,0}$	$\beta^{\alpha}_{m,1}$	$\beta^{\psi}_{d,0}$	$\beta^{\psi}_{d,1}$	$\beta^{\alpha}_{d,0}$	$\beta^{\alpha}_{d,1}$	$\log \underline{h}$	δ
Value	4.29	-3.88	-1.92	-0.09	-2.50	-1.10	-0.42	-0.08	-27.06	5.11
90% CI	[3.25, 5.40]	[-4.15, -3.64]	[-1.97, -1.87]	[-0.11, -0.07]	[-3.03, -2.06]	[-1.17, -1.01]	[-0.53, -0.29]	[-0.10, -0.05]	[-38.17, -24.91]	[4.68, 5.50]

Estimating the Fine-Tuning Scaling Law

Computer Vision Fine-Tuning Experiment: estimating scaling-law parameters $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\underline{h}$

- Fine-tune SwinTransformer model varying task complexity (class number), model size, data size Details

Parametrization:

$$\log \alpha_x(\chi_{i\tau}) \equiv \beta^{\alpha}_{x,0} + \beta^{\alpha}_{x,1} \log \chi_{i\tau} \qquad \log \psi_x(\chi_{i\tau}) \equiv \beta^{\psi}_{x,0} + \beta^{\psi}_{x,1} \log \chi_{i\tau} \qquad x \in \{m, d\}$$

Parameter	$\beta^{\psi}_{m,0}$	$\beta^{\psi}_{m,1}$	$\beta^{\alpha}_{m,0}$	$\beta^{\alpha}_{m,1}$	$\beta^{\psi}_{d,0}$	$\beta^{\psi}_{d,1}$	$\beta^{\alpha}_{d,0}$	$\beta^{\alpha}_{d,1}$	$\log \underline{h}$	δ
Value	4.29	-3.88	-1.92	-0.09	-2.50	-1.10	-0.42	-0.08	-27.06	5.11
90% CI	[3.25, 5.40]	[-4.15, -3.64]	[-1.97, -1.87]	[-0.11, -0.07]	[-3.03, -2.06]	[-1.17, -1.01]	[-0.53, -0.29]	[-0.10, -0.05]	[-38.17, -24.91]	[4.68, 5.50]

Cross-entropy vs. error rates: estimated non-parametrically within the experiment Details

Roadmap

Theory

Application to AI Scaling Laws

Quantitative Results

Conclusion

Data

Task Complexity Survey:

Sum. Stat. | Task estimates | Validation | Sample

- Used to estimate task-level measures of error tolerance ϵ and complexity χ
- Implemented in two waves, between July and November 2023 covering 271 occupations, 445 tasks

Data

Task Complexity Survey:

Sum. Stat. | Task estimates | Validation | Sample

- Used to estimate task-level measures of error tolerance ϵ and complexity χ
- Implemented in two waves, between July and November 2023 covering 271 occupations, 445 tasks

LLM-Sourced Data:

- Used to estimate comparable measures of task frequency ξ and labor productivity Z_τ

Details

Data

Task Complexity Survey:

Sum. Stat. | Task estimates | Validation | Sample

- Used to estimate task-level measures of error tolerance ϵ and complexity χ
- Implemented in two waves, between July and November 2023 covering 271 occupations, 445 tasks

LLM-Sourced Data:

- Used to estimate comparable measures of task frequency ξ and labor productivity Z_τ Details

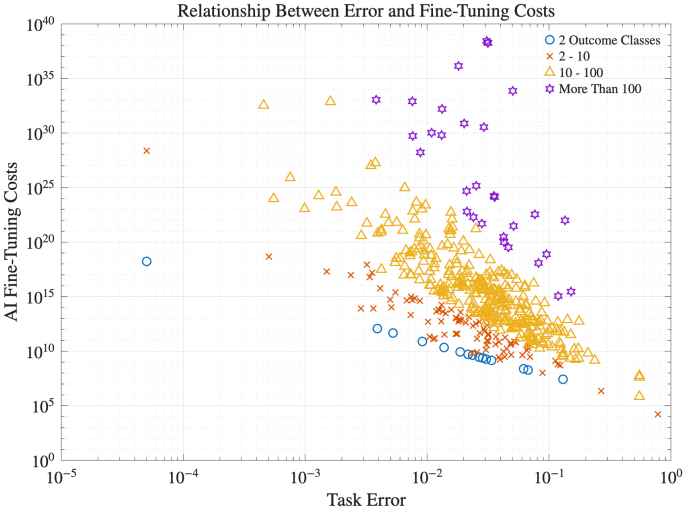
BDS (2023):

- Used to calibrate distribution of firm productivity a_j based on firm size distribution Details | Parameters

Calibrated Parameters

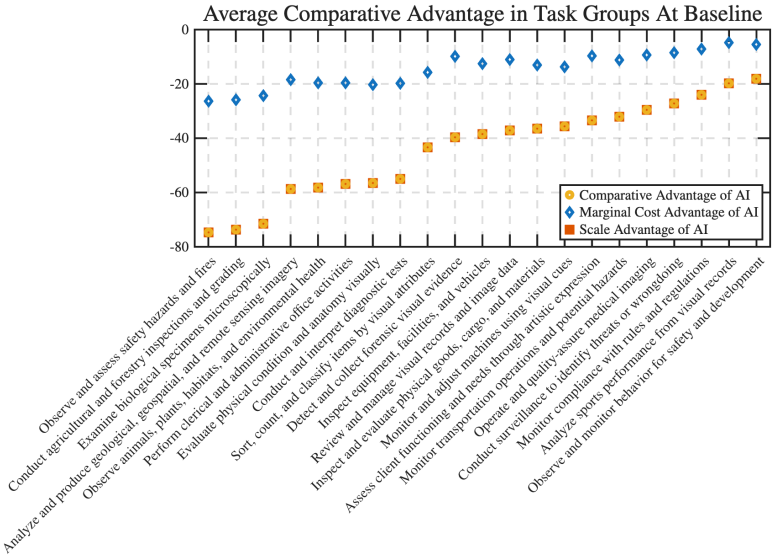
Parameter		Value	Source	Moment	Literature	Model
From Literature						
Task EoS	λ	0.5	Humlum (2019)			
Firm EoS	η	4.0				
Industry EoS	ρ	0.8				
Labor productivity dispersion	θ	1.3	Burstein, et al. (2019)			
Decline in computing prices	$\Delta \log r$	-0.31/-1.35	Del Sozzo, et al. (2025); Ho, et al. (2024)			
Calibration						
Firm productivity	ϑ	1.9	BDS			
Capital price	$\log r$	-13.2	Bonney, et al. (2024)	Fall 2023 ComputerVision adoption rate (unweighted)	0.5%	0.7%
Fixed cost shifter	$\log B$	27.5	Bonney, et al. (2024)	Fall 2023 ComputerVision adoption rate (weighted)	2.4%	2.2%

From Task Error/Complexity to AI Costs



Automation Rate | Task Complexity | Task Error

Comparative and Scale Advantages of AI: 2023

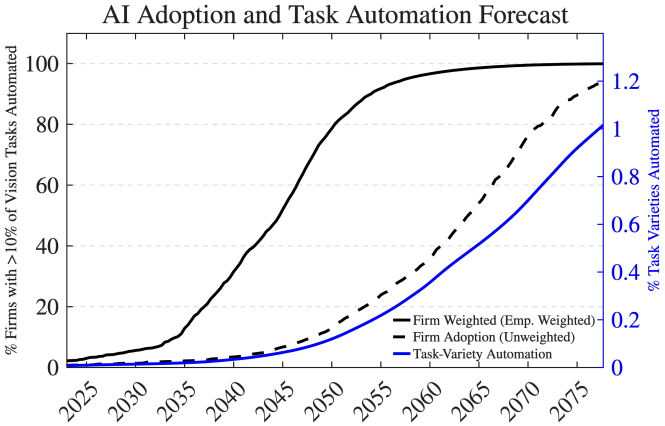


Results for 2050

Predicting Future of Computer Vision Automation

- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP

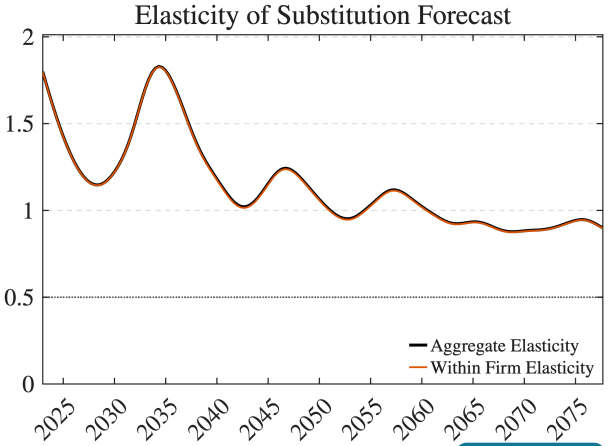
  Del Sozzo, et al. (2025); Ho, et al. (2024)



Predicting Future of Computer Vision Automation

- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP

  Del Sozzo, et al. (2025); Ho, et al. (2024)

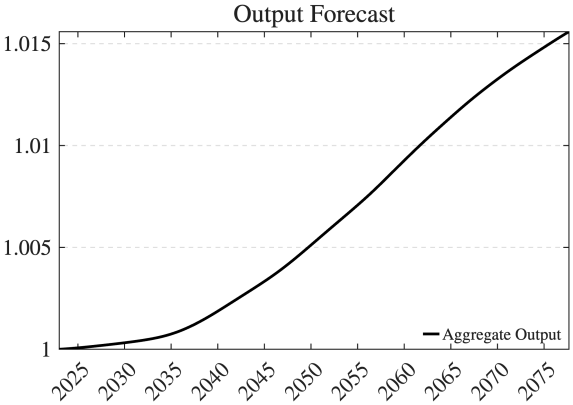


Alternative Calibration | Decline in Algorithmic + Computing Price

Predicting Future of Computer Vision Automation

- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP

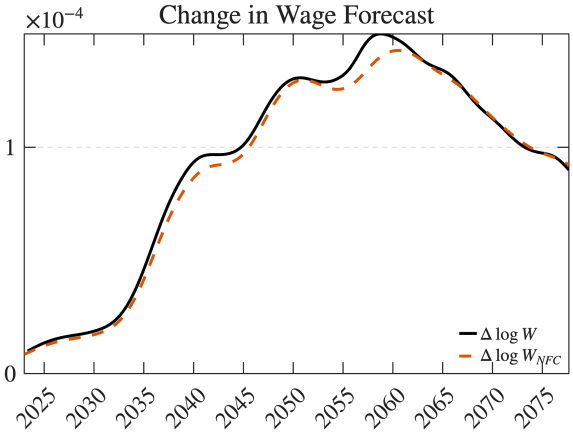
  Del Sozzo, et al. (2025); Ho, et al. (2024)



Predicting Future of Computer Vision Automation

- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP

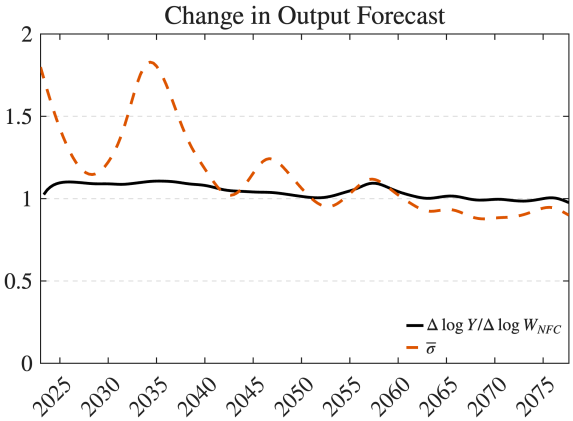
  Del Sozzo, et al. (2025); Ho, et al. (2024)



Predicting Future of Computer Vision Automation

- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP

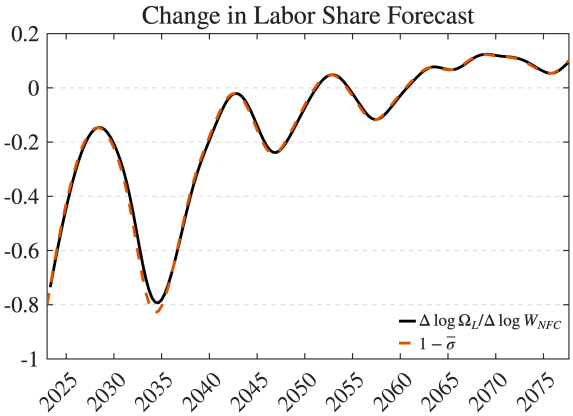
  Del Sozzo, et al. (2025); Ho, et al. (2024)



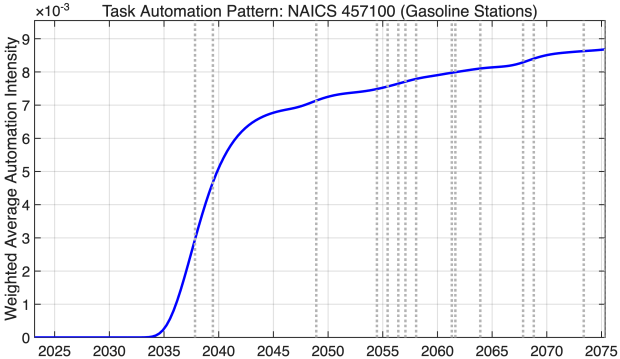
Predicting Future of Computer Vision Automation

- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP

  Del Sozzo, et al. (2025); Ho, et al. (2024)



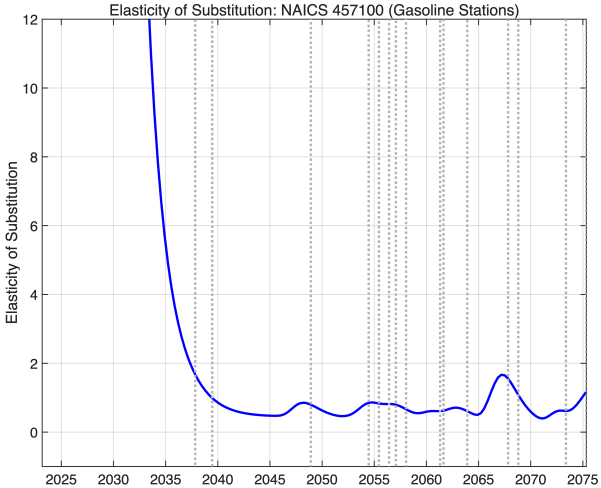
Revisiting Firm-Level Outcomes



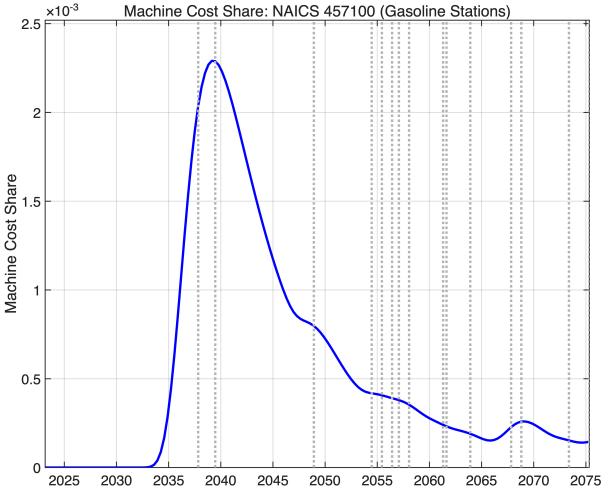
Automated Tasks:

1: Examine merchandise to ensure that it is correctly priced and displayed and that it functions as advertised.
2: Watch for and recognize security risks and thefts and know how to prevent or handle these situations.
3: Inspect supplies, equipment, and work areas to ensure efficient service and conformance to standards.
4: Check conditions of trailers after contents have been unloaded to ensure that there has been no damage.
5: Inspect job sites to determine the extent of maintenance or repairs needed.
6: Sort materials, such as metals, glass, wood, paper or plastics, into appropriate containers for recycling.
7: Observe color of products being baked, and adjust oven temperatures, humidity, or conveyor speeds accordingly.
8: Review work throughout the work process and at completion to ensure that it has been performed properly.
9: Inspect and adjust rental items to meet needs of customer.
10: Monitor and authorize entrance and departure of employees, visitors, and other persons to guard against theft and maintain security of premises.
11: Inspect and clean food preparation areas, such as equipment, work surfaces, and serving areas, to ensure safe and sanitary food-handling practices.
12: Open, sort, and route incoming mail, answer correspondence, and prepare outgoing mail.
13: Examine shipment contents and compare with records, such as manifests, invoices, or orders, to verify accuracy.
14: Inspect loads to ensure that cargo is secure.
15: Inspect equipment for wear and for conformance to specifications.

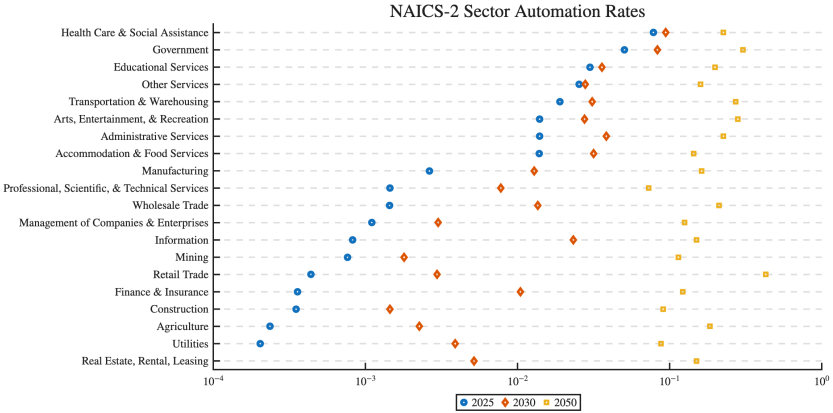
Revisiting Firm-Level Outcomes



Revisiting Firm-Level Outcomes



Industry AI Adoption Rates



Roadmap

Theory

Application to AI Scaling Laws

Quantitative Results

Conclusion

Conclusion

Theory: multi-industry quantitative task-based model of automation

- Potential task-level increasing returns in automation decisions
- Automation determined by patterns of AI marginal cost and scale advantage

Application to Computer Vision AI Automation:

- AI Scaling Law + Task Complexity Survey $\Rightarrow$ AI cost structure
- LLM Data $\Rightarrow$ worker comparative advantage

Results:

- S-curve of adoption of CV over time with declining computing prices
- Scale advantage explains lion's share of cross-task patterns of comparative advantage

Roadmap

Appendix

Appendix

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C}$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C}$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Elasticities of demand:

$$\eta_K \equiv \frac{\partial \log k}{\partial \log y}$$

$$\sigma \equiv \frac{1}{\Omega_K}\frac{\partial \log \ell}{\partial \log r}$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Elasticities of demand:

$$\eta_K \equiv \frac{\partial \log k}{\partial \log y}$$

$$\sigma \equiv \frac{1}{\Omega_K}\frac{\partial \log \ell}{\partial \log r}$$

Elasticities of marginal cost:

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y}$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r}$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Elasticities of demand:

$$\eta_K \equiv \frac{\partial \log k}{\partial \log y}$$

$$\sigma \equiv \frac{1}{\Omega_K}\frac{\partial \log \ell}{\partial \log r}$$

Elasticities of marginal cost: $\widetilde{G}\,(x) \equiv \frac{x\,G(x)}{G(x)}$

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{1}{\lambda}\Upsilon \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}^f_\tau\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}^f_\tau\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]} \qquad \zeta_\tau \equiv \left(\frac{SA_\tau}{\widehat{z}_\tau}\right)^{\lambda-1}$$

$$\Omega^m_K \equiv \frac{\partial \log c}{\partial \log r}$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Elasticities of demand:

$$\eta_K \equiv \frac{\partial \log k}{\partial \log y}$$

$$\sigma \equiv \frac{1}{\Omega_K}\frac{\partial \log \ell}{\partial \log r}$$

Elasticities of marginal cost: $\widetilde{G}\,(x) \equiv \frac{x\,G(x)}{G(x)}$

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{1}{\lambda}\Upsilon \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}^f_\tau\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}^f_\tau\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]} \qquad \zeta_\tau \equiv \left(\frac{SA_\tau}{\widehat{z}_\tau}\right)^{\lambda-1}$$

$$\Omega^m_K \equiv \frac{\partial \log c}{\partial \log r} = \Upsilon + (1-\Upsilon)\left(\omega_K + \omega^f\,\mathbb{E}^f_\tau[\widetilde{G}(\widehat{z}_\tau)]\right)$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Elasticities of demand:

$$\eta_K \equiv \frac{\partial \log k}{\partial \log y} = \frac{1}{\varepsilon}\,\frac{\Omega_K^m}{\Omega_K}$$

$$\sigma \equiv \frac{1}{\Omega_K}\,\frac{\partial \log \ell}{\partial \log r}$$

Elasticities of marginal cost: $\widetilde{G}\,(x) \equiv \frac{x\,G(x)}{G(x)}$

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{1}{\lambda}\Upsilon \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}_\tau^f\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]} \qquad \zeta_\tau \equiv \left(\frac{SA_\tau}{\widehat{z}_\tau}\right)^{\lambda-1}$$

$$\Omega_K^m \equiv \frac{\partial \log c}{\partial \log r} = \Upsilon + (1-\Upsilon)\left(\omega_K + \omega^f\,\mathbb{E}_\tau^f[\widetilde{G}(\widehat{z}_\tau)]\right)$$

Properties of Production Function

Elasticities of cost function:

$$\varepsilon \equiv \frac{\partial \log C}{\partial \log y} = \frac{c\,(y/a)}{C} = \frac{1}{1+\omega^f} \leq 1 \qquad \Rightarrow \qquad (C/I)RS$$

$$\Omega_K \equiv \frac{\partial \log C}{\partial \log r} = \frac{r\,k}{C} = \frac{\omega_K+\omega^f}{1+\omega^f} \geq \omega_K$$

Elasticities of demand: $\widetilde{G}_{\lambda-1}(x) \equiv -\frac{x\,G_{\lambda-1}(x)}{G_{\lambda-1}(x)}, \quad \mathbb{E}^L_\tau[\cdot] \equiv \sum_\tau \psi_\tau \frac{\omega_{L,\tau}}{\omega_L}.$

$$\eta_K \equiv \frac{\partial \log k}{\partial \log y} = \frac{1}{\varepsilon}\,\frac{\Omega^m_K}{\Omega_K}$$

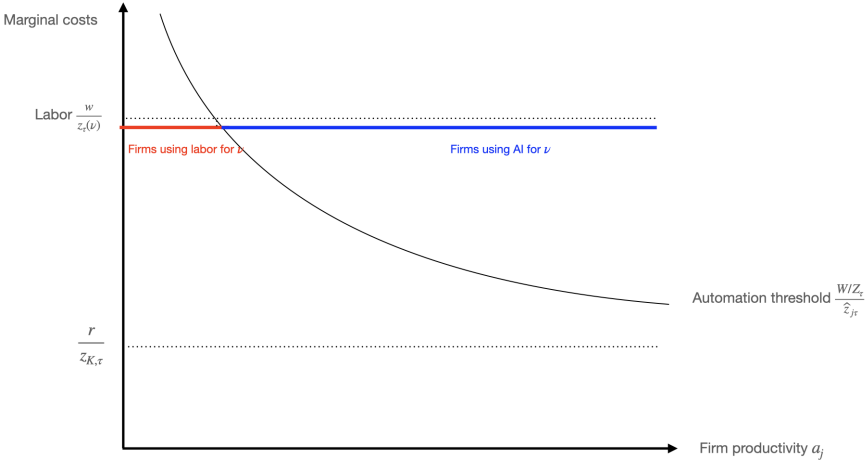
$$\sigma \equiv \frac{1}{\Omega_K}\,\frac{\partial \log \ell}{\partial \log r} = \lambda\,\frac{\Omega^m_K}{\Omega_K} + \frac{1}{\Omega_K}\mathbb{E}^L_\tau[\widetilde{G}_{\lambda-1}(\widehat{z}_\tau)] + \frac{1-\Omega^m_K}{\Omega_K}\,\frac{\lambda}{1-\lambda}\,\mathbb{E}^L_\tau[\zeta_\tau\,\widetilde{G}_{\lambda-1}(\widehat{z}_\tau)]$$

Elasticities of marginal cost: $\widetilde{G}(x) \equiv \frac{x\,G(x)}{G(x)}$

$$\varepsilon^m \equiv \frac{\partial \log c}{\partial \log y} = -\frac{1}{\lambda}\Upsilon \qquad \Upsilon \equiv \frac{\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}^f_\tau\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]}{1+\frac{\lambda}{1-\lambda}\,\omega^f\,\mathbb{E}^f_\tau\left[\zeta_\tau\,\widetilde{G}(\widehat{z}_\tau)\right]} \qquad \zeta_\tau \equiv \left(\frac{SA_\tau}{\widehat{z}_\tau}\right)^{\lambda-1}$$

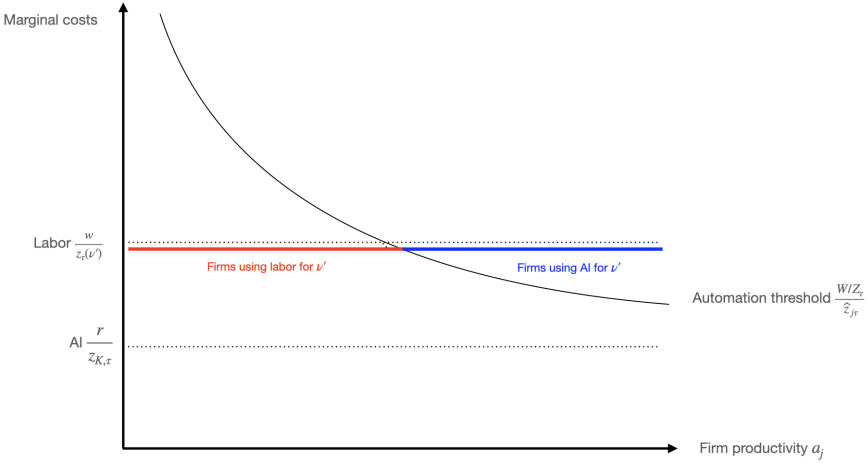
$$\Omega^m_K \equiv \frac{\partial \log c}{\partial \log r} = \Upsilon + (1-\Upsilon)\left(\omega_K + \omega^f\,\mathbb{E}^f_\tau[\widetilde{G}(\widehat{z}_\tau)]\right)$$

Automation



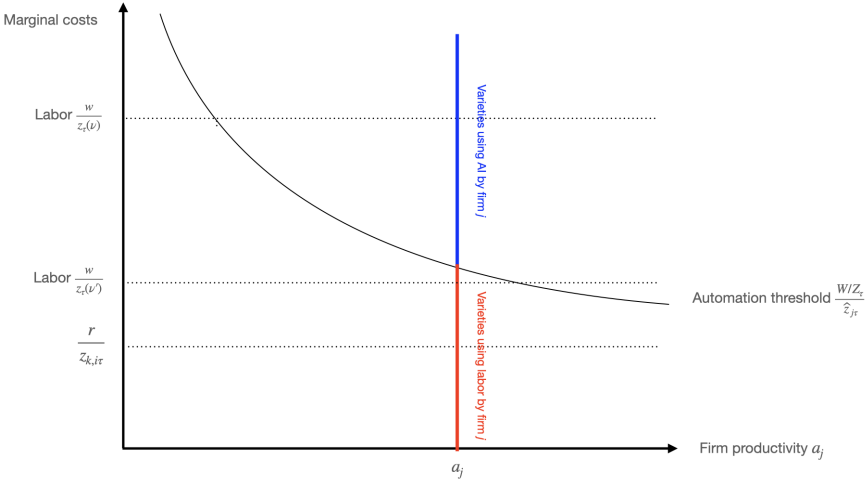
back

Automation



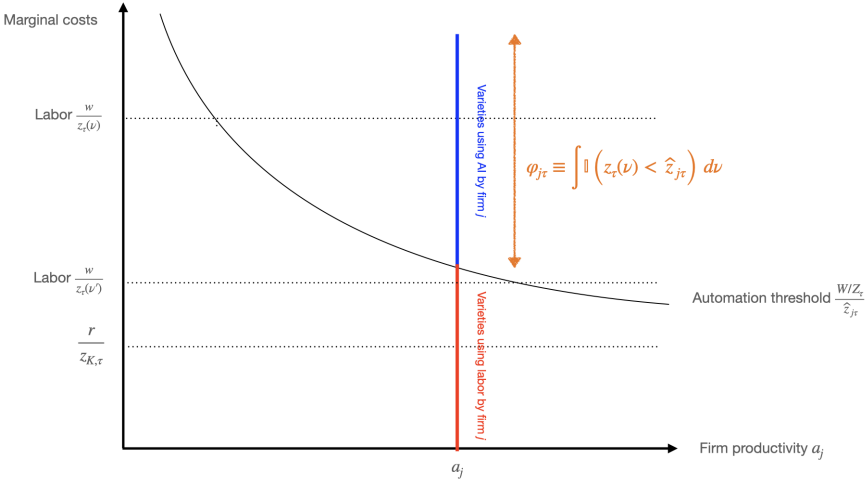
back

Automation



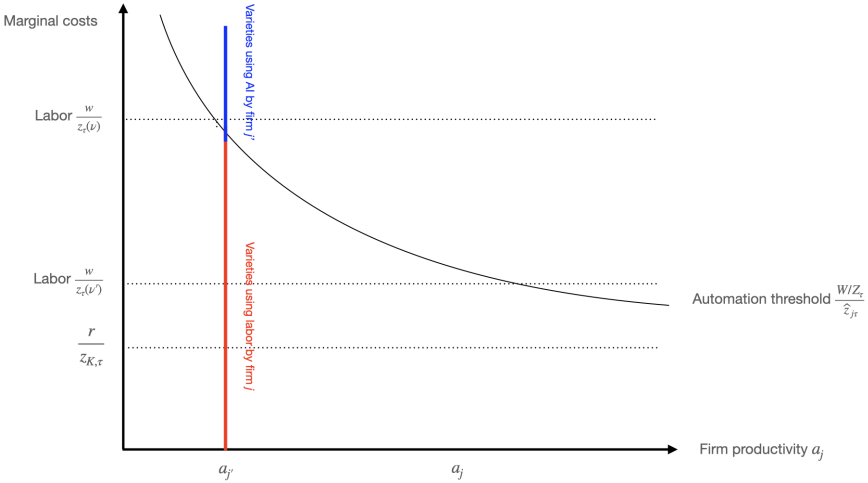
back

Automation



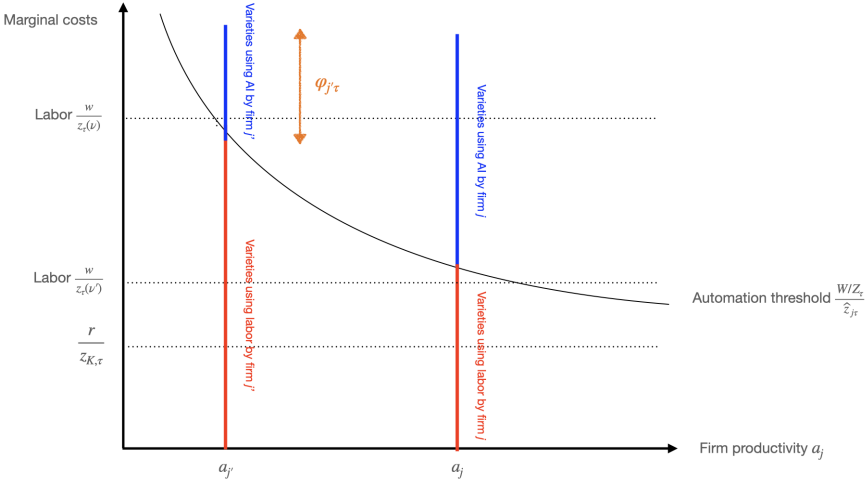
back

Automation



back

Automation



back

Aggregation

Firm-level price and output:

$$p_j = \frac{\eta}{\eta-1}\,\frac{c_j}{a_j} \qquad\qquad y_j = J_i^{\frac{\rho-\eta}{1-\eta}}\, I^{\frac{\rho}{1-\rho}} \left(\frac{c_j}{a_j}\right)^{-\eta} \overline{c}_i^{\eta-\rho}\,\overline{c}^{\rho}\,Y$$

Industry and aggregate average costs: we have $1 = \frac{\eta}{\eta-1}\overline{c}$

$$\overline{c}_i^{1-\eta} = \frac{1}{J_i}\sum_{j\in\mathcal{J}_i}\left(\frac{c_j}{a_j}\right)^{1-\eta} \qquad\qquad \overline{c}^{1-\rho} = \sum_i J_i^{\frac{1-\rho}{1-\eta}}\,\overline{c}_i^{1-\rho}$$

Aggregation

Firm-level price and output:

$$p_j = \frac{\eta}{\eta-1}\frac{c_j}{a_j} \qquad\qquad y_j = J_i^{\frac{\rho-\eta}{1-\eta}} I^{\frac{\rho}{1-\rho}} \left(\frac{c_j}{a_j}\right)^{-\eta} \overline{c}_i^{\eta-\rho}\,\overline{c}^{\rho}\,Y$$

Industry and aggregate average costs: we have $1 = \frac{\eta}{\eta-1}\overline{c}$

$$\overline{c}_i^{1-\eta} = \frac{1}{J_i}\sum_{j\in\mathcal{J}_i}\left(\frac{c_j}{a_j}\right)^{1-\eta} \qquad\qquad \overline{c}^{1-\rho} = \sum_i J_i^{\frac{1-\rho}{1-\eta}}\overline{c}_i^{1-\rho}$$

Firm and industry shares in variable costs:

$$\Psi_{j,i} \equiv \frac{c_j y_j/a_j}{J_i\overline{c}_i Y_i} = \frac{1}{J_i}\left(\frac{c_j/a_j}{\overline{c}_i}\right)^{1-\eta} \qquad\qquad \Psi_i \equiv J_i^{\frac{1}{1-\eta}}\frac{\overline{c}_i Y_i}{\overline{c}\,Y} = J_i^{\frac{1-\rho}{1-\eta}}\left(\frac{\overline{c}_i}{\overline{c}}\right)^{1-\rho}$$

Aggregation

Firm-level price and output:

$$p_j = \frac{\eta}{\eta-1}\frac{c_j}{a_j} \qquad y_j = J_i^{\frac{\rho-\eta}{1-\eta}} I^{\frac{\rho}{1-\rho}} \left(\frac{c_j}{a_j}\right)^{-\eta} \overline{c}_i^{\eta-\rho}\,\overline{c}^{\rho}\,Y$$

Industry and aggregate average costs: we have $1 = \frac{\eta}{\eta-1}\overline{c}$

$$\overline{c}_i^{1-\eta} = \frac{1}{J_i}\sum_{j\in\mathcal{J}_i}\left(\frac{c_j}{a_j}\right)^{1-\eta} \qquad \overline{c}^{1-\rho} = \sum_i J_i^{\frac{1-\rho}{1-\eta}}\overline{c}_i^{1-\rho}$$

Firm and industry shares in variable costs:

$$\Psi_{j,i} \equiv \frac{c_j y_j/a_j}{J_i\overline{c}_i Y_i} = \frac{1}{J_i}\left(\frac{c_j/a_j}{\overline{c}_i}\right)^{1-\eta} \qquad \Psi_i \equiv J_i^{\frac{1}{1-\eta}}\frac{\overline{c}_i Y_i}{\overline{c}\,Y} = J_i^{\frac{1-\rho}{1-\eta}}\left(\frac{\overline{c}_i}{\overline{c}}\right)^{1-\rho}$$

Shares of skill-group and machines in aggregate variable costs:

$$\overline{\omega}_L \equiv \frac{w\,L}{\overline{c}\,Y} = \sum_{i,j}\Psi_i\Psi_{j,i}\,\omega_{L,j} \qquad \overline{\omega}_K \equiv \frac{r\overline{K}-\overline{C}^f}{\overline{c}\,Y} = \sum_{i,j}\Psi_i\Psi_{j,i}\,\omega_{K,j}$$

Firm-Level Scale Elasticities

Firm-level elasticities of cost and marginal cost:

$$\varepsilon_j \equiv \frac{\partial \log C_j}{\partial \log y_j} = \frac{c_j y_j}{C_j a_j} = \frac{1}{1+\omega_{f,j}} \qquad \varepsilon_j^m \equiv \frac{\partial \log c_j}{\partial \log y_j} = \frac{\omega_j^f \, \mathbb{E}_{\tau|j}^f \left[\theta \log \varphi_{j\tau} \, \zeta_{j\tau}\right]}{1-\lambda-\lambda \, \omega_j^f \, \mathbb{E}_{\tau|j}^f \left[\theta \log \varphi_{j\tau} \, \zeta_{j\tau}\right]}$$

with $\mathbb{E}_{\tau|j}^f [\cdot]$ task average w.r.t to dist. of fixed costs $\psi_{\tau,j} \, \omega_{j\tau}^f / \omega_j^f$ and $\zeta_{j\tau}$ importace of fixed costs:

$$\zeta_{j\tau} \equiv \frac{(1-\lambda) \, z_{K,\tau} / z_{K,\tau}^f y_{K,j\tau}}{1+(1-\lambda) \, z_{K,\tau} / z_{K,\tau}^f y_{K,j\tau}}$$

Firm-level scale elasticities of demand:

$$\varepsilon_{K,j} \equiv \frac{\partial \log k_j}{\partial \log y_j} = \frac{\omega_{K,j} - \lambda \varepsilon_j^m \, \omega_{L,j}}{\omega_{K,j} + \omega_j^f} - \theta \left(1 + \lambda \varepsilon_j^m\right) \frac{\omega_j^f}{\omega_{K,j} + \omega_j^f} \mathbb{E}_{\tau|j}^f \left[\log \varphi_{j\tau}\right]$$

$$\varepsilon_{L,j} \equiv \frac{\partial \log \ell_{s,j}}{\partial \log y_j} = \left(1 + \lambda \varepsilon_j^m\right) \left(1 + \theta \, \frac{\omega_j^f}{\omega_{L,j}} \, \mathbb{E}_{\tau|j}^f \left[\log \varphi_{j\tau}\right]\right)$$

back

Aggregate-Level Scale Elasticities

Industry- and aggregate-level scale elasticities of cost:

$$\overline{\varepsilon}_i^m \equiv \frac{1}{\eta}\left(\mathbb{E}_{j|i}\left[\left(1+\eta\,\varepsilon_j^m\right)^{-1}\right]-1\right) \qquad \overline{\varepsilon}^m \equiv \frac{1}{\rho}\left(\mathbb{E}_i\left[\left(1+\rho\,\overline{\varepsilon}_i^m\right)^{-1}\right]-1\right)$$

Pass-through distribution:

$$\Psi_{j,i}^p \equiv \Psi_{j,i}\frac{1+\eta\,\overline{\varepsilon}_i^m}{1+\eta\,\varepsilon_j^m} \qquad \Psi_i^p \equiv \Psi_i\frac{1+\rho\,\overline{\varepsilon}^m}{1+\rho\,\overline{\varepsilon}_i^m}$$

Industry- and aggregate-level scale elasticities of demand:

$$\overline{\eta}_{K,i} \equiv \frac{1}{\overline{\omega}_{K,i}}\mathbb{E}_{j|i}^p\left[\omega_{K,j}\varepsilon_{K,j}\right] \qquad \overline{\eta}_{L,i} \equiv \frac{1}{\overline{\omega}_{L,i}}\mathbb{E}_{j|i}^p\left[\omega_{L,j}\varepsilon_{L,j}\right]$$

$$\overline{\eta}_K \equiv \frac{1}{\overline{\omega}_K}\mathbb{E}_i^p\left[\overline{\omega}_{K,i}\overline{\varepsilon}_{K,i}\right] \qquad \overline{\eta}_L \equiv \frac{1}{\overline{\omega}_s}\mathbb{E}_{j|i}^p\left[\omega_{s,j}\varepsilon_{s,j}\right]$$

where $\mathbb{E}_{j|i}^p[\cdot]$ and $\mathbb{E}_i^p[\cdot]$ denote the pass-through-weighted averages within and across industries

back

Aggregate-Level Substitution Elasticities

Aggregate-level substitution elasticities of demand:

$$\overline{\sigma} = \mathbb{E}_j^c \left[\frac{\Omega_{L,j}}{\Omega_L} \frac{\Omega_{K,j}}{\Omega_K} \sigma_j \right] - \mathbb{E}_i \left[\frac{\eta \overline{\varepsilon}}{1+\eta \overline{\varepsilon}_i^m} \mathbb{C}_{j|i}^p \left(\frac{\omega_{L,j} \varepsilon_{L,j}}{\Omega_L}, \frac{\omega_{K,j} \varepsilon_{K,j}}{\Omega_K} \right) \right] - \frac{\rho \overline{\varepsilon}}{1+\rho \overline{\varepsilon}^m} \mathbb{C}_i^p \left(\frac{\overline{\omega}_{L,i} \overline{\varepsilon}_{L,i}}{\Omega_s}, \frac{\overline{\omega}_{K,i} \overline{\varepsilon}_{K,i}}{\Omega_K} \right)$$

where $\mathbb{E}_j^c [\cdot]$ denotes the total-cost-weighted average across all firms under the distribution $\frac{\overline{\varepsilon}}{\varepsilon_j} \Psi_i \Psi_{ji}$, $\mathbb{C}_{j|i}^p (\cdot, \cdot)$ and $\mathbb{C}_i^p (\cdot, \cdot)$denote the within-industry and across-industry covariances under the pass-through distributions $\Psi_{j,i}^p$ and Ψ_i^p, respectively, and where $\overline{\varepsilon} \equiv 1/(1 + \overline{\omega}^f)$ denotes the aggregate scale elasticity of costs.

back

Deriving $\overline{\Omega}_L^m$ and $\overline{\eta}_L$

Industry- and aggregate-level scale elasticities of cost:

$$\overline{\Omega}_L^m = \frac{\overline{\eta}_K - \overline{\Omega}_K^m}{1-\overline{\Omega}_K}\left(\frac{\overline{\varepsilon}}{\overline{\eta}_K} - \overline{\Omega}_K\right)$$

$$\overline{\eta}_L = \frac{1}{1-\overline{\Omega}_K}\left(\overline{\varepsilon} - \overline{\Omega}_K\overline{\eta}_K\right)$$

back

Cost Structure of AI

AI Cost Minimization: firms choose $d_{i\tau}$, $m_{i\tau}$ to minimize training costs $\equiv 6 \cdot m_{i\tau} d_{i\tau}$

- FLOPs during training may be approximated by $m_{i\tau} d_{i\tau}$

  Rosenfeld (2021)

- Scalar 6 arises from number of passes in standard fine-tuning process

  Sevilla, et al. (2022)

Cost Structure of AI

AI Cost Minimization: firms choose $d_{i\tau}$, $m_{i\tau}$ to minimize training costs $\equiv 6 \cdot m_{i\tau} d_{i\tau}$

- FLOPs during training may be approximated by $m_{i\tau} d_{i\tau}$

  Rosenfeld (2021)

- Scalar 6 arises from number of passes in standard fine-tuning process

  Sevilla, et al. (2022)

Implied AI fixed (training) and marginal (inference) productivity:

$$z^f_{K,\tau} = \frac{1}{6}\left(\frac{1}{\psi_m}\frac{\alpha_d}{\alpha_m+\beta_d}\right)^{\frac{1}{\alpha_m}}\left(\frac{1}{\psi_d}\frac{\alpha_m}{\alpha_m+\alpha_d}\right)^{\frac{1}{\alpha_d}}(h_{i\tau}-\underline{h})^{\frac{1}{\alpha_m}+\frac{1}{\alpha_d}}$$

$$z_{K,\tau} = \left(\frac{1}{\psi_m}\frac{\alpha_d}{\alpha_m+\alpha_d}\right)^{\frac{1}{\alpha_m}}(h_{i\tau}-\underline{h})^{\frac{1}{\alpha_m}}$$

back

Estimating the Entropy Function

Computer Vision Fine-Tuning Experiment: estimating $h\left(\epsilon_{i\tau}; \chi_{i\tau}\right)$

- Experiment varies task complexity (class number), achieved error, and cross-entropy loss

Details

Estimating the Entropy Function

Computer Vision Fine-Tuning Experiment: estimating $h(\epsilon_{i\tau}; \chi_{i\tau})$

- Experiment varies task complexity (class number), achieved error, and cross-entropy loss

Details

Parametrization:

$$\log h(\epsilon_{i\tau}; \chi_{i\tau}) \equiv \beta_0^h + \beta_1^h \log \epsilon_{i\tau} + \beta_2^h \log \chi_{i\tau} + \beta_3^h \log \epsilon_{i\tau} \times \log \chi_{i\tau}$$
$$+ \beta_3^h \log \epsilon_{i\tau} \times \log \chi_{i\tau} + \beta_4^h \frac{1}{\chi_{i\tau}} + \beta_5^h \epsilon_{i\tau} \times \log \epsilon_{i\tau}$$

Estimated Parameters of the Mapping from Error to Cross-Entropy Loss

Parameter	β_1^h	β_2^h	β_3^h	β_0^h	β_4^h	β_5^h
Estimate	0.28	0.21	0.13	0.51	-1.75	0.36

back

Roadmap

Appendix

Appendix

Task estimates from survey data

Task accuracy: worst error rate at which worker can complete task & still be considered qualified

Task complexity: number of possible choices/categories in outcome

Residualize outcomes $y_{o\tau,\ell}$ for occupation o, task τ, respondent i:

$$y_{o\tau,\ell} = \alpha + \beta \cdot Duration_{o\tau,i} + \gamma_{o\tau} + \Gamma' \boldsymbol{X_{o\tau,i}} + \epsilon_{o\tau,i}$$

- $\boldsymbol{X_{o\tau,i}}$ includes confidence level, language, age, gender, experience, & income fixed effects
- $\gamma_{o\tau}$ are occupation-task estimates excluding effect of respondent characteristics

Aggregate $\gamma_{o\tau}$ using occupation share in industry by wage bill using CPS

Comparison with other measures | Correlation with O*NET importance | Back

Relationship Between Variables

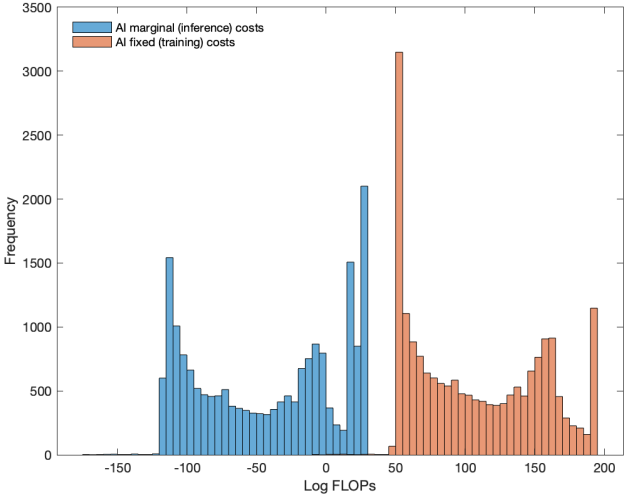
	Log CV Task Characteristics				
	(1)	(2)	(3)	(4)	(5)
	Max. Error	Typ. Error	Complexity	Frequency	Time Use
Log O*NET Importance	-1.584**	-1.059	1.091	0.725	-0.202
	(0.494)	(0.628)	(0.727)	(0.778)	(0.236)
N	505	504	418	505	505
Element FE	✓	✓	✓	✓	✓
2-digit SOC FE	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

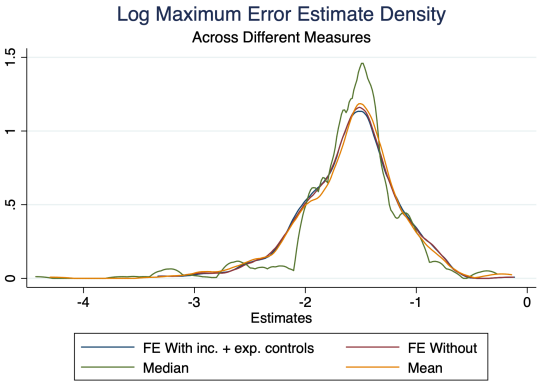
Back

Distribution of Estimated AI Costs



Back

Task accuracy kernel densities under different measures



Back

Correlation of survey data with O*NET Importance

	Log CV Task Characteristics				
	(1)	(2)	(3)	(4)	(5)
	Max. Error	Typ. Error	Complexity	Frequency	Time Use
Log O*NET Importance	-1.584**	-1.059	1.091	0.725	-0.202
	(0.494)	(0.628)	(0.727)	(0.778)	(0.236)
N	505	504	418	505	505
Element FE	✓	✓	✓	✓	✓
2-digit SOC FE	✓	✓	✓	✓	✓

Standard errors in parentheses

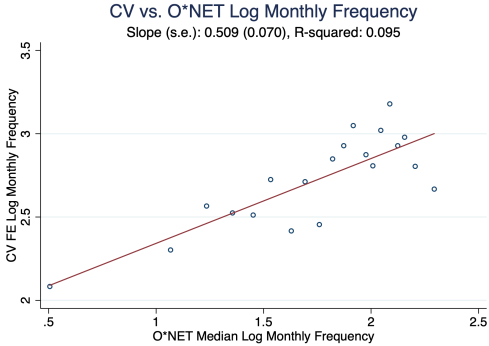
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Back

Validation of Survey Data with O*NET

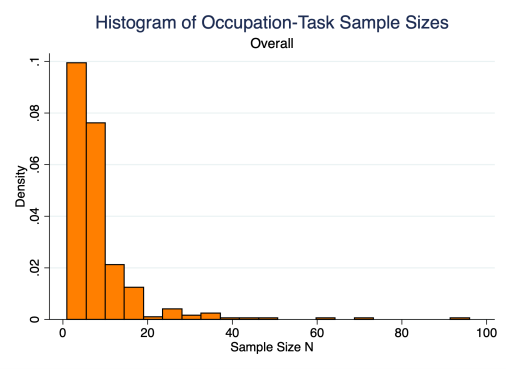
Task frequency: both O*NET and our survey measure monthly occurence of tasks

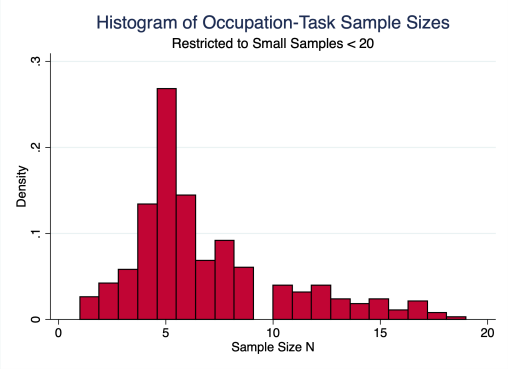
- Survey question: "How often do workers make judgments for the visual part of this task?"
- O*NET question: "How often do workers perform this task?"



Back

Survey Sample Sizes





Back

Firm Productivity Calibration

Suppose productivity a_{ij} of firm j in industry i follow a Pareto distribution:

$$G(a) = 1 - \left(\frac{\underline{a}_i}{a}\right)^{\psi}$$

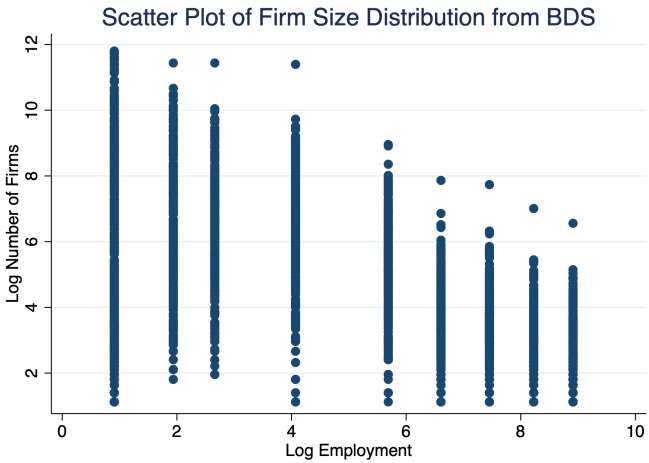
Since $\ell_{ij} \propto a_{ij}^{\eta-1}$,

Scatter plot in BDS

$$\log G_{ij}^{-1} = -\frac{\psi}{\eta - 1} \log \ell_{ij} + \psi \log \underline{a}_i$$

Back

Scatter Plot of Firm Size using BDS



Back

LLM Generated Labor Productivity

AI Model: GPT-5

System and User Prompt: Generates task frequency and time use for all O*NET tasks

Prompt | Data Validation

- Task Frequency:

  Calculate task frequencies at the O*NET task and industry level

$$\xi_{i\tau} = \sum \frac{w_{o,i} l_{o,i}}{\sum_{o',i} w_{o',i} l_{o',i}} \cdot \xi_{o,\tau} \qquad \xi_{o,\tau} = \frac{Freq_{\tau,o}}{\sum_{\tau'} Freq_{\tau',o}}$$

- Task-Level Labor Productivity:

  Calculate productivity for set of vision tasks ($\mathcal{T}_V$) at the O*NET task level

  Calculate productivity for non-vision tasks at the GWA (Generalized Work Activity) level

$$Z_\tau = \frac{Freq_\tau}{TimeUse_\tau \times w_o} \qquad \tau \in \mathcal{T}_V$$

$$Z_{GWA} = \frac{1}{|GWA|} \sum_{\tau \in GWA} \frac{Freq_\tau}{TimeUse_\tau \times w_o} \qquad GWA \not\subset \mathcal{T}_V$$

Back

LLM Prompt

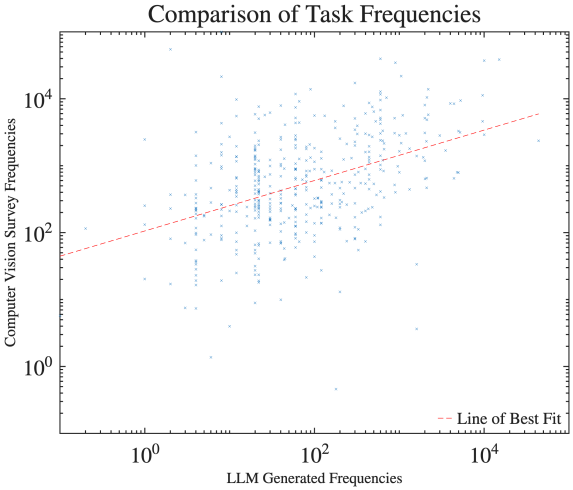
System Prompt: You are a(n) {occupation} participating in a research study about job tasks. You will be asked to answer questions about job tasks associated with your occupation. Think through your answer carefully. You must return a strictly positive (non-zero) numeric answer. Provide no explanation for your answer. Keep in mind that the following tasks are associated with your occupation: {List of Tasks}

Question 1: How often does a worker in your occupation perform this task? Answer with a single positive (non-zero) number representing the number of times per month.

Question 2: How many hours per month does a worker actively spend performing this task? Exclude background or implicit activities (e.g., general awareness, compliance, or principles). Answer with a single strictly positive (non-zero) number representing the hours per month.

Back

LLM Data Validation: Computer Vision Survey



Back

AI Scaling Law Experiment

AI Model: SwinTransformer fine-tuned on unseen classes

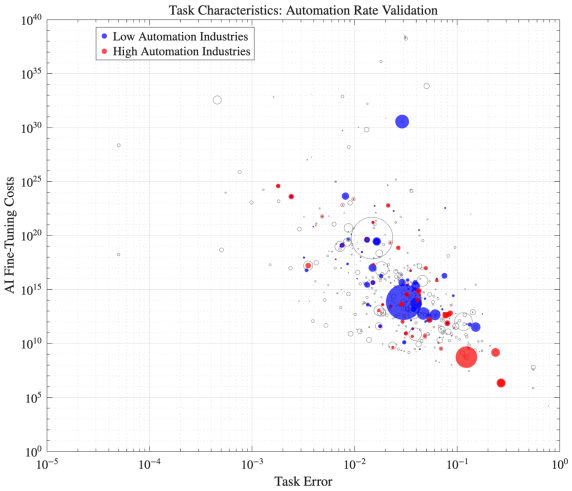
Dataset: ImageNet

Experimental setup: 80 configurations varying:

- Data size: 5 levels (13/class to 1,300/class)
- Model size: 4 levels (7.3K to 87.8M parameters)
- Task complexity: 4 levels (2 to 500 classes)
- Performance tracked across training steps

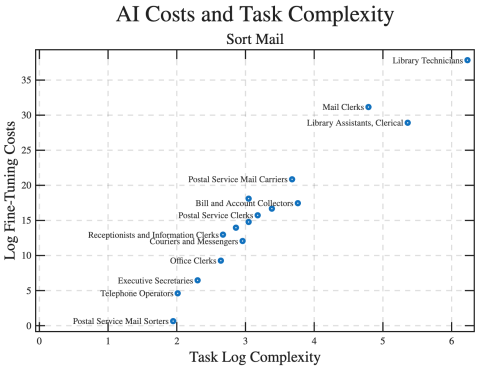
Back

Scaling Law Costs by Automation Rate



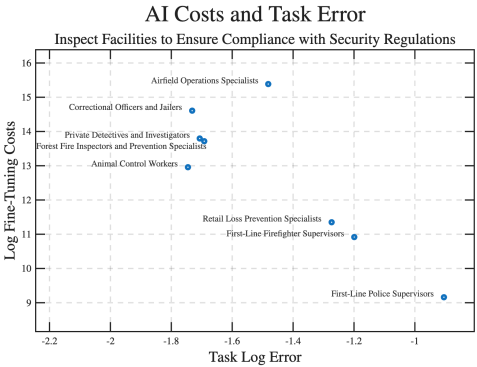
Back

Task Complexity Example



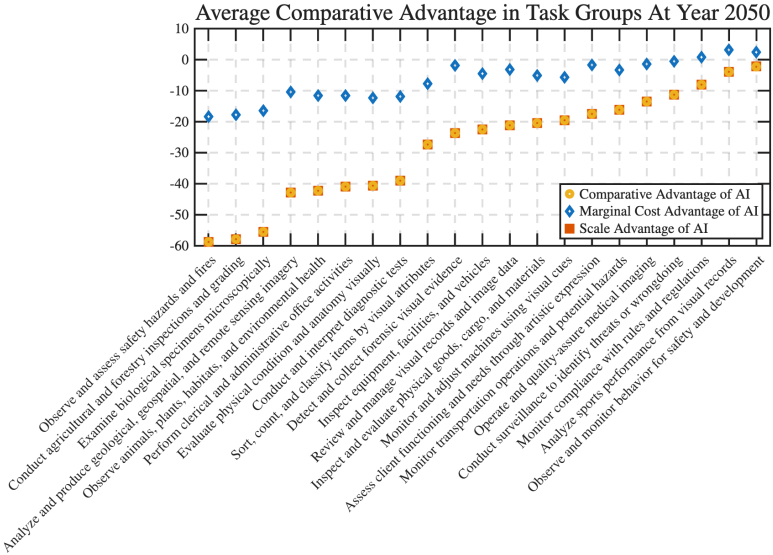
Back

Task Error Example



Back

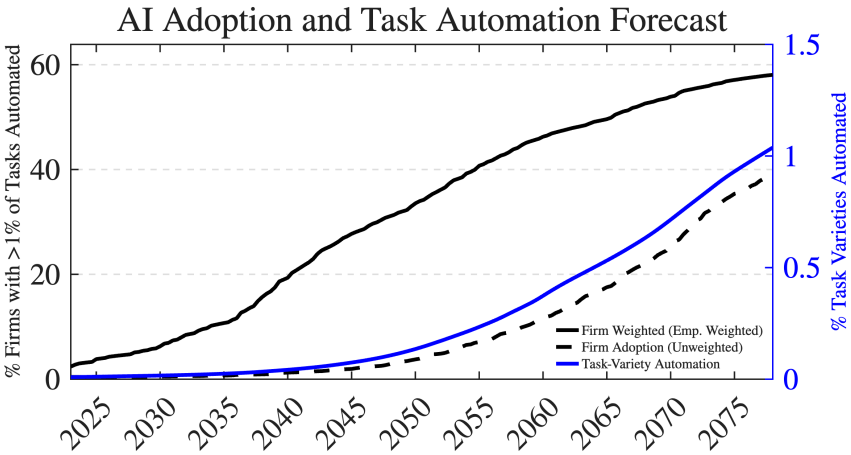
Comparative and Scale Advantage of AI: 2050



Back

Predicting Future of Computer Vision Automation

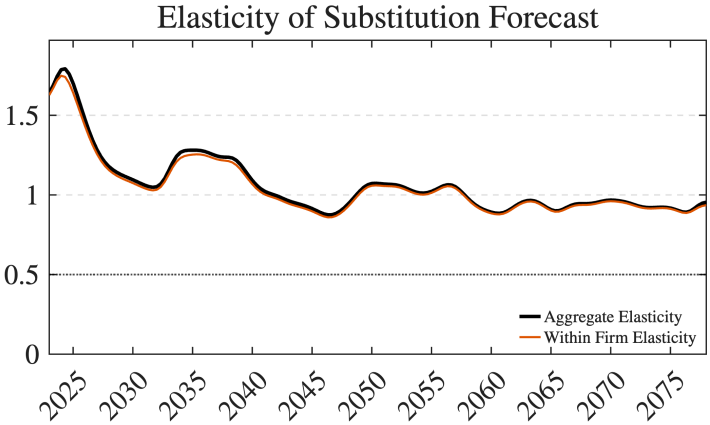
- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP
  Del Sozzo, et al. (2025); Ho, et al. (2024)



Back

Predicting Future of Computer Vision Automation

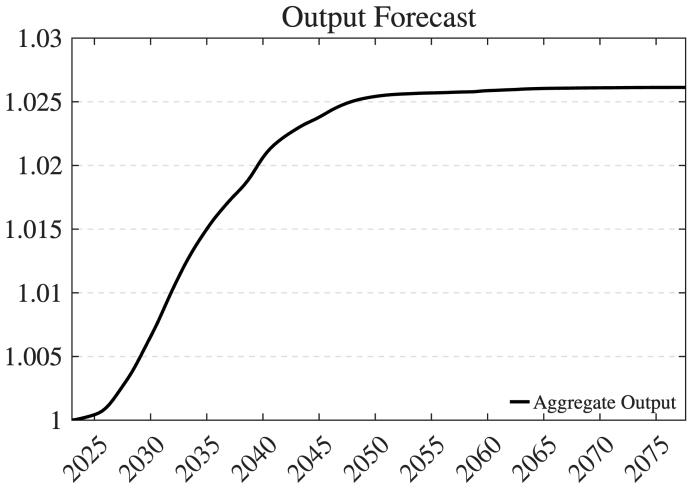
- Trends in computing price r: annual decline of 26.4% in $ cost of FLOP
  Del Sozzo, et al. (2025); Ho, et al. (2024)



Back

Predicting Future of Computer Vision Automation

- Trends in computing cost r: annual decline of 74.0% in (hardware + algorithmic) $ cost of FLOP
  Del Sozzo, et al. (2025); Ho, et al. (2024)

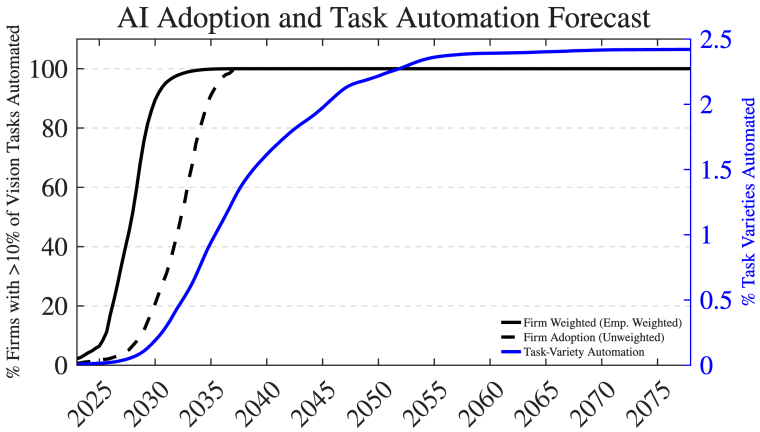


Back

Predicting Future of Computer Vision Automation

- Trends in computing cost r: annual decline of 74.0% in (hardware + algorithmic) $ cost of FLOP

  Del Sozzo, et al. (2025); Ho, et al. (2024)

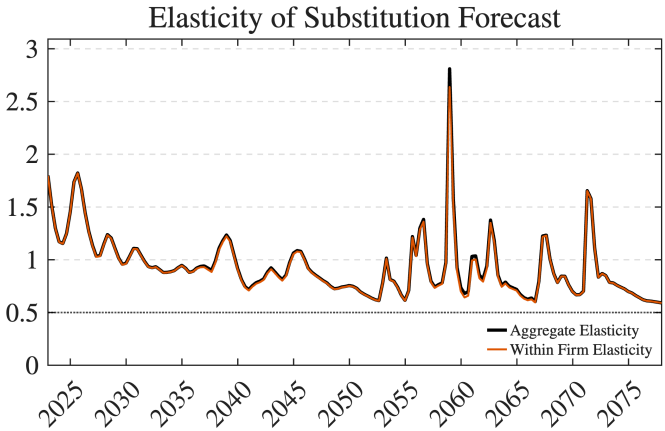


Back

Predicting Future of Computer Vision Automation

- Trends in computing cost r: annual decline of 74.0% in (hardware + algorithmic) $ cost of FLOP

  Del Sozzo, et al. (2025); Ho, et al. (2024)



Back