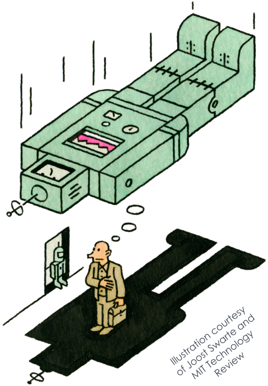
What Happens to Workers at Firms that Automate?

James Bessen, *Boston University*
Maarten Goos, *Utrecht University & Instituut Gak*
Anna Salomons, *Utrecht University & Instituut Gak*
Wiljan van den Berge, *Utrecht University & CPB*

forthcoming in Review of Economics and Statistics

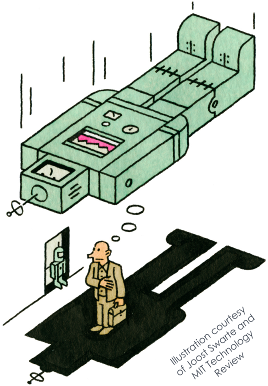
Does automation threaten work?



- Task-based theories of **automation as labor-replacing technology**

  Autor&al.(03), Acemoglu&Autor(11), Acemoglu&Restrepo(18,22)

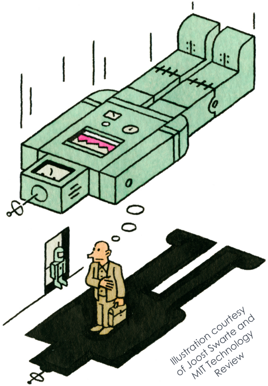
Does automation threaten work?



- Task-based theories of **automation as labor-replacing technology**
  Autor&al.(03), Acemoglu&Autor(11), Acemoglu&Restrepo(18,22)

- Different from previous work assuming **Harrod-, Solow-, or Hicks-neutral** tech progress Uzawa(61), Katz&Murphy(92), Piketty(14), Krusell&al.(00)

Does automation threaten work?



- Task-based theories of **automation as labor-replacing technology** Autor&al.(03), Acemoglu&Autor(11), Acemoglu&Restrepo(18,22)
- Different from previous work assuming **Harrod-, Solow-, or Hicks-neutral** tech progress Uzawa(61), Katz&Murphy(92), Piketty(14), Krusell&al.(00)
- Models of automation more easily predict decreases in labor share and labor demand Restrepo(23), Grossman&Oberfield(22)

Automation and labor markets: emerging evidence overview

- **Macro-level evidence** on aggregate changes in occupations, sectors, labor share, wage inequality Acemoglu&Restrepo(20,22), Boustan&al.(22), Hubmer&Restrepo(21), Autor&al.(20)

Automation and labor markets: emerging evidence

overview

- **Macro-level evidence** on aggregate changes in occupations, sectors, labor share, wage inequality Acemoglu&Restrepo(20,22), Boustan&al.(22), Hubmer&Restrepo(21), Autor&al.(20)
- **Micro-level evidence** on firm-level adoption of robots in manufacturing on firm-level outcomes Aghion&al.(23), Bonfiglioli&etal.(22), Humlum(21), Hirvonen&etal.(22)

Automation and labor markets: emerging evidence overview

- **Macro-level evidence** on aggregate changes in occupations, sectors, labor share, wage inequality Acemoglu&Restrepo(20,22), Boustan&al.(22), Hubmer&Restrepo(21), Autor&al.(20)
- **Micro-level evidence** on firm-level adoption of robots in manufacturing on firm-level outcomes Aghion&al.(23), Bonfiglioli&etal.(22), Humlum(21), Hirvonen&etal.(22)
- Challenges for micro-level evidence of automation on labor demand:
  - measures of **automation beyond robotics**
  - **worker-level** adjustments
  - **credible research design** given larger firms invest more in automation

Contributions

1. Measure of **firm-level automation expenditures** across sectors

Contributions

1. Measure of **firm-level automation expenditures** across sectors
2. Examine **worker-level impacts** of automation

Contributions

1. Measure of **firm-level automation expenditures** across sectors
2. Examine **worker-level impacts** of automation
3. **Event-study DiD design** leveraging the timing of automation events

Contributions

1. Measure of **firm-level automation expenditures** across sectors
2. Examine **worker-level impacts** of automation
3. **Event-study DiD design** leveraging the timing of automation events
4. Compare impacts of **automation versus computerization**

Contributions

1. Measure of **firm-level automation expenditures** across sectors
2. Examine **worker-level impacts** of automation
3. **Event-study DiD design** leveraging the timing of automation events
4. Compare impacts of **automation versus computerization**
5. Ideas for examining **role of worker power**

Outline

Data

Defining and explaining automation events

Stacked DiD estimates of worker-level impacts

Automation versus computerization events

Automation in distorted labor markets

Conclusions

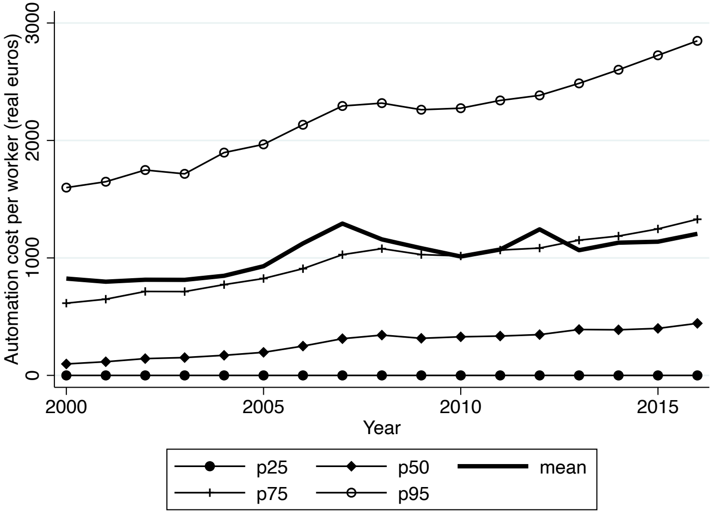
Data from Statistics Netherlands

- Annual **survey of private non-financial firms**, incl. **automation costs**:
  - Described as "expenditures on third-party automation services"
  - Automation expenditures are an official book-keeping entry $\Rightarrow$ well measured
  - Pervasive across time, sectors and firm sizes
  - Correlated with process innovation and automation technologies more
  - Correlated with automation imports more

Data from Statistics Netherlands

- Annual **survey of private non-financial firms**, incl. **automation costs**:
  - Described as "expenditures on third-party automation services"
  - Automation expenditures are an official book-keeping entry $\Rightarrow$ well measured
  - Pervasive across time, sectors and firm sizes
  - Correlated with process innovation and automation technologies more
  - Correlated with automation imports more
- Administrative daily **matched employer-employee records** more
- Years **2000-2016**

Automation costs per worker over time more



Automation occurs in all sectors

Sector	**Mean cost level** Total	Per worker	**Cost share (%)** *Mean*	*SD*	**Nr of obs** *Firms*	*Firms × yrs*
Manufacturing	430,091	1,076	0.36	0.58	5,522	44,393
Construction	78,128	451	0.20	0.36	4,429	28,200
Wholesale & retail trade	116,308	1,177	0.31	0.80	10,903	75,135
Transportation & storage	279,324	907	0.41	1.06	3,125	21,268
Accommodation & food serving	55,714	245	0.30	0.50	1,182	6,535
Information & communication	444,364	1,789	0.85	2.92	2,646	16,929
Prof'l, scientific, & technical activities	150,766	1,285	1.02	1.75	3,935	23,367
Administrative & support activities	133,437	839	0.50	1.19	3,825	22,796

Notes: Automation cost level in 2015 euros, automation cost shares as a percentage of total costs, excluding automation costs. Total firms is N=35,567; Total firms × years is 238,623.

Automation costs by firm size

Firm size class	Total cost	Cost per worker		Cost share (%)		Nr of obs	
	Mean	Mean	SD	Mean	SD	*Firms*	*Firms × yrs*
1-19 employees	12,270	921	14,571	0.4	1.3	9,495	48,052
20-49 employees	27,693	893	4,547	0.42	1.34	13,424	86,540
50-99 employees	61,460	953	4,345	0.42	0.96	6,186	47,038
100-199 employees	144,912	1,135	5,813	0.44	0.94	3,412	28,660
200-499 employees	406,534	1,574	21,314	0.51	1.11	1,941	17,852
$\geq$500 employees	3,161,867	2,124	14,294	0.76	1.6	1,109	10,481

Notes: Automation cost level in 2015 euros, automation cost shares as a percentage of total costs, excluding automation costs. Total firms is N=35,567; Total firms $\times$ years is 238,623.

Outline

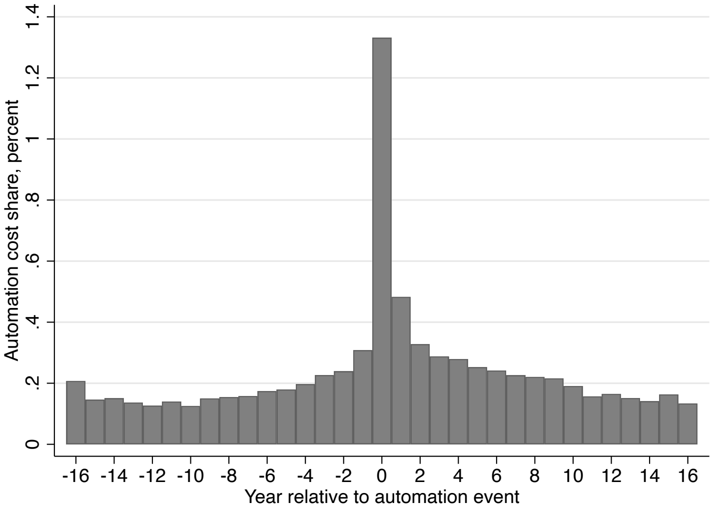
Defining spikes in automation cost shares

- Firms have **spikes in automation cost shares over time**
- Firm j has an automation cost share spike in year τ if:

$$spike_{j\tau} = \mathbb{1}\left\{ \frac{AC_{j\tau}}{\overline{TC}_j} \geq 3 \times \frac{1}{T-1} \sum_{t \neq \tau}^{T} \left(\frac{AC_{jt}}{\overline{TC}_j} \right) \right\}$$

- Of 35K firms, 10K have at least 1 spike, 8K have exactly 1 spike more
- A firm's **first spike is its automation event**

Automation cost shares around automation events



A model to explain automation events model

A model of monopolistic competition with endogenous firm-level automation:

- **Automation**: task-based model in which K directly substitutes for L in tasks (ignoring other types of tech progress) Acemoglu&Restrepo(18,22)
- **Automation events**: automation is fixed and irreversible investment, spikes in automation cost shares within firms over time Haltiwanger(99), Doms&Dunne(98)
- **Product demand shocks** to explain why firms with automation events grow faster than firms without Bonfiglioli&al.(22)

The firm's decision to automate

- If firm j automates, **its output price decreases** to technology frontier:

$$P_{jt} = \begin{cases} P_{jt-1} & \text{if } D_{jt-1} = 0 \\ \mathcal{P}_t & \text{if } D_{jt-1} = 1 \text{ with } \mathcal{P}_t = \mu \mathcal{P}_{t-1} \text{ with } \mu < 1 \end{cases}$$

The firm's decision to automate

- If firm j automates, **its output price decreases** to technology frontier:

$$P_{jt} = \begin{cases} P_{jt-1} & \text{if } D_{jt-1} = 0 \\ \mathcal{P}_t & \text{if } D_{jt-1} = 1 \text{ with } \mathcal{P}_t = \mu \mathcal{P}_{t-1} \text{ with } \mu < 1 \end{cases}$$

- Firm j chooses $D_{j0}, D_{j1}, \ldots$ to **maximize expected net profits**:

$$\max_{D_{j0}, D_{j1}, \ldots} \mathbb{E} \sum_{t=0}^{\infty} \beta^t \left[\sigma^{-1} Y_t \epsilon_{jt}^{\sigma-1} \left[\frac{P_{jt}}{\mathcal{P}_t} \right]^{(1-\sigma)} - D_{jt} F_{jt} \right]$$

with F_{jt} the cost of automation incl. employment adjustment

The firm's decision to automate

- If firm j automates, **its output price decreases** to technology frontier:

$$P_{jt} = \begin{cases} P_{jt-1} & \text{if } D_{jt-1} = 0 \\ \mathcal{P}_t & \text{if } D_{jt-1} = 1 \text{ with } \mathcal{P}_t = \mu \mathcal{P}_{t-1} \text{ with } \mu < 1 \end{cases}$$

- Firm j chooses $D_{j0}, D_{j1}, \dots$ to **maximize expected net profits**:

$$\max_{D_{j0}, D_{j1}, \dots} \mathbb{E} \sum_{t=0}^{\infty} \beta^t \left[\sigma^{-1} Y_t \epsilon_{jt}^{\sigma-1} \left[\frac{P_{jt}}{\mathcal{P}_t} \right]^{(1-\sigma)} - D_{jt} F_{jt} \right]$$

with F_{jt} the cost of automation incl. employment adjustment

- F_{jt} is fixed and irreversible s.t. **spikes in automation cost shares** over time

The impact of automation events on labor demand

- **Unconditional labor demand** is given by:

$$L_{jt} = \left[\frac{\sigma - 1}{\sigma}\right]^{\sigma} Y_t \epsilon_{jt}^{\sigma-1} W_t^{-\sigma} [1 - I_{jt}] \left[\left[\frac{W_t}{R_t}\right]^{I_{jt}} \Psi_H(I_{jt})\right]^{\sigma-1}$$

  with $I_{jt} \in [0, 1]$ share of tasks that are automated

- In $t-1$, the firm chooses $I_{jt} = I_{jt-1}$ or $I_{jt} = \mathcal{I}_t$ which increases over time
- Increase in I_{jt} reduces labor demand if displacement effect $>$ productivity effect

The impact of automation events on labor demand

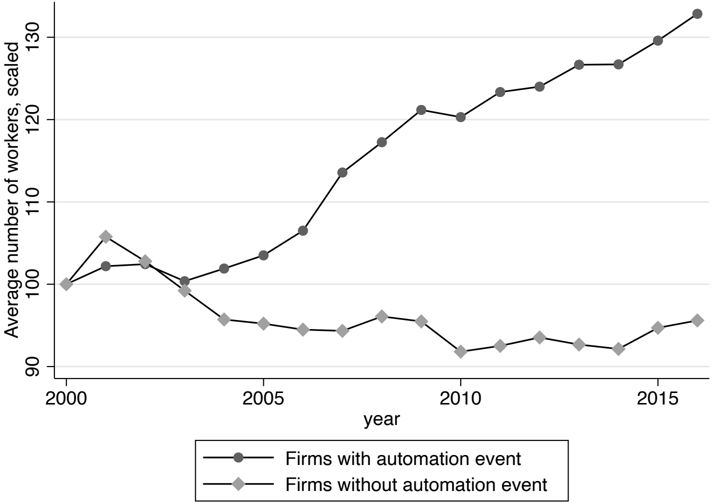
- **Unconditional labor demand** is given by:

$$L_{jt} = \left[\frac{\sigma - 1}{\sigma}\right]^{\sigma} Y_t \epsilon_{jt}^{\sigma-1} W_t^{-\sigma} [1 - I_{jt}] \left[\left[\frac{W_t}{R_t}\right]^{I_{jt}} \Psi_H(I_{jt})\right]^{\sigma-1}$$

  with $I_{jt} \in [0, 1]$ share of tasks that are automated

- In $t-1$, the firm chooses $I_{jt} = I_{jt-1}$ or $I_{jt} = \mathcal{I}_t$ which increases over time
- Increase in I_{jt} reduces labor demand if displacement effect $>$ productivity effect
- **Product demand shock** affects both labor demand and automation

Ever-automators have faster employment growth than never-automators



Identifying assumptions for an event-study DiD design

1. **Parallel trends in post-treatment periods**:
   - Average outcomes for treated would change same as for controls if no treatment
   - Not true when comparing ever-automating with never-automating firms
   - Only use firms with automation events and exploit event timing not incidence

Identifying assumptions for an event-study DiD design

1. **Parallel trends in post-treatment periods**:
   - Average outcomes for treated would change same as for controls if no treatment
   - Not true when comparing ever-automating with never-automating firms
   - Only use firms with automation events and exploit event timing not incidence
2. **No anticipation in pre-treatment periods**:
   - Average outcomes for treated same if no treatment
   - Firms do not invest in automation before an automation event
   - Focus on incumbent workers employed at their firm 3 yrs prior automation

Outline

Stacking the data

- Consider event window of 3 yrs before and 5 yrs after
- For each year $2003 \leq t \leq 2011$, create 9 **group-specific data sets** of workers treated in t and control workers treated at least 5 yrs later

Stacking the data

- Consider event window of 3 yrs before and 5 yrs after
- For each year $2003 \leq t \leq 2011$, create 9 **group-specific data sets** of workers treated in t and control workers treated at least 5 yrs later
- We have excluded forbidden comparisons Borusyak&al.(23), Goodman-Bacon(21) more

Stacking the data

- Consider event window of 3 yrs before and 5 yrs after
- For each year $2003 \leq t \leq 2011$, create 9 **group-specific data sets** of workers treated in t and control workers treated at least 5 yrs later
- We have excluded forbidden comparisons Borusyak&al.(23), Goodman-Bacon(21) more
- Stack 9 group-specific data sets into a single **stacked data set**

Stacking the data

- Consider event window of 3 yrs before and 5 yrs after
- For each year $2003 \leq t \leq 2011$, create 9 **group-specific data sets** of workers treated in t and control workers treated at least 5 yrs later
- We have excluded forbidden comparisons Borusyak&al.(23), Goodman-Bacon(21) more
- Stack 9 group-specific data sets into a single **stacked data set**
- Do this for incumbent workers (with $\geqslant 3$ yrs of tenure at their firm in year $t-1$)

TWFE event-study DiD specification using stacked data

- Using stacked data, regress standard TWFE event-study DiD specification:

$$Y_{i,j,t} = \alpha_i + \alpha_t + \sum_{e=-3}^{-2} \gamma_e^{PRE} D_e \times D_i + \sum_{e=0}^{4} \gamma_e^{POST} D_e \times D_i + \lambda X_{i,j,t} + \varepsilon_{i,j,t}$$

  with α_i individual-group FE and α_t calendar year-group FE

TWFE event-study DiD specification using stacked data

- Using stacked data, regress standard TWFE event-study DiD specification:

$$Y_{i,j,t} = \alpha_i + \alpha_t + \sum_{e=-3}^{-2} \gamma_e^{PRE} D_e \times D_i + \sum_{e=0}^{4} \gamma_e^{POST} D_e \times D_i + \lambda X_{i,j,t} + \varepsilon_{i,j,t}$$

  with α_i individual-group FE and α_t calendar year-group FE

- $\hat{\gamma}_e$ is a variance-weighted average of group-specific ATTs alternatives

TWFE event-study DiD specification using stacked data

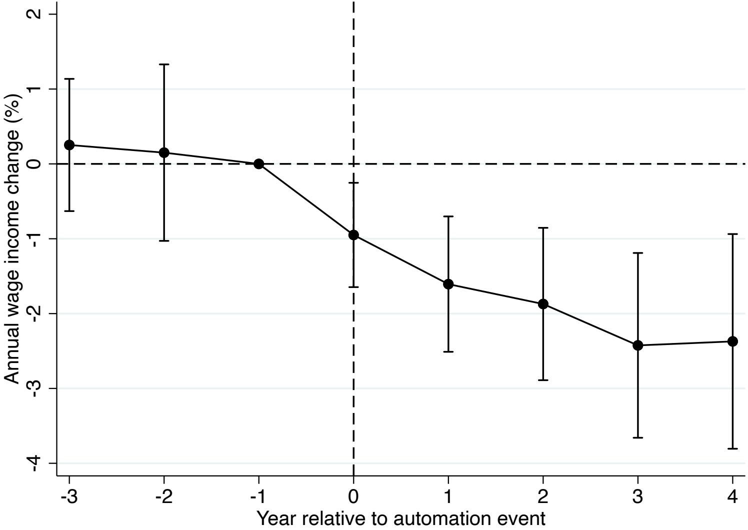
- Using stacked data, regress standard TWFE event-study DiD specification:

$$Y_{i,j,t} = \alpha_i + \alpha_t + \sum_{e=-3}^{-2} \gamma_e^{PRE} D_e \times D_i + \sum_{e=0}^{4} \gamma_e^{POST} D_e \times D_i + \lambda X_{i,j,t} + \varepsilon_{i,j,t}$$

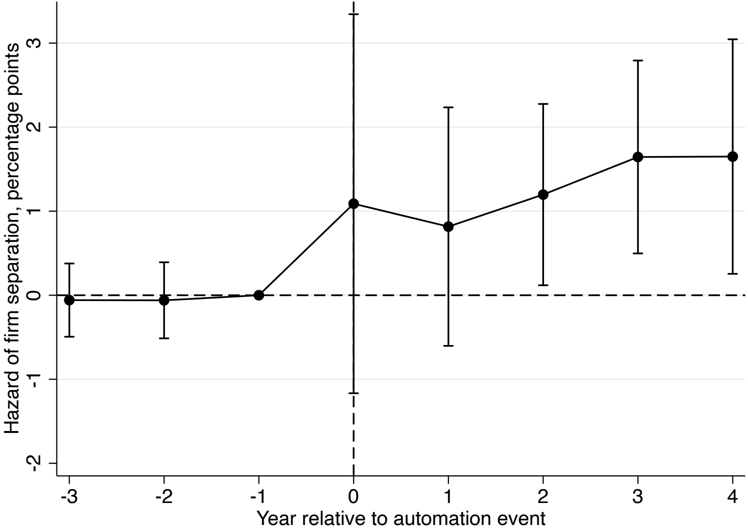
  with α_i individual-group FE and α_t calendar year-group FE

- $\hat{\gamma}_e$ is a variance-weighted average of group-specific ATTs alternatives
- X includes age, age squared (with time-invariant char. absorbed by α_i)
- S.e. are clustered at the treatment-level (i.e. all workers at a firm in $t-1$)

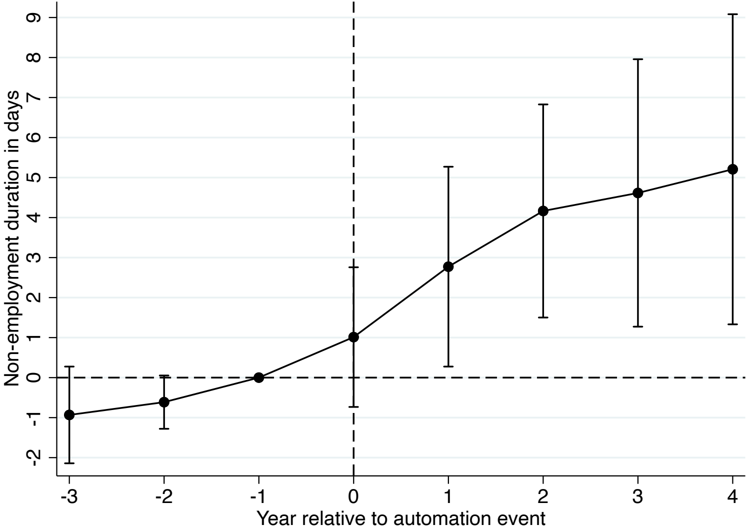
Loss in annual earnings totals 10% of one annual wage after 5 yrs



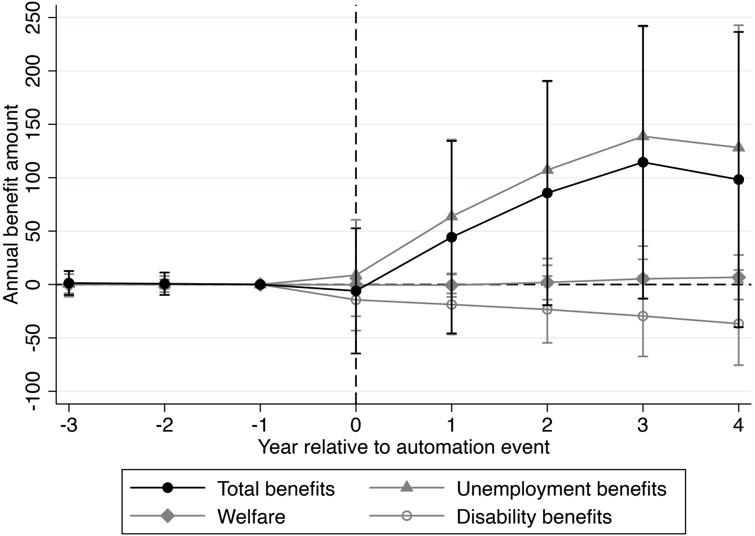
Hazard of leaving the firm increases by a total of 6.5ppt after 5 yrs



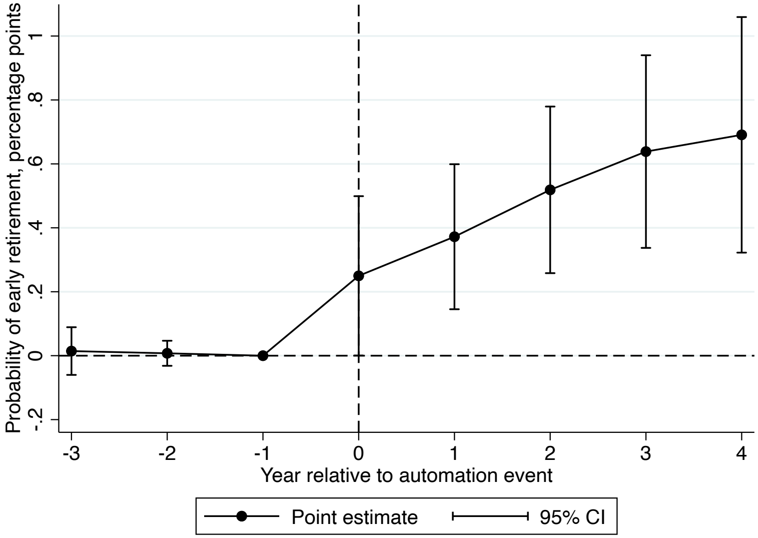
Annual days in non-emp. increase by a total of 18 days after 5 yrs



Annual income from unemployment benefits increases



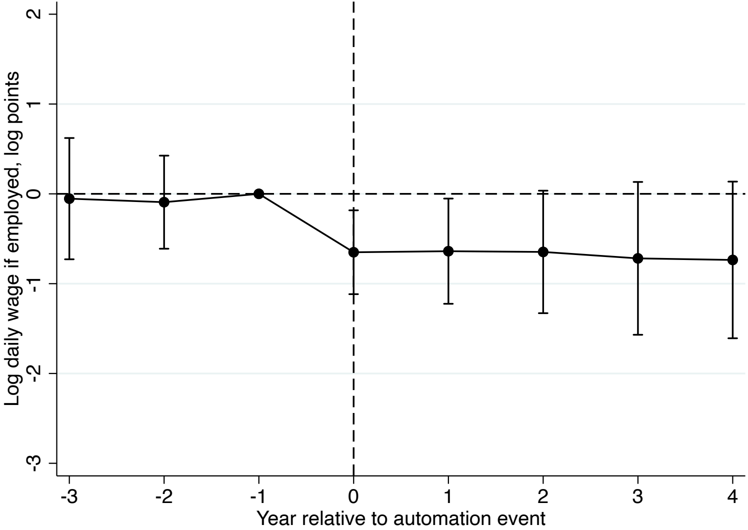
Probability of early retirement increases by a total of 2.5ppt after 5 yrs



A summary of findings

	All workers	Displaced workers
A firm that automates later	100	
A firm that automates after 5 years		
Stay	94	
Displaced	6	
New job		3
Unemployed		1
Early retirement		2

Little effect on log daily wage if employed



Effect heterogeneity

Annual earnings losses are:

1. Pervasive across sectors estimates
2. Larger for workers at smaller firms estimates
3. Larger for older workers estimates
4. Larger for less-educated workers estimates
5. Similar for men and women estimates

Additional analyses

- Other measures of employment (firm-level employment, new hires) more
- Placebo events (investment in other material fixed assets) more
- Robustness tests (spikes, model specification, other firm-level events) more
- Clustering, FRTs, and random treatment timing more

Outline

Data

Defining and explaining automation events

Stacked DiD estimates of worker-level impacts

Automation versus computerization events

Automation in distorted labor markets

Conclusions

Computerization is less likely to decrease labor demand

- Tech progress can also be capital-augmenting Piketty(14), Karabarbounis&Neiman(14)

Computerization is less likely to decrease labor demand

- Tech progress can also be capital-augmenting Piketty(14), Karabarbounis&Neiman(14)
- If $F(\Psi_K K, L)$ with factors paid their marginal products and CRS:

$$\frac{d\ln(W)}{d\ln(\Psi_K)} = \frac{s^K}{\sigma_{KL}} > 0 \qquad \frac{d\ln(s_L)}{d\ln(\Psi_K)} = s^K\left[\frac{1}{\sigma_{KL}} - 1\right]$$

such that computerization must increase labor demand

Computerization is less likely to decrease labor demand

- Tech progress can also be capital-augmenting Piketty(14), Karabarbounis&Neiman(14)
- If $F(\Psi_K K, L)$ with factors paid their marginal products and CRS:

$$\frac{d\ln(W)}{d\ln(\Psi_K)} = \frac{s^K}{\sigma_{KL}} > 0 \qquad \frac{d\ln(s_L)}{d\ln(\Psi_K)} = s^K\left[\frac{1}{\sigma_{KL}} - 1\right]$$

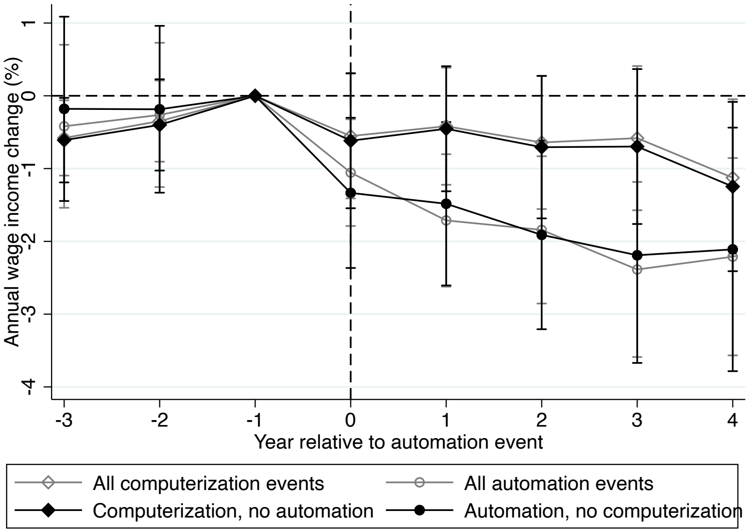
  such that computerization must increase labor demand
- However, a model of capital-skill complementarity: Krusell&al.(00)

$$\frac{dW_S}{d\Psi_K} > 0 \qquad \frac{dW_U}{d\Psi_K} < 0$$

  such that computerization could decrease labor demand for some workers

Computerization versus automation

more



Outline

Data

Defining and explaining automation events

Stacked DiD estimates of worker-level impacts

Automation versus computerization events

Automation in distorted labor markets

Conclusions

Automation in distorted labor markets

- Results consistent with competitive labor markets:

  1. **Automation $\Rightarrow$ marginal product of labor** $\downarrow$ because it displaces workers more than it increases allocative efficiency

  2. **Marginal product of labor** $\downarrow \Rightarrow L \downarrow$ **or** $W \downarrow$ because workers lack the power to benefit from increased allocative efficiency if labor markets are competitive

Automation in distorted labor markets

- Results consistent with competitive labor markets:

  1. **Automation $\Rightarrow$ marginal product of labor $\downarrow$** because it displaces workers more than it increases allocative efficiency

  2. **Marginal product of labor $\downarrow$ $\Rightarrow$ $L \downarrow$ or $W \downarrow$** because workers lack the power to benefit from increased allocative efficiency if labor markets are competitive

- If **workers have wage bargaining power**, the impact of automation on labor demand and welfare may be different model

- Merging collective agreements since 2000 into CBS data

Outline

Conclusions

1. **Automation leads to displacement** for incumbent workers

2. Annual earnings $\downarrow$ $\Rightarrow$ firm separation $\uparrow$ $\Rightarrow$ non-employment $\uparrow$ $\Rightarrow$ unemployment $+$ early retirement $\uparrow$

Conclusions

1. **Automation leads to displacement** for incumbent workers
2. Annual earnings $\downarrow$ $\Rightarrow$ firm separation $\uparrow$ $\Rightarrow$ non-employment $\uparrow$ $\Rightarrow$ unemployment + early retirement $\uparrow$
3. Effects are pervasive across sectors and larger for workers in smaller firms, older workers, less-educated workers

Conclusions

1. **Automation leads to displacement** for incumbent workers
2. Annual earnings $\downarrow \Rightarrow$ firm separation $\uparrow \Rightarrow$ non-employment $\uparrow \Rightarrow$ unemployment $+$ early retirement $\uparrow$
3. Effects are pervasive across sectors and larger for workers in smaller firms, older workers, less-educated workers
4. Automation appears to be **more labor-displacing than computerization**

Conclusions

1. **Automation leads to displacement** for incumbent workers
2. Annual earnings $\downarrow \Rightarrow$ firm separation $\uparrow \Rightarrow$ non-employment $\uparrow \Rightarrow$ unemployment $+$ early retirement $\uparrow$
3. Effects are pervasive across sectors and larger for workers in smaller firms, older workers, less-educated workers
4. Automation appears to be **more labor-displacing than computerization**
5. Impact of automation may depend on **role of worker power**

Thank you!

Appendices

Appendix: New literature on automation

Automation and... back

- **the changing labor share**

  Acemoglu&Restrepo'20,'22; Graetz&Michaels'18; Boustan&al'22; Kogan&al'21; Hubmer&Restrepo'21; Autor&al'20; Kehrig&Vincent'20
- **the changing occupational structure**

  Autor&al'03; Goos&Manning'07; Goos&al'14; Webb'20; Kogan&al'21; Autor&al'22; Acemoglu&al'22; Dillinder&Forsythe'23
- **firm-level outcomes**

  Acemoglu&al'20;Koch&al'21;Humlum'21;Bonfiglioli&al'22;Acemoglu&al'23; Cheng&al'21;Dinlersoz&Wolf'23; Acemoglu&al'22;Aghion&al'23;Hirvonen&al'22
- **exposed workers**

  Cortes'16, Kogan&al'21; Feigenbaum&Gross'20; Acemoglu&Restrepo'20; Boustan&al'22; Mann&Puttmann'23, Coelli'19; Acemoglu&Autor'11; Acemoglu&Restrepo'22; Webb'20

Appendix: Automation costs and innovation

Automation costs and type of innovation

	Dependent variable: Standardized automation cost share
Process innovations	0.203***
	(0.048)
Product innovations	0.098**
	(0.036)
Organizational innovations	0.099*
	(0.041)
N	7,160

Notes: Automation cost shares as a percentage of total costs, excluding automation costs. Model controls for one-digit industry fixed effects and the log number of workers at the firm, and is weighted by survey weights.

Automation costs and technology usage

back

Dependent variable: Standardized automation cost share			
Use of electronic data suited to automated processing	0.236***	Received orders for goods or services through EDI	0.106**
	(0.053)		(0.0339)
N	4,313	Ordered through Electronic Data Interchange (EDI)	-0.099**
			(0.032)
CRM, inventory and distribution analysis	0.200***	N	14,172
	(0.041)		
Customer Relationship Mngmnt (CRM), customer analysis	0.055	Sales software	0.088**
	(0.048)		(0.030)
N	11,927	Purchasing software	0.006
			(0.03)
Enterprise Resource Planning (ERP) software	0.164***	N	7,831
	(0.027)		
N	12,535	Radio Frequency Identification (RFID)	0.056
			(0.083)
Automated records used for value chain integration	0.200**	N	4,149
	(0.066)		
Value chain integration	-0.008	Local Area Network (LAN)	0.015
	(0.047)		(0.026)
N	7,879	N	7,653
Big data analysis	0.127*	Internet for financial transactions	0.016
	(0.054)		(0.025)
N	4,680	N	7,526
Cloud-services: Software for customer information mngmnt	0.168*	Internet for training and education (incl. e-learning)	0.035
	(0.084)		(0.031)
Cloud-services: Software for accounting and financial mngmnt	0.136*	N	8,385
	(0.062)		
N	6,711		

Appendix: Automation costs and automation imports

Comparing automation costs to automation imports by sector

Sector	**Mean share in total costs** Automation costs	Imports	Net imports
Manufacturing	0.346	0.081	0.043
Construction	0.193	0.001	0.001
Wholesale & retail trade	0.300	0.058	0.051
Transportation & storage	0.353	0.134	0.095
Accommodation & food serving	0.268	0.000	0.000
Information & communication	0.804	0.004	0.004
Prof'l, scientific & technical activities	1.006	0.007	0.005
Administrative & support activities	0.437	0.003	0.003

Notes: Total N firms is 30,267. Net automation imports are defined as imports minus re-exports. Total costs include automation costs.

Comparing automation costs to automation imports at the firm level

	Dependent variable: Automation costs (IHS)			
	(1)	(2)	(3)	(4)
Automation imports (IHS)	0.0178**	0.0177**	-0.001	-0.002
	(0.007)	(0.007)	(0.004)	(0.004)
	(5)	(6)	(7)	(8)
Net automation imports (IHS)	0.0158*	0.0157*	-0.003	-0.003
	(0.006)	(0.006)	(0.004)	(0.004)
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Log total costs	Yes	Yes	Yes	Yes

Notes: N=110,698 (firm-year). Automation costs, imports, and net imports are transformed using the inverse hyperbolic sine (IHS). Net automation imports are defined as imports minus re-exports. All models control for log total costs at the firm-year level. Standard errors are clustered at the firm-level.

Importers are much larger than firms with automation events

Dependent variable: Log firm-level number of employees				
	Automation cost spike		Automation imports	
	(1)	(2)	(3)	(4)
Automating	0.078***	0.085***	0.857***	0.838***
	(0.013)	(0.013)	(0.022)	(0.022)
Sector fixed effects	No	Yes	No	Yes

Notes: N = 30,267 firm-level observations. Automation imports measured as non-zero mean automation imports at the firm level. Sector fixed effects are two-digit sector dummies. *p<0.10, **p<0.05, ***p<0.01.

Between firms: automation events and automation importer correlation

	Dependent variable: Dummy for firm having an automation cost spike			
	(1)	(2)	(3)	(4)
Importer	0.022*	0.028**		
	(0.010)	(0.011)		
Net importer			0.022*	0.028**
			(0.010)	(0.011)
Controls	No	Yes	No	Yes

Notes: N = 30,267 firm observations, where 31% of firms have automation cost spikes, and 8.2% (7.9%) have non-zero (net) imports. Controls are log total costs and sector fixed effects. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

Within firms: automation events and automation importers back

Dependent variable: Dummy for firm having an automation cost spike				
	(1)	(2)	(3)	(4)
Importer	0.005	0.002	0.003	0.000
	(0.005)	(0.005)	(0.005)	(0.005)
	(5)	(6)	(7)	(8)
Net importer	0.003	0.000	0.001	-0.001
	(0.005)	(0.005)	(0.005)	(0.005)
Firm fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	No	Yes
Log total costs	No	No	Yes	Yes

Notes: N = 110,698 firm-year observations. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

Appendix: Data cleaning

Data cleaning back

We remove the following observations:

- Workers enrolled in full-time studies earning either less than EUR 5K annually or EUR 10 daily on average across the year
- Workers with earnings above EUR 500K annually or EUR 2K daily on average across the year
- Later, we further exclude workers at firms that have:
  - Not a single spike in automation cost shares
  - No event window (7 yrs of consecutive data)
  - Other events in the event window (mergers, takeovers, splits, restructuring)
  - Large ($>$90%) annual employment changes in the event window or also outside the event window

Appendix: Descriptive statistics on automation costs

Distribution of automation costs back

	All observations			Automation costs > 0		
	Cost level	Cost per worker	Cost share (%)	Cost level	Cost per worker	Cost share (%)
p5	0	0	0	2,211	59	0.04
p10	0	0	0	3,987	101	0.06
p25	0	0	0	10,487	256	0.14
p50	11,736	283	0.16	30,000	641	0.32
p75	52,824	986	0.47	93,711	1,447	0.68
p90	192,393	2,256	1.06	305,111	2,949	1.37
p95	453,172	3,625	1.69	713,121	4,590	2.13
mean	*211,326*	*1,045*	*0.44*	*307,840*	*1,522*	*0.64*
N firms $\times$ years		238,623			163,810	
N with 0 costs		31%			0%	

Appendix: Automation cost spike frequences

Automation cost spike frequencies

back

Spike frequency	N firms	% of N firms
0	25,145	70.7
1	8,351	23.5
2	1,772	5.0
3	266	0.7
4	29	0.1
5	4	0.0
Total	35,567	100

Notes: Spike frequency is defined as the total number of spikes occurring over 2000-2016. The total number of firms is 35,567 and the total number of firms with at least one automation cost share spike is 10,422.

Appendix: A model of monopolistic competition with endogenous automation

Consumption and product demand

- Utility is given by:

$$U(Y_1, ..., Y_J) = \left[\sum_{j=1}^{J} [\epsilon_j Y_j]^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \text{ such that } \sum_{j=1}^{J} P_j Y_j = PY$$

  where $\sigma > 1$

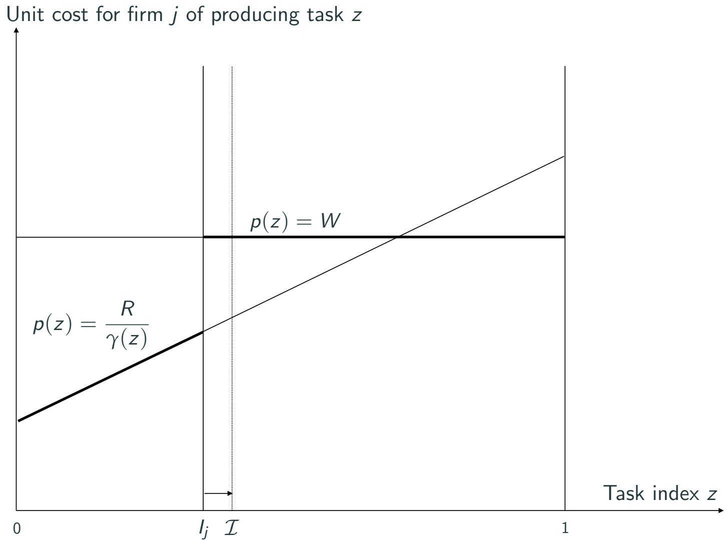
- The ideal price index given by:

$$P(P_1, ..., P_J) \equiv \left[\sum_{j=1}^{J} [P_j/\epsilon_j]^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = 1$$

- Demand for firm j is given by:

$$Y_j = Y\epsilon_j^{\sigma-1} P_j^{-\sigma}$$

Firm-level allocation of capital and labor across tasks in production



Factor bills, prices, output, and profits back

- Conditional factor demands are given by:

$$RK_j = l_j \frac{\sigma - 1}{\sigma} P_j Y_j \quad \text{and} \quad WL_j = [1 - l_j] \frac{\sigma - 1}{\sigma} P_j Y_j$$

- The (relative) output price is given by:

$$P_j = \frac{\sigma}{\sigma - 1} \frac{W^{1-l_j} R^{l_j}}{\Psi_H(l_j)}$$

- Output is given by:

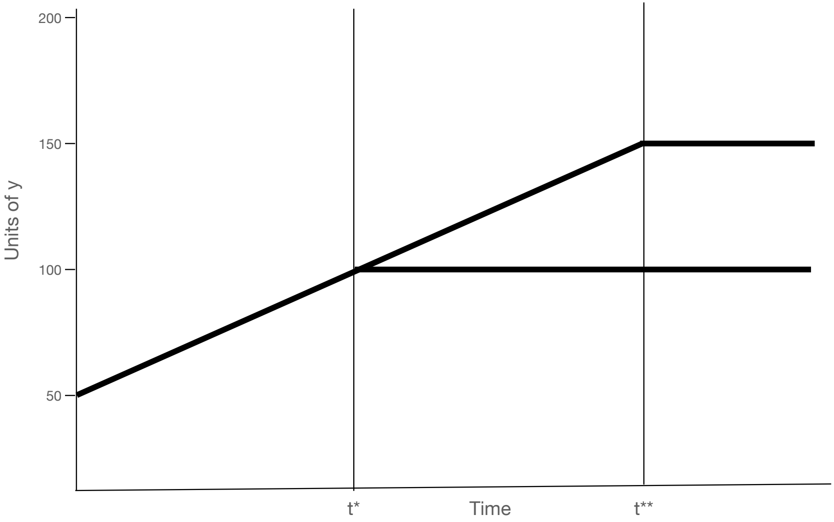
$$Y_j = Y \epsilon_j^{\sigma-1} P_j^{-\sigma} = Y \epsilon_j^{\sigma-1} \left[\frac{\sigma}{\sigma - 1} \frac{W^{1-l_j} R^{l_j}}{\Psi_H(l_j)} \right]^{-\sigma}$$

- Profits are given by:

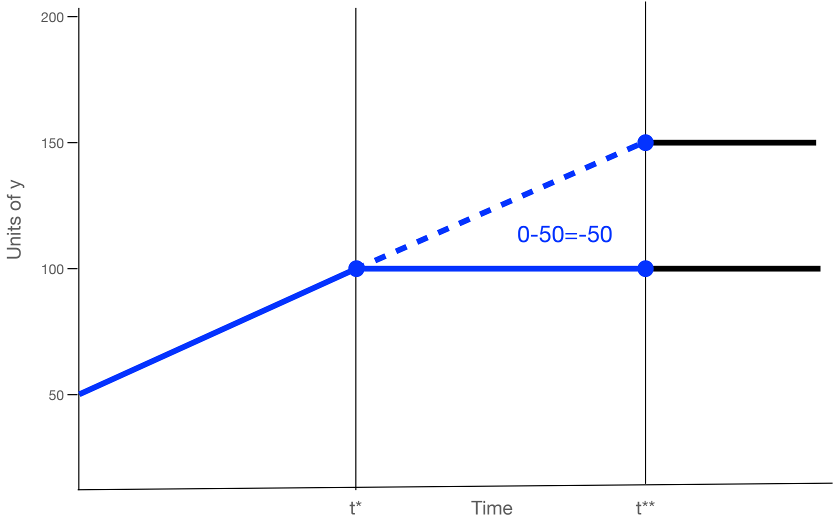
$$\Pi_j = \frac{P_j Y_j}{\sigma}$$

Appendix: Forbidden comparisons

Time-varying homogeneous effects

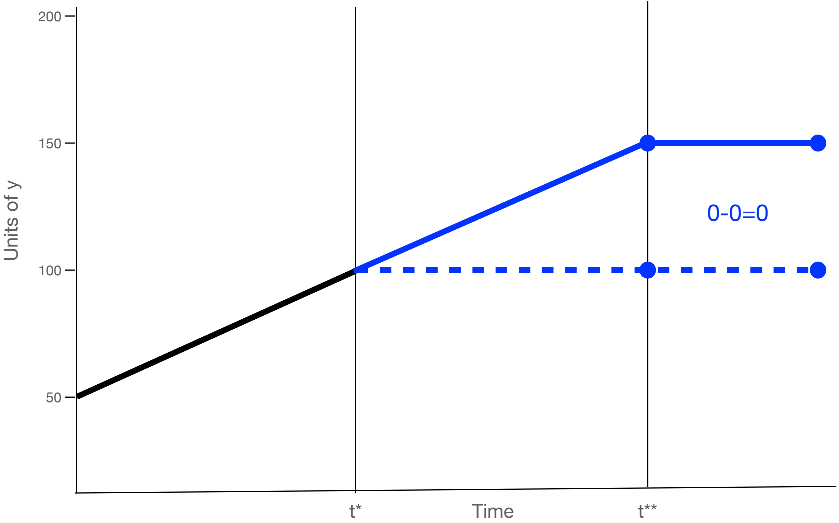


Good comparisons



Forbidden comparisons

back



Appendix: Other estimators for staggered DiD designs

Other estimators for staggered DiD designs

1. **Callaway&Sant'Anna** (`csdid`): doubly robust estimator, flexible aggregation, covariates, bootstrapping, simultaneous CI
2. **Sun&Abraham** (`eventstudyinteract`): 3-step estimator, Interaction-Weighted regression, event-studies
3. **Chaisemartin&D'Haultfoeuille** (`did_multiplegt_dyn`): Wald-TC estimator of treatment effects on switchers, instantaneous treatment effects, non-staggered designs, multi-valued treatments
4. **Roth&Sant'Anna** (`staggered`): general DiD/DiM plugin estimator, efficient estimator if treatment timing is random
5. **Borusyak&al.** (`did_imputation`): 3-step imputation estimator (`event_plot` for plotting event-study graphs)

For an overview, go to https://asjadnaqvi.github.io/DiD/

Other estimators for staggered DiD designs back

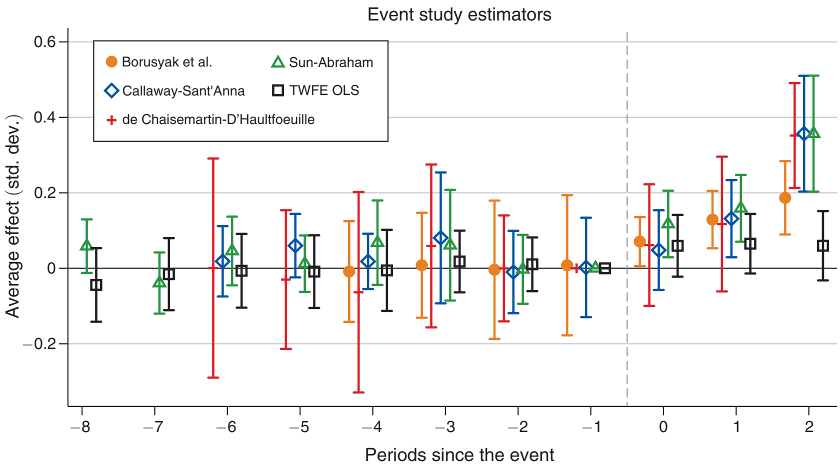


FIGURE 2. EFFECTS OF FACEBOOK ON THE INDEX OF POOR MENTAL HEALTH BASED ON DISTANCE TO/FROM FACEBOOK INTRODUCTION

Braghieri&al.'22, "Social Media and Mental Health", *American Economic Review* 2022, 112(11)

Appendix: Matching details

CEM statistics

back

- Coarsened Exact Matching (CEM):
  1. In each of the three pre-treatment years, separate strata for each 5 percentiles of annual wage + separate bins for the 99th and 99.5th percentiles
  2. One year prior to treatment, matched workers must be observed in the same calendar year and work in the same sector
- 30,247 strata
- 98% of treated incumbents are matched; and 93% of control group incumbents are assigned a non-zero weight

Appendix: Effect heterogeneity

Heterogeneity by sector, contract type, gender and wages

back

(1) Sector	
Manufacturing (reference)	-1.61*
	(0.83)
Deviations from reference group for:	
Construction	0.16
	(1.49)
Wholesale & retail trade	-0.69
	(1.14)
Transportation & storage	1.40
	(1.50)
Accommodation & food serving	2.88**
	(1.43)
Information and communication	-0.87
	(1.55)
Prof'l, scientific, & technical activities	-1.19
	(1.55)
Administrative & support activities	-1.08
	(2.45)

(2) Gender	
Male (reference)	-1.52***
	(0.56)
Deviation from reference group for:	
Female	-0.94
	(0.74)

(3) Contract type	
Open-ended contract (reference)	-1.75***
	(0.44)
Deviation from reference group for:	
Flexible contract	-2.12
	(3.15)

(4) Overall age-specific wage quartile	
Bottom quartile (reference)	-2.12*
	(1.25)
Deviations from reference group for:	
Second quartile	-0.03
	(1.21)
Third quartile	0.49
	(1.24)
Top quartile	0.17
	(1.47)

(5) Within-firm age-specific wage quartile	
Bottom quartile (reference)	-1.44
	(1.78)
Deviations from reference group for:	
Second quartile	-0.77
	(2.13)
Third quartile	-0.96
	(2.23)
Top quartile	-0.19
	(1.77)

Heterogeneity by firm size, age and education level back

A. Firm size		B. Worker age	
1–19 employees (reference)	-3.16***	Age ≥50 (reference)	-3.96***
	(0.76)		(1.25)
Deviations from reference group for:			
20–49 employees	0.22	Age 40–49	2.63*
	(0.91)		(1.36)
50–99 employees	2.39**	Age 30–39	2.27*
	(0.96)		(1.27)
100–199 employees	1.33	Age 20–29	3.13*
	(1.11)		(1.71)
200–499 employees	2.25*		
	(1.16)		
≥500 employees	0.76		
	(1.51)		
N	8,792,616		8,022,952

C. Worker education level	
Medium education (reference)	-2.60***
	(0.77)
Deviations from reference group for:	
Low education	0.92
	(1.48)
High education	1.32*
	(0.70)
N	2,178,168

Appendix: Other measures of employment

Incumbents versus recent hires and firm-level employment

- Incumbents leave because **firms lower their long-run optimal level of employment** after automation

  $\Rightarrow$ net decrease in **firm-level employment**

  $\Rightarrow$ adverse impacts on annual wage income for **recent hires**

Incumbents versus recent hires and firm-level employment

- Incumbents leave because **firms lower their long-run optimal level of employment** after automation

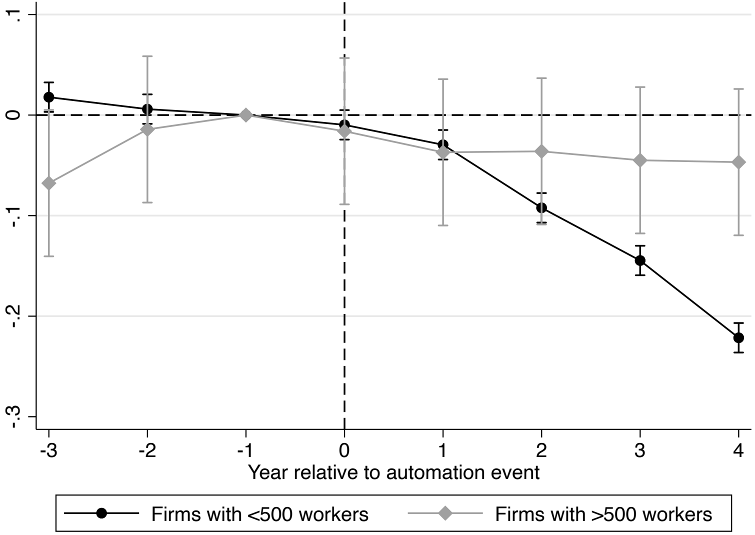
  $\Rightarrow$ net decrease in **firm-level employment**

  $\Rightarrow$ adverse impacts on annual wage income for **recent hires**

- Adverse **effects can be different** if firms foresee shocks (even if common) of expected cost in hiring when labor demand rebounds

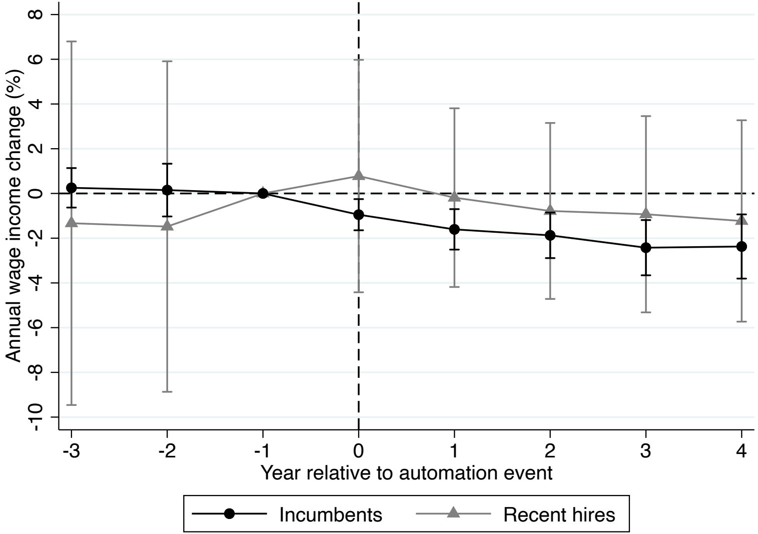
  e.g. effects of automation in large firms muted if they have stronger employment trend growth so will want to hire more workers in the future

Estimates for firm-level employment (%)



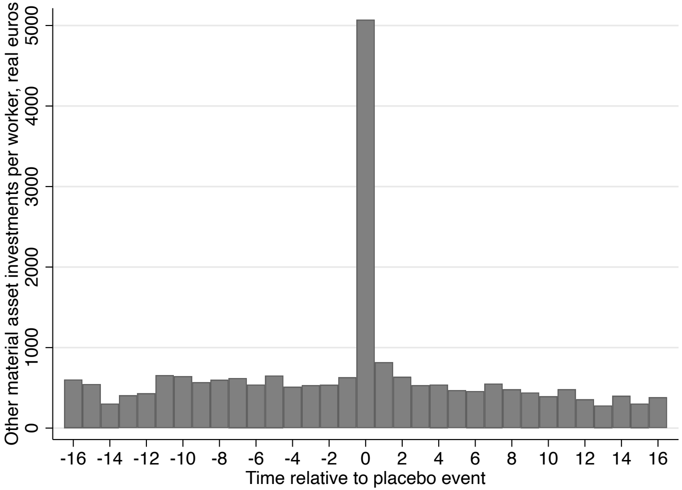
Incumbents versus recent hires

back

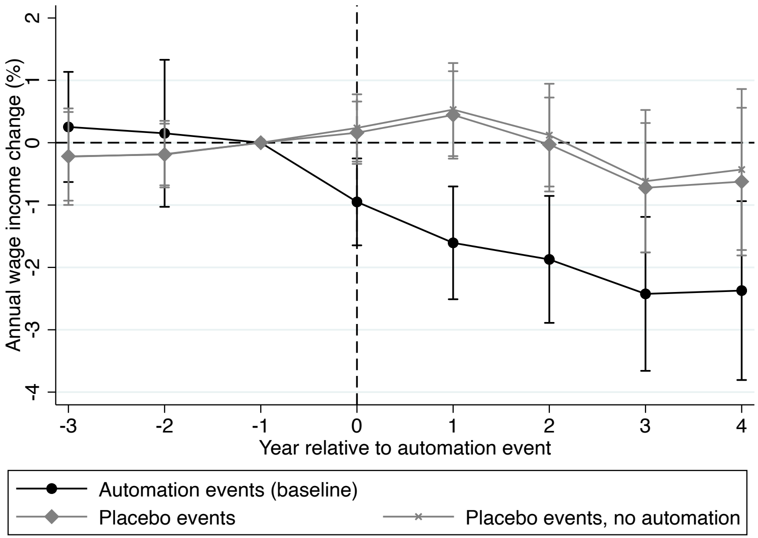


Appendix: Placebo events

Spikes in other material fixed assets



Automation versus other material fixed assets back



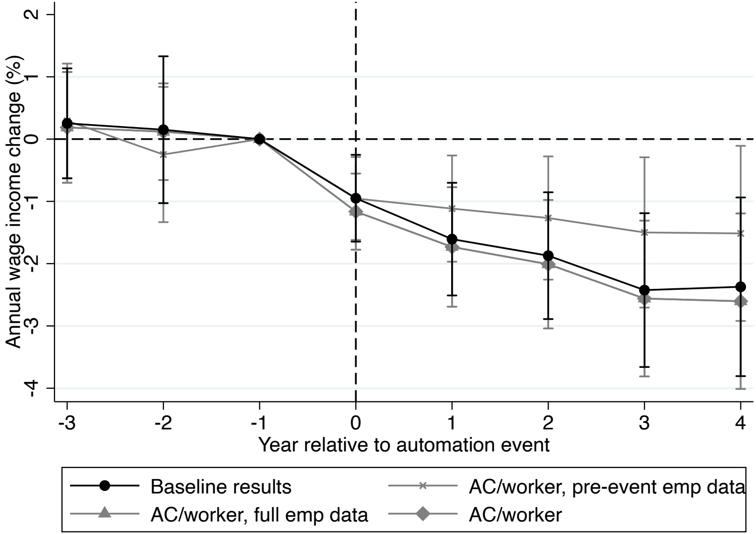
Appendix: Robustness tests

Robustness tests back

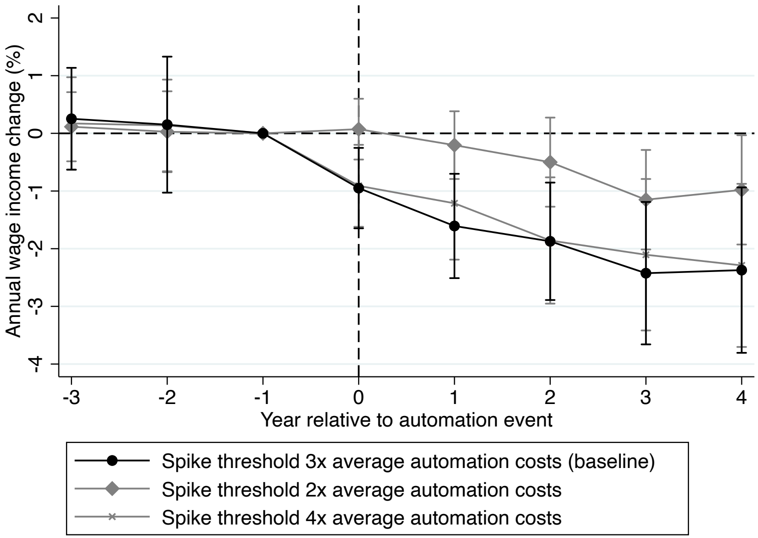
Results for annual earnings (and other worker outcomes) are robust to:

1. Different spike definitions
2. Different spike sizes
3. Different model specifications
4. Eliminating other firm-level events

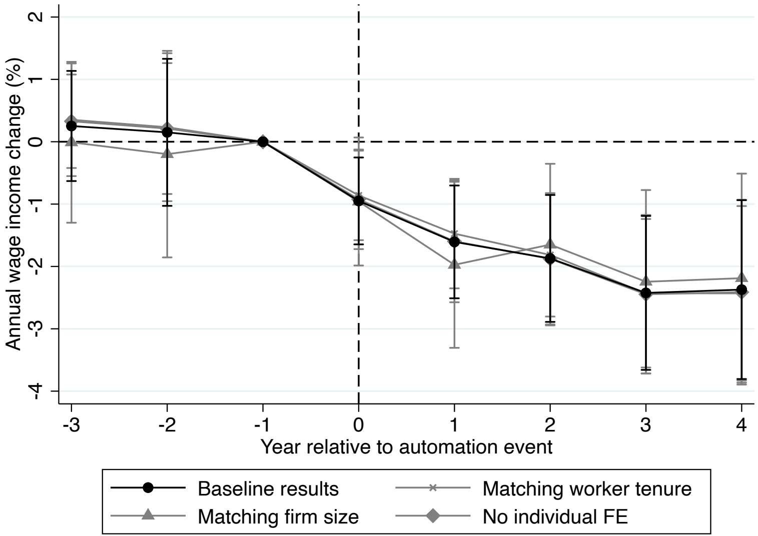
Robustness to spike definition



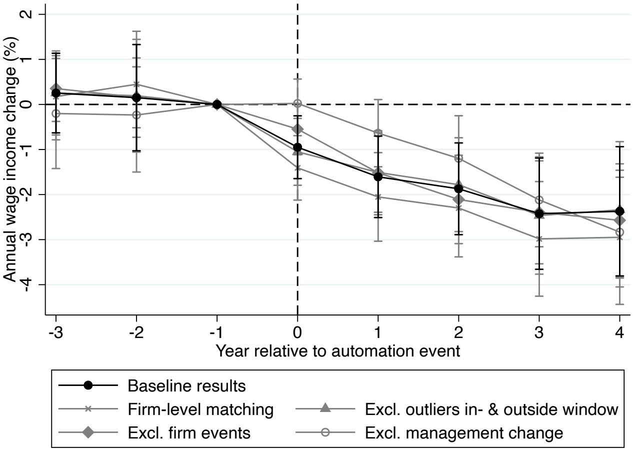
Robustness to spike size



Robustness to model specification



Eliminating other firm-level events



Appendix: Clustering, FRTs, and random treatment timing

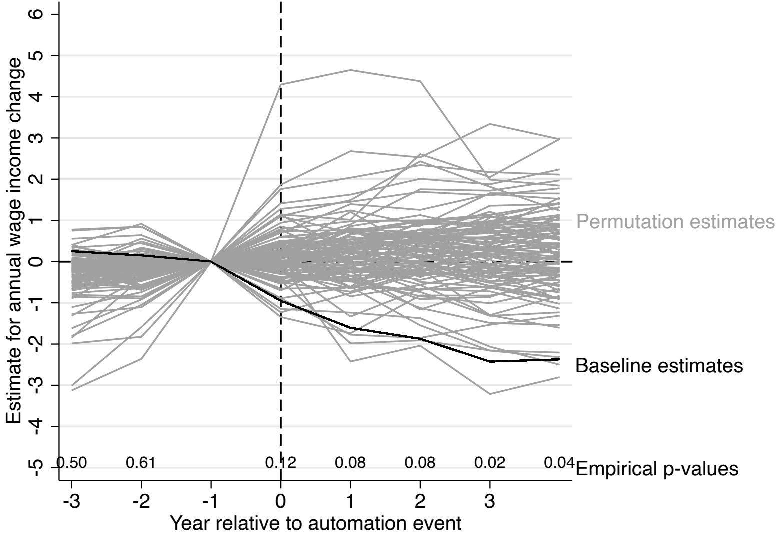
Design-based clustering and random automation

- S.e. are clustered at the **treatment level**
- Alternative for inference is **Fischer Randomization Test (FRT)** which plots permutation estimates after randomly assigning treatment
- FRT is test of the **null hypothesis that all ATTs are 0**

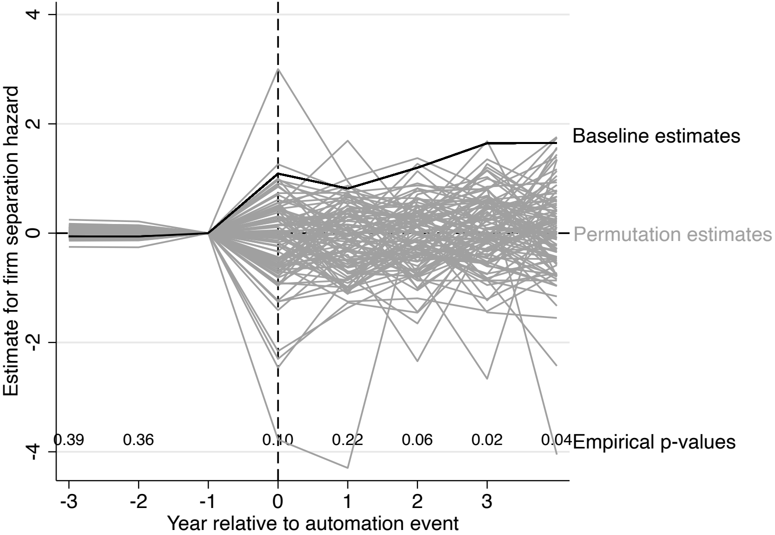
Design-based clustering and random automation

- S.e. are clustered at the **treatment level**
- Alternative for inference is **Fischer Randomization Test (FRT)** which plots permutation estimates after randomly assigning treatment
- FRT is test of the **null hypothesis that all ATTs are 0**
- FRT (implicitly) imposes **treatment timing is random**
- If treatment timing trully random, use **other more efficient estimators**

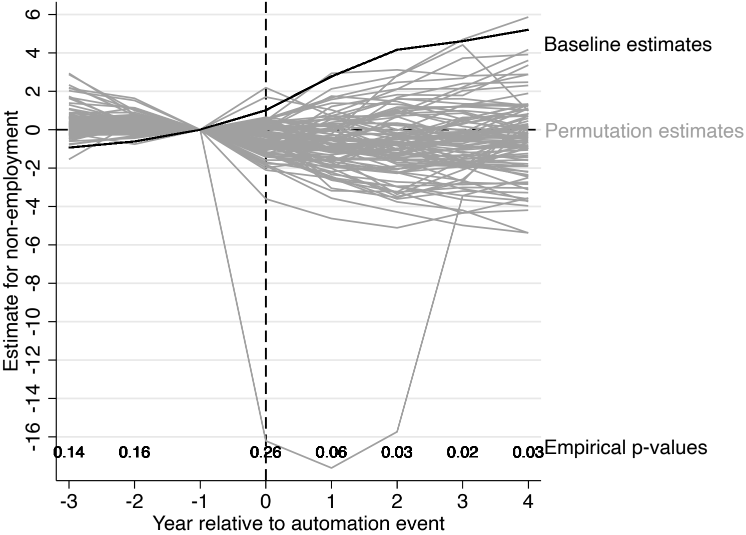
Fischer Randomization Test: Annual wage income



Fischer Randomization Test: Firm separation

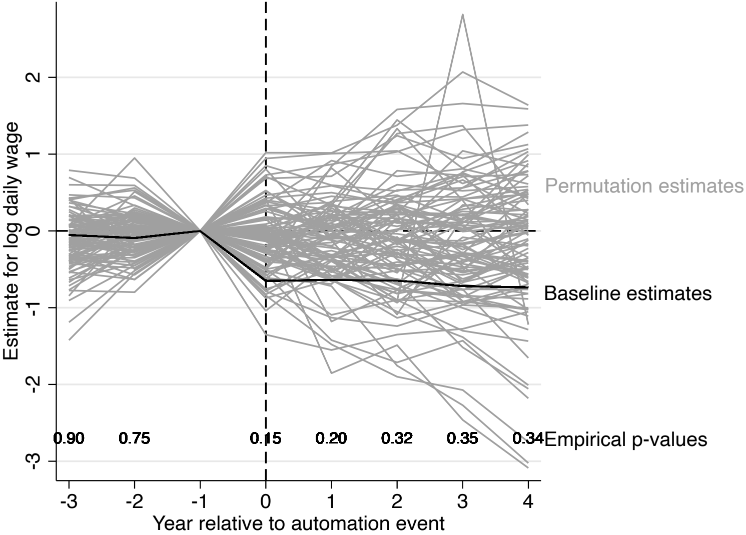


Fischer Randomization Test: Annual days in non-employment



Fischer Randomization Test: Daily wages

back



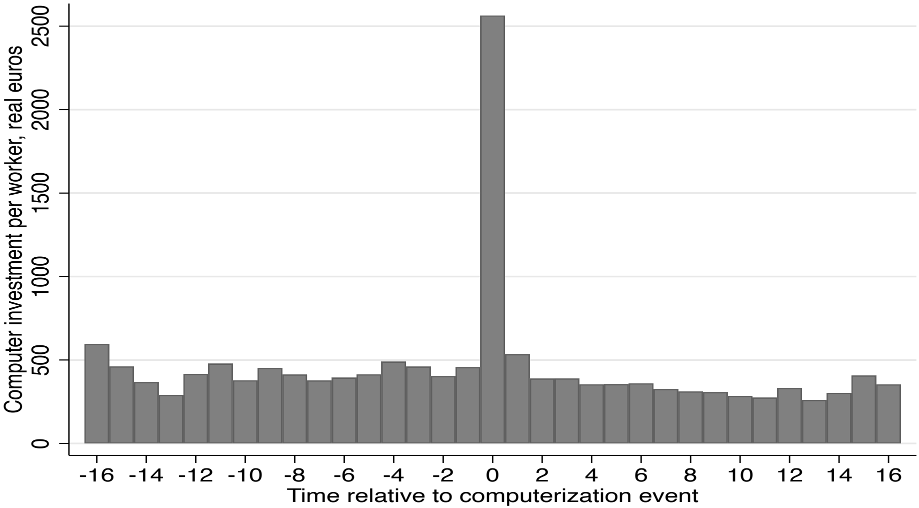
Appendix: Computer investments

Spike frequencies, overlapping sample

	Percentage of firms with event type:	
Nr of spikes	Automation	Computerization
0	71.9	47.9
1	22.5	41.9
2	4.8	9.1
3	0.7	1.1
4	0.1	0.1

Notes: Overlapping sample of firms. N=25,107.

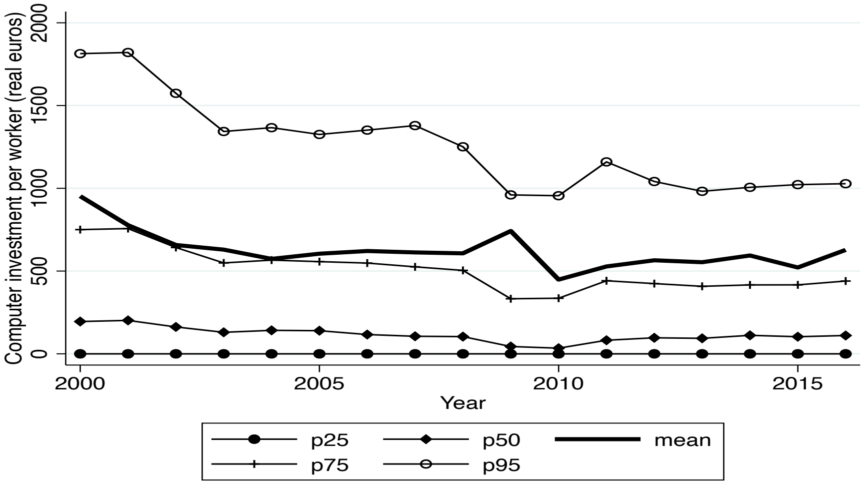
Computer investment spikes



Summary statistics on overlapping sample

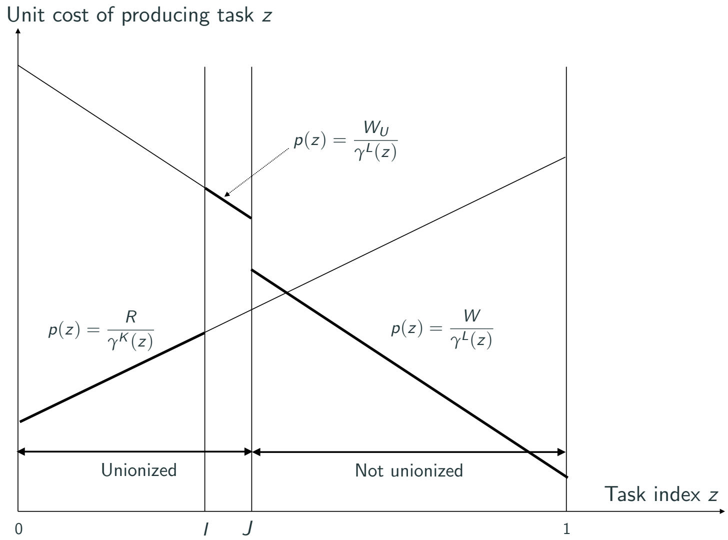
	Automation cost		**Computer investment**	
	level	*per worker*	*level*	*per worker*
p5	0	0	0	0
p10	0	0	0	0
p25	0	0	0	0
p50	18,285	324	6,046	108
p75	75,758	1,043	33,892	488
p90	263,000	2,372	123,065	1,229
p95	620,508	3,837	273,263	2,040
mean	271,929	1,125	109,415	615
mean excl. zeros	378,036	1,564	170,846	960
N firms $\times$ yrs	171,797		171,797	
N firms $\times$ yrs with 0 costs	48,220		61,773	

Computer investment per worker over time back



Appendix: A model of automation with wage bargaining

Automation of union jobs



Assumptions and equilibrium output

- Tasks 0 to I are produced with K, and tasks I to 1 with L
- In union tasks I to J, workers receive a union wage premium
- Capital K and labor L are supplied inelastically
- If tasks are combined Cobb-Douglas, equilibrium output can be written as:

$$Y = \Psi_H(I) \left[\frac{K}{I}\right]^I \left[\frac{L_U}{J-I}\right]^{J-I} \left[\frac{L-L_U}{1-J}\right]^{1-J}$$

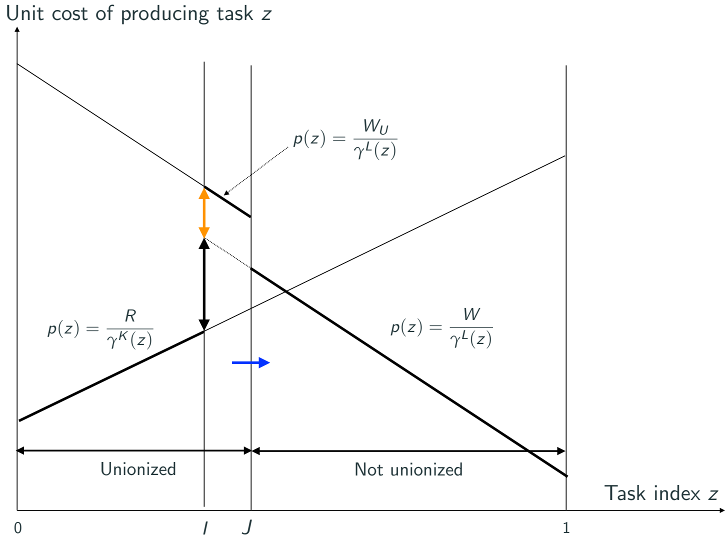
with L_U employment in union jobs and with

$$\Psi_H(I) \equiv \exp\left[\int_0^I \ln(\gamma^K(z))dz + \int_I^1 \ln(\gamma^L(z))dz\right]$$

Automation of union jobs and labor demand

- A union worker earns $W_U > W$ and wages equal marginal product
- Automation of union jobs increases the gain in allocative efficiency such that automation of union jobs is less likely to decrease their marginal product of labor
- Union workers displaced to non-union jobs experience stronger wage decreases
- Impact on welfare is ambiguous because the direct allocative efficiency gain from automation opposes the loss in allocative efficiency from union workers moving to non-union jobs

The impact of wage rents on allocative efficiency



So-so automation of union jobs and allocative efficiency back

- The change in $Y|K, L$ due to automation is given by:

$$\frac{dY}{dI} = \frac{dY}{dI}|_{L_U} + W_U \frac{dL_U}{dI} + W \frac{d[L - L_U]}{dI}$$

- Using the expression for aggregate output above gives:

$$\frac{dY}{dI} = \underbrace{\left[\ln\left(\frac{W}{\gamma^L(I)}\right) - \ln\left(\frac{R}{\gamma^K(I)}\right)\right] Y}_{\text{Gain in allocative efficiency without union wage premium}}$$

$$\underbrace{\left[\ln\left(\frac{W_U}{\gamma^L(I)}\right) - \ln\left(\frac{W}{\gamma^L(I)}\right)\right] Y}_{\text{Extra gain in allocative efficiency}} + \underbrace{[W_U - W]\frac{dL_U}{dI}}_{\text{Extra loss in allocative efficiency}}$$

where the last term is negative given that $dL_U/dI < 0$.