Appendices

Appendices occur in the order as they are referred to in the main text.

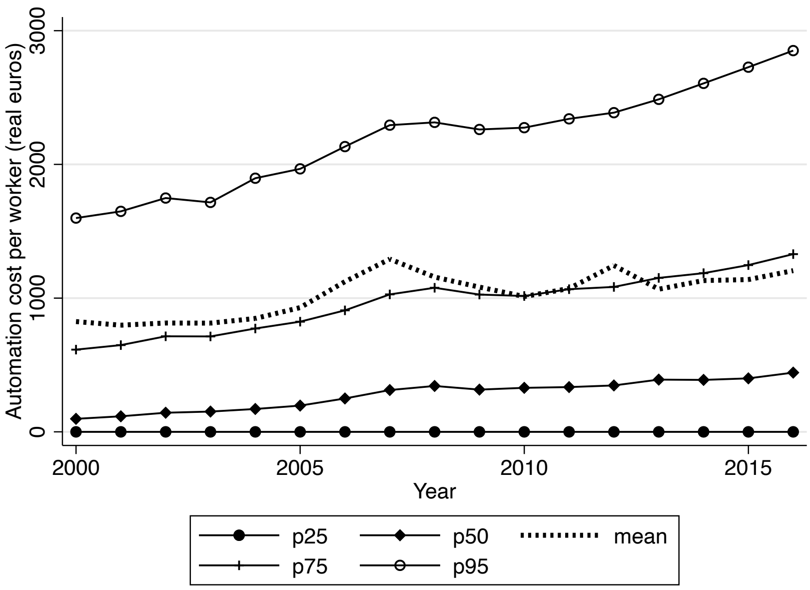
A Data

A.1 Automation costs over time

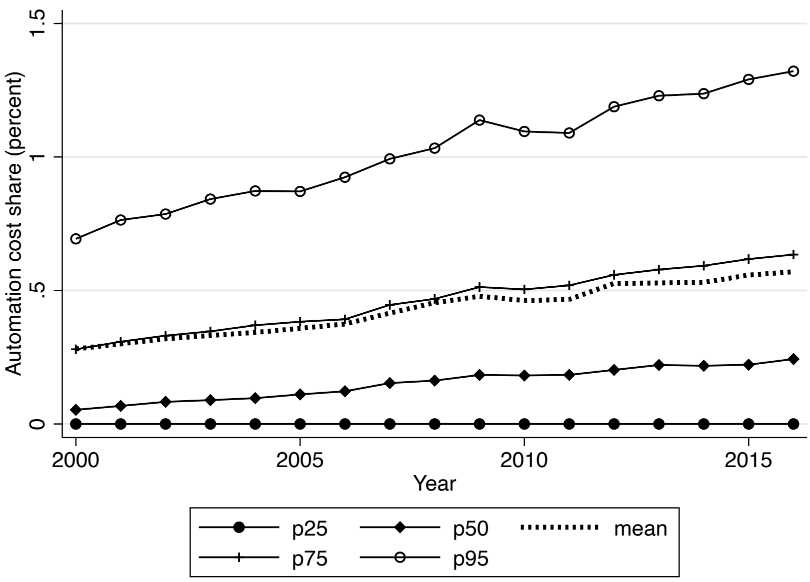
Figure A.1 shows how the distribution of automation costs per worker (top panel) and automation cost shares (bottom panel) have changed over time. Mean automation costs per worker and the mean automation cost share are rising over 2000–2016. Furthermore, besides increases in means, there is a fanning out of the distributions with automation costs rising faster for higher percentiles.

Figure A.1. Automation costs over time

(a) Automation costs per worker



(b) Automation costs shares



A.2 Correlations between automation costs and specific technologies

Table A.1 reports correlations between firms' (standardized) automation cost shares and their self-reported implementation with self- or externally developed process, product, and organizational innovations (all three measured as dummies). The model controls for sector fixed effects and firm size.

Table A.2 estimates the same model– again controlling for sector fixed effects and firm size– for a host of self-reported uses of specific technologies, all of which are given as dummy variables in the data. Note that since these variables originate from different survey years, they have varying overlap with the firms where we observe automation costs; and cannot be entered jointly in a single regression.

Table A.1. Firm-level correlations between automation cost shares and type of innovation

Dependent variable: Standardized automation cost share	
Process innovations	0.205***
	(0.048)
Product innovations	0.098**
	(0.036)
Organizational innovations	0.100*
	(0.041)
N	7,163

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.

Table A.2. Firm-level correlations between automation cost shares and technology usage

Dependent variable: Standardized automation cost share			
Use of electronic data suited to automated processing	0.246***	Received orders for goods or services through EDI	0.104**
	(0.054)		(0.034)
N	4,315	Ordered through Electronic Data Interchange (EDI)	-0.097**
			(0.032)
CRM, inventory and distribution analysis	0.202***	N	14,180
	(0.041)		
Customer Relationship Management (CRM), customer analysis	0.052	Sales software	0.089**
	(0.048)		(0.030)
N	11,934	Purchasing software	0.002
			(0.030)
Enterprise Resource Planning (ERP) software	0.162***	N	7,834
	(0.027)		
N	12,542	Radio Frequency Identification (RFID)	0.051
			(0.083)
Automated records used for value chain integration	0.201**	N	4,149
	(0.066)		
Value chain integration	-0.012	Local Area Network (LAN)	0.015
	(0.047)		(0.027)
N	7,883		7,656
Big data analysis	0.126*	Internet for financial transactions	0.015
	(0.053)		(0.025)
N	4,684	N	7,530
Cloud-services: Software for customer information mngmnt	0.170*	Internet for training and education (incl. e-learning)	0.036
	(0.085)		(0.031)
Cloud-services: Software for accounting and financial mngmnt	0.141*	N	8,388
	(0.063)		
N	6,715		

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.

A.3 Technology usage by sector

Table A.3 shows that the overall use of data for automated processing is relatively common in all sectors, reflecting this is a general characteristic of automation. CRM software for inventory management is used in Professional activities and Information & communication services as well as in Manufacturing. Further, over 60 percent of manufacturing firms use ERP, more than 20 percentage points more than in Wholesale & retail, the sector which has the second-highest rate of use. Automation-compatible value chain integration is also most prevalent in Manufacturing, Wholesale & retail, and Transportation & storage. By contrast, the use of cloud software for accounting and CRM is most widespread in service sectors: Professional activities, Information & communication, and Administrative activities. Applications of big data analysis vary: in Information & communication and Manufacturing, the use of internal firm data is most common, in Transportation this is (unsurprisingly) location data, and the use of social data is most common in Professional activities as well as Accommodation & food serving.

Table A.3. Technology usage by sector

	Data for automated processing	CRM for inventory & distribution	ERP software	Automation-compatible value chain integration	Cloud software for CRM	Cloud software for accounting	Big data: Internal firm	Big data: Location	Big data: Social	Big data: Other
Manufacturing	0.36	0.47	0.63	0.15	0.09	0.11	0.14	0.05	0.05	0.03
Construction	0.23	0.24	0.24	0.05	0.09	0.12	0.10	0.14	0.07	0.01
Wholesale & retail trade	0.43	0.42	0.41	0.17	0.12	0.14	0.12	0.07	0.13	0.05
Transportation & storage	0.39	0.29	0.27	0.18	0.10	0.14	0.16	0.20	0.07	0.05
Accommodation & food serving	0.27	0.18	0.10	0.05	0.14	0.23	0.06	0.05	0.18	0.01
Information & communication	0.43	0.64	0.35	0.10	0.35	0.35	0.15	0.13	0.21	0.07
Prof'l, scientific, & technical activities	0.40	0.52	0.28	0.08	0.20	0.29	0.08	0.06	0.12	0.06
Administrative & support activities	0.34	0.39	0.23	0.08	0.19	0.25	0.07	0.05	0.11	0.04
N	4,315	11,938	12,542	11,028	6,715	6,715	4,684	4,684	4,684	4,684

Notes: Observation numbers differ across columns because technology usage variables are obtained from different survey waves; and because we have discarded missing and imputed values.

A.4 Automation imports

We obtain data on firms' imports, exports, and re-exports of intermediates from Statistics Netherlands: unlike our automation cost measure, which starts in 2000, we can only identify these trade variables from 2010 onward. Following the literature, we define automation machinery using CN-2018 product codes. In particular, we follow the categorization of Acemoglu and Restrepo (2021) and include automatically controlled machines, automatic transfer machines, automatic welding machines, numerically controlled machines, and robots as automated machinery. Examples of descriptions of automatically controlled machines are "Automatic regulating or controlling instruments and apparatus"; examples of automatic transfer machines are "Continuous-action elevators and conveyors, for goods or materials"; examples of automatic welding machines are "Machines and apparatus for arc (including plasma arc) welding of metals"; examples of numerically controlled machines are "Numerically controlled bending, folding, straightening or flattening machines (including presses)"; and robots are described as "Industrial robots, not elsewhere specified or included".

Detailed product codes for each of these are as follows:

- Automatically controlled machines:
  90321080, 90321000, 90328100, 90320000, 90321020, 90328900, 90328100, 90329000, 90322000

- Automatic transfer machines:
  84283100, 84283900, 84573090, 84283300, 84283200, 84283990, 84580000, 84283100, 84283920, 84573000, 84573010

- Automatic welding machines:
  85153100, 85153100, 85152100, 85152100

- Numerically controlled machines:
  845811000080, 845811200080, 845811410010, 845811410080, 845811490080, 845811800080, 845891000010, 845891000080, 845891200080, 845891800080, 845921000010, 845921000080, 845931000010, 845931000080, 845941000010,

845941000080, 845951000010, 845951000080, 845961000010, 845961000080, 845961100080, 845961900080, 846012000010, 846012000080, 846022000010, 846022000080, 846023000080, 846024000080, 846031000010, 846031000080, 846040100080, 846221000010, 846221000080, 846221100080, 846221800080, 846231000010, 846231000080, 846241000010, 846241000080, 846241100080, 846241900080

- Robots: 84795000

Figure A.2 shows real total imports, exports, and re-exports for automation technology over 2010–2016. Re-exports are defined as goods transported via the Netherlands which are (temporarily) owned by a resident of the Netherlands without any significant industrial processing (including, for example, goods that are cleared by Dutch distribution centers and exported to other (European) countries). This shows that exports of automation technologies exceed imports in the Netherlands, and that there is also a substantial amount of re-exports.

Table A.4 compares automation costs and automation imports as a percentage of total operating cost for the overlapping subsample of firms at the sector level[53], revealing that average automation expenditures are substantially higher than average automation imports – since few firms are importers–, and observed across a wider range of sectors. Automation imports and automation expenditures are somewhat correlated at the firm-level, as shown in the first two columns of Table A.5 where firm-by-year automation expenditures are regressed onto (net) automation imports while controlling for firms' total operating cost, and additionally year fixed effects in the second column. However, this correlation disappears when adding firm fixed effects (seen from the last two columns of Table A.5): that is, firms are not more likely to have higher automation costs when they (net) import more automation technology.

[53]We construct firm-level averages and remove firms which cease operations before 2009 when comparing our automation cost data to automation imports.

Figure A.2. Total automation imports, exports, and re-exports over time

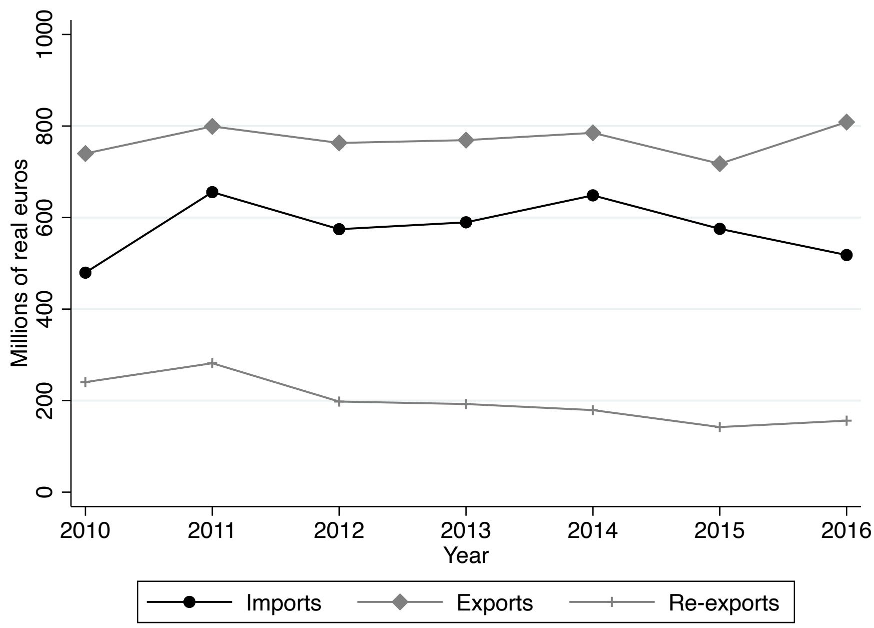


Table A.4. Comparing automation costs to automation imports by sector

	Mean percentage in total costs, for automation:		
Sector	Costs	Imports	Net imports
Manufacturing	0.347	0.080	0.042
Construction	0.196	0.001	0.001
Wholesale & retail trade	0.300	0.058	0.051
Transportation & storage	0.352	0.134	0.095
Accommodation & food serving	0.268	0.000	0.000
Information & communication	0.810	0.004	0.004
Prof'l, scientific, & technical activities	1.000	0.007	0.005
Administrative & support activities	0.434	0.003	0.003

Notes: Total N firms is 30,291. Net automation imports are defined as imports minus re-exports. Total costs include automation costs.

Table A.5. Comparing automation costs to automation imports
between and within firms

Dependent variable: Automation costs (IHS)				
	(1)	(2)	(3)	(4)
Automation imports (IHS)	0.018**	0.018**	-0.001	-0.002
	(0.007)	(0.007)	(0.004)	(0.004)
	(5)	(6)	(7)	(8)
Net automation imports (IHS)	0.016*	0.016*	-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: 110,805 firm-year observations for each model. 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. *p<0.10, **p<0.05, ***p<0.01.

B Theoretical model of firm-level automation spikes

As in Bonfiglioli et al. (2021), our model adds monopolistic competition and firm heterogeneity to task models of endogenous automation developed by Acemoglu and Restrepo (2018, 2019). As in Bonfiglioli et al. (2021), we abstract from new task creation as this is not our object of study. However, the assumptions in our model also differ from Bonfiglioli et al. (2021) in two ways. First, our model assumes fixed instead of convex adjustment costs of automation. Second, our model formulates the firm's decision to automate as a dynamic instead of a static profit maximization problem. In particular, it assumes that the fixed costs of automation are irreversible (i.e. cannot be recouped other than through higher profits in the future) and that the lowest possible output price using the most recent automation technologies falls over time.

Together, this implies that a firm will automate when the expected gain from moving to the lowest possible output price outweighs its costs of adjustment (abstracting from the importance of product demand shocks). However, because adjustment costs are fixed and irreversible, a firm will not automate in every period. Instead, immediately after the firm decides to automate it will wait a number of periods until the increase in expected profit again outweighs its fixed adjustment costs. Therefore, our model predicts spikes in automation cost shares over time. This prediction of spikes in automation costs shares is key to our empirical identification strategy. It is also consistent with the literature on lumpy investment in capital, as in Haltiwanger et al. (1999); Doms and Dunne (1998), or in robots, as in Humlum (2021). However, these papers do not capture the task-based approach to endogenous automation.

B.1 Exogenous automation

B.1.1 Consumption

Assume consumers derive utility from consuming goods Y_1,...,Y_J, according to the following CES utility function:

$$U(Y_1, ..., Y_J) = \left[\sum_{j=1}^{J} [\epsilon_j Y_j]^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{B.1}$$

$$\text{such that } \sum_{j=1}^{J} P_j Y_j = PY$$

where $\sigma > 1$.[54] ϵ_j captures specific preferences across goods, Y is utility or real income spent on Y_1,...,Y_J, and P is the ideal price index[55] The price index P is given by:

$$P(P_1, ..., P_J) \equiv \left[\sum_{j=1}^{J} [P_j/\epsilon_j]^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = 1 \tag{B.2}$$

where the last equality follows from choosing consumption as the numeraire such that all prices are relative to P.

From equations (B.1) and (B.2) we obtain that product demand for firm j is given by:

$$Y_j = Y \epsilon_j^{\sigma-1} P_j^{-\sigma} \tag{B.3}$$

where Y captures demand shocks that are common across all firms $j = 1, ..., J$, and ϵ_j captures shocks in the relative demand for Y_j relative to $Y_{\tilde{j}}$ with $\tilde{j} \neq j$ and $\tilde{j} = 1, ..., J$.[56]

[54] Humlum (2021) assumes that $\sigma = 4$.

[55] Note that $\sum_{j=1}^{J} P_j Y_j = E(P(1), ..., P(J), U) = e(P(1), ..., P(J))U = PY$ with $E(P(1), ..., P(J), U)$ the expenditure function, $P \equiv e(P(1), ..., P(J))$ the expenditure function per unit of utility, and $Y \equiv U$ utility or real income spent on Y_1,...,Y_J.

[56] If ϵ_j would be the same for all j, it would disappear from equations (B.2) and (B.3). Also, if ϵ_j's are different across j but would all increase by the same proportion, equation (B.3) simplifies to $Y_j = Y \epsilon_j^{-1} P_j^{-\sigma}$ capturing that utility but not real income has

B.1.2 Production

For each firm $j = 1, ..., J$, the production of Y_j is given by the following Cobb-Douglas production function:

$$Y_j = \exp\left(\int_0^1 \ln(y_j(z))dz\right) \tag{B.4}$$

with $y_j(z)$ a quantity of task z used in the production of Y_j.[57]

Each task is produced using a quantity of capital, $k_j(z)$, or labor, $\ell_j(z)$, according to:

$$y_j(z) = \begin{cases} \ell_j(z) + \gamma_j(z)k_j(z) & \text{if} \quad z \in [0, I] \\ \ell_j(z) & \text{if} \quad z \in (I, 1] \end{cases} \tag{B.5}$$

with $\gamma_j(z)$ a firm-specific task productivity schedule of capital. Assume that $\gamma_j(z)$ is decreasing in z. That is, tasks are ordered on the unit-interval such that capital has a comparative advantage in the production of lower-indexed tasks.

Task $I \in (0, 1)$ is a task threshold such that all tasks $z \leq I$ can be produced by labor or capital (and will be produced by capital in equilibrium), and all tasks $z > I$ can only be produced by labor. Also assume that for all relevant levels of I:

$$W > \frac{R}{\gamma_j(I)} \tag{B.6}$$

with W is the price of one unit of $\ell_j(z)$ and R is the price of one unit of $k_j(z)$. That is, as new automation technologies make it feasible for labor tasks just above I to be performed by capital, cost minimizing firms have an incentive to adopt these new automation tech-

increased by the same proportion.

[57]Equation (B.4) implicitly assumes that the elasticity of substitution between tasks is unity. Although direct estimates do not exist, Goos et al. (2014) report an elasticity of substitution between 21 occupations of 0.85, and Humlum (2021) finds an elasticity of substitution between production workers, tech workers and other workers of 0.49. Assuming a more general CES production function would not qualitatively change the analysis.

nologies. Consequently, all tasks $z \in [0, I]$ will exclusively be performed by capital. We also assume that firm j takes I as given when deciding how much labor and capital to use, i.e. that technology is a quasi-fixed factor. Section B.2 relaxes this assumption by endogenizing the firm's decision to automate.

Finally, firm-level employment of capital K_j and labor L_j is given by:

$$\int_0^1 k_j(z)dz = K_j \quad \text{and} \quad \int_0^1 \ell_j(z)dz = L_j \tag{B.7}$$

and we assume that each firm takes the wage W and the rental rate of capital R as given.

B.1.3 Conditional factor demands

If factors are paid their revenue marginal products and firms minimize costs, the unit-cost of producing task z, $p_j(z)$, is given by:

$$p_j(z) = \begin{cases} R/\gamma_j(z) & \text{if} \quad z \in [0, I] \\ W & \text{if} \quad z \in (I, 1] \end{cases} \tag{B.8}$$

Given that equation (B.4) is a Cobb-Douglas production function using a continuum of tasks on a unit-interval, cost shares must be constant and equal across all tasks in equilibrium. In particular, we must have that:

$$\forall z : p_j(z)y_j(z) = \frac{\sigma - 1}{\sigma} P_j Y_j \tag{B.9}$$

where $[\sigma - 1]/\sigma < 1$ accounts for the fact that firm j earns a profit $P_j Y_j/\sigma$ because it charges a constant price mark-up $\sigma/[\sigma - 1] > 1$ over marginal costs.[58]

[58] Given constant returns to scale and no fixed costs, marginal and average costs are the same such that a constant mark-up $\sigma/[\sigma - 1]$ over marginal costs implies that average costs can be written as $[\sigma - 1]/\sigma P_j$. Consequently, profits per unit of output are given by $P_j - [\sigma - 1]/\sigma P_j = P_j/\sigma$ and profits can be written as $P_j Y_j/\sigma$.

Using equation (B.9) together with equations (B.5) and (B.8), it holds that:

$$k_j(z) = \begin{cases} \frac{\sigma-1}{\sigma}\frac{P_jY_j}{R} & \text{if} \quad z \in [0, I] \\ 0 & \text{if} \quad z \in (I, 1] \end{cases} \qquad \ell_j(z) = \begin{cases} 0 & \text{if} \quad z \in [0, I] \\ \frac{\sigma-1}{\sigma}\frac{P_jY_j}{W} & \text{if} \quad z \in (I, 1] \end{cases} \tag{B.10}$$

which gives demand for capital and labor for each task z, respectively.

Using equations (B.7) and (B.10) then solves for RK_j and WL_j:

$$RK_j = I\frac{\sigma - 1}{\sigma}P_jY_j \tag{B.11}$$

and

$$WL_j = [1 - I]\frac{\sigma - 1}{\sigma}P_jY_j \tag{B.12}$$

which gives the firm's conditional factor demands.

B.1.4 Output as a Cobb-Douglas aggregate

In equilibrium, P_j is a constant mark-up $\sigma/[\sigma-1] > 1$ over the marginal cost of producing Y_j. Using the corresponding expression for the marginal cost of producing Y_j given the Cobb-Douglas production function in equation (B.4), we get that:

$$P_j = \frac{\sigma}{\sigma - 1}\exp\left(\int_0^1 \ln(p_j(z))dz\right) \tag{B.13}$$

Substitute expressions for R and W from equations (B.11) and (B.12) into equation (B.8). Then substitute equation (B.8) into equation (B.13). Taking logarithms, we obtain:

$$\ln(Y_j) = \int_0^I \ln(\gamma_j(z))dz + I\ln(\frac{K_j}{I}) + [1 - I]\ln(\frac{L_j}{1 - I})$$

Taking the exponential on both sides yields the following expression for firm output:

$$Y_j = \varphi_j \left[\frac{K_j}{I}\right]^I \left[\frac{L_j}{1 - I}\right]^{1-I} \tag{B.14}$$

where φ_j is defined as:

$$\varphi_j \equiv \exp\left(\int_0^I \ln(\gamma_j(z))dz\right)$$

B.1.5 Unconditional labor demand

Dividing equation (B.11) by equation (B.12) gives:

$$\frac{K_j}{I} = \frac{1}{1-I}\frac{W}{R}L_j \tag{B.15}$$

and substituting equation (B.15) into equation (B.14) gives:

$$Y_j = \varphi_j \frac{L_j}{1-I}\left[\frac{W}{R}\right]^I \tag{B.16}$$

Substituting equation (B.16) into equation (B.12) and using equation (B.3):

$$L_j = \left[\frac{\sigma-1}{\sigma}\right]^\sigma Y \epsilon_j^{\sigma-1} W^{-\sigma}[1-I]\left[\left[\frac{W}{R}\right]^I \varphi_j\right]^{\sigma-1} \tag{B.17}$$

with $\sigma > 1$ and $I \in [0, 1]$.

Equation (B.17) is the firm's unconditional demand for labor. It shows that firm-level labor demand increases if there is an increase in the firm's product demand captured by an increase in Y if the shock is common across firms, or in ϵ_j if the product demand shock is firm-specific. Equation (B.17) also shows that an increase in the automation possibility frontier I has an ambiguous impact on labor demand. On the one hand, $1 - I$ decreases capturing a direct displacement effect. On the other hand, $[W/R]^I\varphi_j$ increases capturing a productivity effect from automation. To see that the productivity effect is positive, take logs of $[W/R]^I\varphi_j$ and differentiate with respect to I. We then get $\ln(W\gamma_j(I)/R)$ which is positive given equation (B.6). Finally note that the productivity effect is increasing in $\gamma_j(I)$. That is, the productivity effect is larger and labor demand is more likely to increase in a firm where capital is more productive at the automation possibility frontier.

B.1.6 Output

Substitute equation (B.17) into equation (B.16) we get the following expression for output:

$$Y_j = Y\epsilon_j^{\sigma-1}\left[\frac{\sigma}{\sigma-1}\frac{W^{1-I}R^I}{\varphi_j}\right]^{-\sigma} = Y\epsilon_j^{\sigma-1}P_j^{-\sigma} \tag{B.18}$$

with

$$P_j = \frac{\sigma}{\sigma-1}\frac{W^{1-I}R^I}{\varphi_j} \tag{B.19}$$

The first ratio on the right-hand side of equation (B.19) is the constant price mark-up and the second ratio is the marginal cost of producing Y_j.

Equation (B.18) shows that firm-level output increases if there is an increase in the firm's product demand for a given output price. Equations (B.18) and (B.19) also show that firm-level output increases if the automation possibility frontier I increases because of a productivity effect from automation that decreases marginal costs and therefore the output price. To see that the productivity effect decreases marginal costs, take logs of $W^{1-I}R^I\varphi_j^{-1}$ and differentiate with respect to I. We then get $\ln(R/[W\gamma_j(I)])$ which is negative given equation (B.6).

B.1.7 Profits

Using equations (B.18) and (B.19), profits can be written as:

$$\Pi_j = \frac{P_jY_j}{\sigma} = \frac{Y\epsilon_j^{\sigma-1}}{\sigma}\left[\frac{\sigma}{\sigma-1}\frac{W^{1-I}R^I}{\varphi_j}\right]^{1-\sigma} \tag{B.20}$$

showing that profits increase following an increase in product demand for a given output price or following a decrease in the output price because of automation.

B.2 Endogenous automation

B.2.1 Expected profits

Denote time periods by t and assume that firm j chooses if and when to automate by maximizing expected profit:

$$\max_{D_{j0},D_{j1},\dots} E\sum_{t=0}^{\infty}\beta^t\left[\Pi_{jt}-D_{jt}F_j\right]$$

with $D_{jt}=1$ if firm j decides to automate at time t and $D_{jt}=0$ otherwise. F_j is a fixed firm-specific cost of automation.[59] $\beta<1$ is the discount rate.[60] Substituting Π_{jt} with equation (B.20) gives:

$$\max_{D_{j0},D_{j1},\dots} E\sum_{t=0}^{\infty}\beta^t\left[\frac{Y_t\epsilon_{jt}^{\sigma-1}}{\sigma}P_{jt}^{1-\sigma}-D_{jt}F_j\right] \tag{B.21}$$

with P_{jt} given by equation (B.19).[61]

The adjustment cost of adopting a new technology F_j is fixed (in real terms) and therefore indivisible. We also assume that these fixed costs cannot be recouped other than through higher future profits, which also makes them irreversible. It is plausible that investments in automation meet these two criteria: automation likely requires fixed adjustment costs from reorganizing production processes, and these costs are irreversible if they require e.g. the development of custom software or worker training.

[59]The model presented in Bonfiglioli et al. (2021) is similar to ours but differs in that automation costs are assumed to be convex instead of fixed.

[60]Humlum (2021) assumes an annual discount rate $\beta=0.96$.

[61]Note that we implicitly assume that automation does not result in a wage premium for workers (who remain or are newly) employed at automating firms. This is in line with our empirical findings.

B.2.2 Technological progress

When a firm decides not to automate, it keeps producing at the same marginal cost. But when a firm decides to automate, its marginal cost and therefore its output price decrease to minimum values determined by the technological frontier. We also assume that these minimum values decrease over time because of underlying exogenous technological progress captured by an increase over time in I_t (and possibly also $\gamma_{jt}(z)$ for all or some z that is common across firms). In particular, we assume that:

$$P_{jt} = \begin{cases} P_{jt-1} & \text{if } D_{jt-1} = 0 \\ \mathcal{P}_t & \text{if } D_{jt-1} = 1 \end{cases} \tag{B.22}$$

where $\mathcal{P}_t$ is the lowest possible output price using the most recent automation technologies and given that the firm decided to automate in period $t-1$.[62] Further assume that technological progress is captured by a decrease in $\mathcal{P}_t$ over time given by $\mathcal{P}_t = \mu \mathcal{P}_{t-1}$ with $\mu < 1$.[63]

If the firm last adopted automating technologies in period $\tilde{t}$, the "age" of its technology in period t is $t - \tilde{t}$. Its output price in period t can be written as $P_{jt} = \mathcal{P}_{j\tilde{t}} = \mathcal{P}_t \mu^{\tilde{t}-t}$. That is, the firm's relative output price increases as its technology ages. Normalizing $\mathcal{P}_t = 1$ in every period, we can rewrite the firm's expected profit in equation (B.21) as:

$$\max_{D_{j0}, D_{j1}, \ldots} E \sum_{t=0}^{\infty} \beta^t \left[\sigma^{-1} Y_t \epsilon_{jt}^{\sigma-1} \mu^{(\tilde{t}-t)(1-\sigma)} - D_{jt} F_j \right] \tag{B.23}$$

[62] The assumption that the decision to automate or otherwise invest happens one period in advance is common in the literature estimating firm-level production functions, including Humlum (2021); Doraszelski and Jaumandreu (2013); Olley and Pakes (1996).

[63] Strictly speaking $\mathcal{P}_t$ is each firm's lowest possible *relative* output price. This implicitly assumes that not all firms decide to automate (or not) at the same time despite common technological progress, which in our model will be true because firms differ in their fixed adjustment costs.

where the firm can reset $\tilde{t}$ to t such that $\tilde{t} - t = 0$ if it chooses to automate ($D_{jt-1} = 1$).

B.2.3 Spikes in automation expenditures

The indivisibility and irreversibility of automation costs imply that automation occurs in relatively infrequent episodes of disproportionately large quantities. Consistent with the literature on lumpy investment (Haltiwanger et al. 1999; Doms and Dunne 1998), our model therefore predicts spikes in firms' automation cost expenditures. In particular, equation (B.23) captures that the firm's expected profit decreases over time if it does not automate given that $\mu < 1$ and $(\tilde{t} - t)(1 - \sigma) > 0$ (because $t \geq \tilde{t}$ and $\sigma > 1$). The reason for this is that the firm falls further behind the technology frontier as its technology ages.

To see this more formally, derive the firm's per-period profit excluding adjustment cost with respect to t to get (for given Y_t and ϵ_{jt}):

$$\frac{\partial[\sigma^{-1} Y_t \epsilon_{jt}^{\sigma-1} \mu^{(\tilde{t}-t)(1-\sigma)}]}{\partial t} = \frac{\sigma - 1}{\sigma} Y_t \epsilon_{jt}^{\sigma-1} \mu^{(\tilde{t}-t)(1-\sigma)} \ln(\mu) < 0 \tag{B.24}$$

where the last inequality follows given that $\sigma > 1$, $t \geq \tilde{t}$, and $\mu < 1$. Said differently, the increase in expected profits from automation implies that a firm will automate. However, the firm will not automate in every period given its fixed costs of automation. In particular, immediately after the firm decides to automate it will wait a number of periods until the increase in expected profit again outweighs its fixed adjustment cost.

In sum, our model predicts spikes in firms' automation cost shares over time because automation involves fixed costs and automation events are preceded and followed by periods in which firms will not automate. Moreover, firms will automate at different points in time if they have different fixed adjustment costs.

B.2.4 Shocks to product demand

Not only technological progress and the firm's fixed cost of automation determine when a firm automates. Shocks in the firm's product demand will also increase profits which could induce the firm to automate. To illustrate this, assume an increase in real income Y_t which increases the firm's product demand. If the firm, after observing Y_t, expects that

Y_{t+1} will also be higher (e.g. because shocks to product demand follow a Markov process), equation (B.24) shows that expected profits in $t+1$ will be higher. Importantly, equation (B.24) further shows that this increases the firm's increase in expected profits if it decides to automate. The same is true for an increase in ϵ_{jt}. In sum, positive shocks to product demand that persist over time increase the likelihood that the firm will automate.[64]

Therefore, persistent shocks to product demand are likely to (in part) predict automation events. However, they also directly affect firms' future outcomes, making them potential confounders for estimating the causal impact of automation. For example, equation (B.17) shows that a persistent positive product demand shock directly increases the firm's future labor demand. If this product demand shock is unobserved by the econometrician while it also induces the firm to automate, estimates of the impact of automation on labor demand will be upward biased.

B.3 Empirical implications

Automation can empirically be measured as spikes in automation costs which we directly observe at firm-level, capturing F_j when $D_{jt} = 1$. Assuming that F_j is larger for larger firms, we express automation costs as a share of the firm's total costs excluding automation costs.[65] When a firm decides to automate, our model predicts a spike in the firm's automation cost share because automation involves fixed costs and is preceded and followed by periods in which the firm will not automate.

Further, our model highlights that common demand shocks (captured by Y_t) that are persistent over time may trigger automation in some firms, depending on the level of firm-specific fixed costs. These common shocks are a threat to identification when using

[64] Only if product demand shocks are independently and identically distributed will they not be correlated with the firm's decision to automate. If product demand shocks are i.i.d., they are not informative about the future and therefore will not affect the firm's decision at time t whether or not to automate.

[65] This is similar to Bonfiglioli et al. (2021) who construct a proxy for automation intensity defined in their model as the firm's chosen level of automation over capital expenditure.

a firm-level event-study design: automation events may be correlated with such positive common demand shocks, confounding the effects on firm-level outcomes. This motivates using difference-in-differences, comparing firms that do and do not automate in any given year, removing the common shock component.

However, firm-specific demand shocks (captured by ϵ_{jt}) may confound identification in such a difference-in-differences set-up. In particular, our model points out that firms we do not observe automating may fail to do so exactly because they do not experience (large enough) persistent positive firm-specific demand shocks within our observation window. If automating firms experience more positive demand shocks than do non-automating firms, we will obtain a biased estimate of the impact of automation when comparing automators to non-automators. This motivates our empirical approach of exploiting firm automation timing among firms that we do observe automating: that is, we use a difference-in-differences event-study design.

We can interpret the two identification assumptions of parallel trends and no anticipation for our difference-in-differences event-study design, outlined in Section 4.2, in the context of this model:

1. Firm-specific product demand shocks must be identically and independently distributed for automating firms. That is, ϵ_{jt} is i.i.d. for firms which automate at some point in our observation window. If these firm-specific product demand shocks are not i.i.d. among automators, the firm's decision of when to automate will be positively correlated with persistent firm-specific product demand shocks and its direct impact on firm outcomes in the future. I.i.d. ϵ_{jt} result in parallel trends between firms that automate at different points in time. Our empirical analyses provide evidence for this parallel trends assumption by showing that pre-event trends are mostly similar for firms that have an automation event now compared to those that have an automation event later. We also match individual incumbent workers on observable characteristics and show there are no pre-trends in their labor market outcomes.

   Related theories make similar assumptions. Bonfiglioli et al. (2021) present a static model, thereby implicitly assuming that firm-specific shocks in product demand are i.i.d. Humlum (2021) assumes firm-specific productivity (instead of product demand) shocks that evolve according to a Markov chain of length three. He then draws on the literature that estimates firm-level production functions to estimate parameters in the Markov chain. Assuming that any remaining unobserved productivity shocks are i.i.d., he then matches pairs of firms on initial firm-level outcomes before one (but not the other) firm automates to causally estimate the impact of automation on firm-level outcomes.

2. No anticipation implies firms do not anticipate automation when determining how many workers to employ and how much to produce. This is captured by our assumption (in section B.1) that firms treat their technology as given. Moreover, we implicitly assume that not only firms but also workers take the firm's technology as given when making decisions about labor supply.

   Similar assumptions are made in related papers. Humlum (2021) makes similar no-anticipation assumptions for firms and workers, while also explicitly modeling workers' labor supply in general dynamic equilibrium. Although Bonfiglioli et al. (2021) present a static model, they also implicitly assume that firms treat their technology as quasi-fixed when maximizing operating profits and that, just as we do here, firm-level labor supply is perfectly elastic.

   Arguably, the non-anticipation assumption is less likely to hold for firms than for their incumbent workers (i.e. workers with at least 3 years of tenure). One reason is that firms (instead of all their incumbent workers) decide whether or not to automate and that firms are better informed about the likely consequences of automation (and even control them). Another reason is that employment at the firm-level is more flexible, e.g. because of regular turnover or fixed-term contracts, than incumbent workers' perceived outside options. Finally, incumbent workers (rather than recent hires employed at the firm) are less likely to adjust their labor supply in anticipation of an automation event. Therefore, we are more cautious in interpreting effects of automation at the firm level as causal but are more confident at the incumbent worker level.

C Automation events

C.1 Automation spike frequency

Table C.1. Spike frequency

Spike frequency	N firms	% of N firms
0	25,155	70.7
1	8,354	23.5
2	1,772	5.0
3	266	0.7
4 or 5	33	0.1
Total	35,580	100

Notes: Spike frequency is defined as the total number of spikes occurring over 2000–2016. The total number of firms is 35,580 and the total number of firms with at least one automation cost share spike is 10,425.

C.2 Automation events across sectors and firm size classes

Table C.2. Share of firms ever having an automation spike

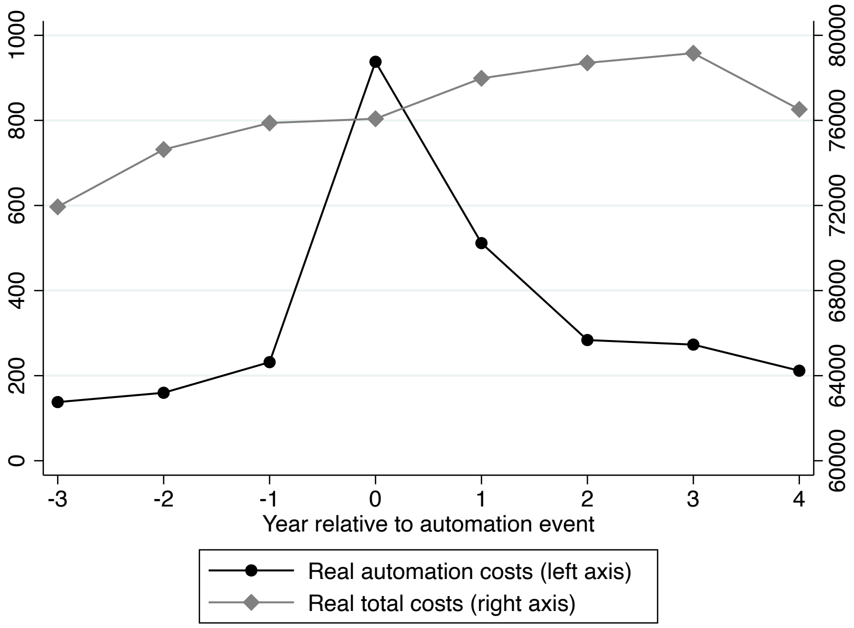
By sector	
Manufacturing	0.29
Construction	0.25
Wholesale & retail trade	0.27
Transportation & storage	0.30
Accommodation & food serving	0.27
Information & communication	0.39
Prof'l, scientific & technical activities	0.33
Administrative & support activities	0.30
By firm size	
1-19 employees	0.26
20-49 employees	0.30
50-99 employees	0.31
100-199 employees	0.29
200-499 employees	0.32
$\geq$500 employees	0.28

Notes: N=35,580 firms.

C.3 Automation and total costs around automation events

Figure C.1 shows the two components making up our automation spike measure, automation costs and total costs, for a balanced sample of firms around automation events. This highlights that automation spikes are driven by increases in automation costs, not decreases in total costs.

Figure C.1. Automation costs and total costs around automation spikes



Notes: Balanced sample of firms.

D Firm-level analyses

D.1 Predicting automation events

Table D.1 estimates a firm-level linear probability model where the dependent variable is a dummy for the firm having at least one automation spike over 2000–2016. This table highlights that firms that have automation events are different from those that do not. In particular, the probability of having an automation event is higher for firms with younger and more highly educated workers and with a higher fraction of men, firms that pay higher wages, larger firms, and firms in Information & communication, Professional, scientific & technical activities, Transportation & storage, and Administrative & support activities.

Table D.1. Correlates of a firm ever having an automation spike

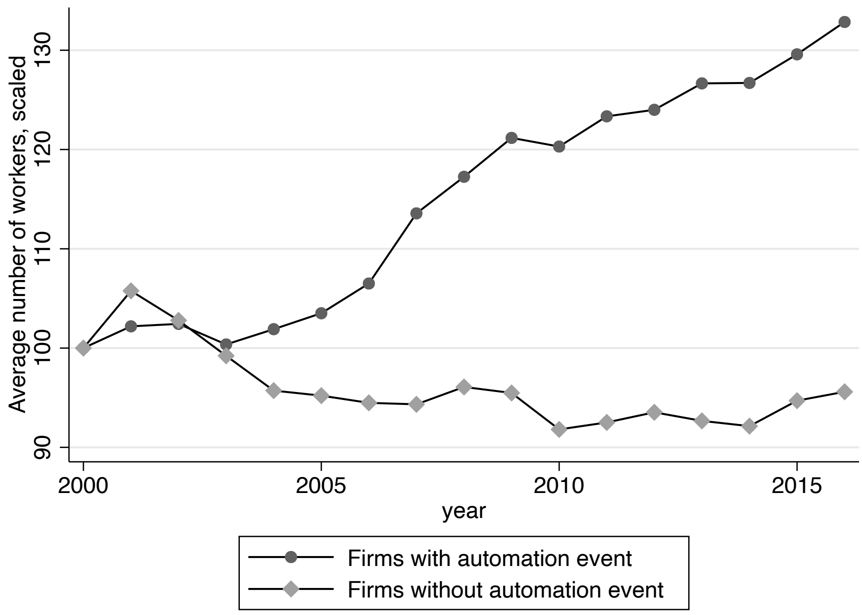
Mean annual wage	0.0009***	Manufacturing	*reference*
	(0.0002)		
Share of women	-0.0259**	Construction	-0.0003
	(0.0126)		(0.0094)
Mean worker age	-0.0034***	Wholesale & retail trade	0.0161**
	(0.0005)		(0.0080)
Share high educated	0.0368*	Transportation & storage	0.0414***
	(0.0197)		(0.0102)
1–19 employees	*reference*	Accommodation & food serving	-0.0022
			(0.0155)
20–49 employees	0.1146***	Information & communication	0.1094***
	(0.0060)		(0.0123)
50–99 employees	0.1218***	Prof'l, scientific, & technical activities	0.0580***
	(0.0074)		(0.0112)
100–199 employees	0.1174***	Administrative & support activities	0.0262***
	(0.0090)		(0.0101)
200–499 employees	0.1317***		
	(0.0113)		
≥500 employees	0.1133***	Constant	0.2856***
	(0.0141)		(0.0230)

Notes: 35,577 observations, each observation is a unique firm. The dependent variable is having an automation spike at any point in the sample. Mean real annual wage in thousands of euros. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

D.2 Employment growth for automating and non-automating firms: balanced panel

Figure D.1 uses the balanced panel of firms existing over the entire 17-year period and plots a time series of firm-level employment averaged across automating and non-automating firms with both series normalized to 100 in 2000.

Figure D.1. Average firm-level employment for firms with and without automation events



Notes: All firms existing over the entire 17-year period 2000–2016. $N = 399$ for firms with an automation event and $N = 623$ for firms without an automation event.

D.3 Difference-in-differences with non-automating firms

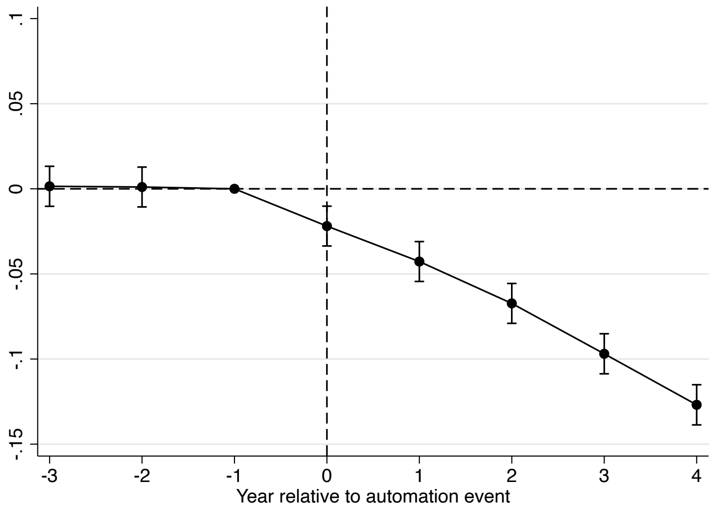
Here, we construct a different control group for our main analysis. Instead of using firms that automate later, we use firms that do not automate over the period we observe them. We follow the same steps in constructing the sample as a stacked difference-in-differences as discussed in the main text and estimate the same models.

Figure D.2 reports the findings for employment and wages. All estimates are weighted by firm size. The top figures reveal that for employment there are clear divergent pre-trends between automating and non-automating firms (right), which do not seem to exist when comparing automating with later-automating firms (left). By contrast, for wages the pre-trends appear quite similar, while we find almost no significant impacts on wages using non-automating firms and small positive impacts using later-automating firms as control group.

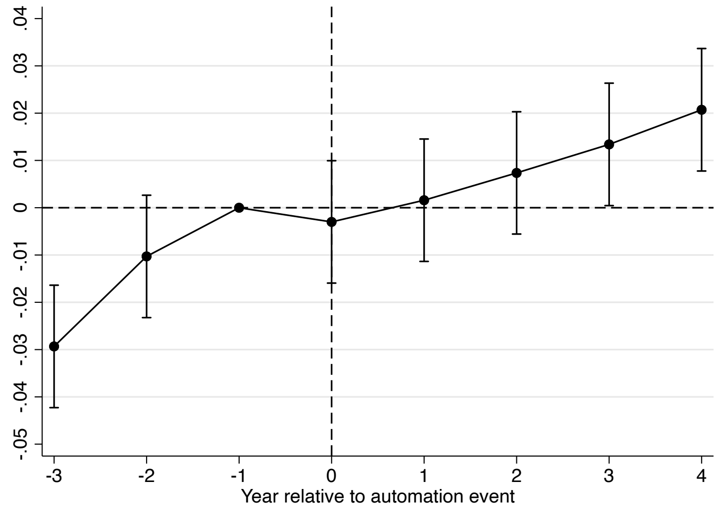
An important limitation of this analysis is that we do not observe automation costs for firms in each year. Hence, we cannot be sure that firms do not automate in a year where we do not observe them in the survey. Restricting the sample to firms that we observe each year would leave us with a very small dataset. This, along with the clear differential employment trends between automating and non-automating firms observed here and in the balanced panel in Appendix D.2, supports our design using later-treated firms.

Figure D.2. Using later-automating (left) vs non-automating firms (right) as control group

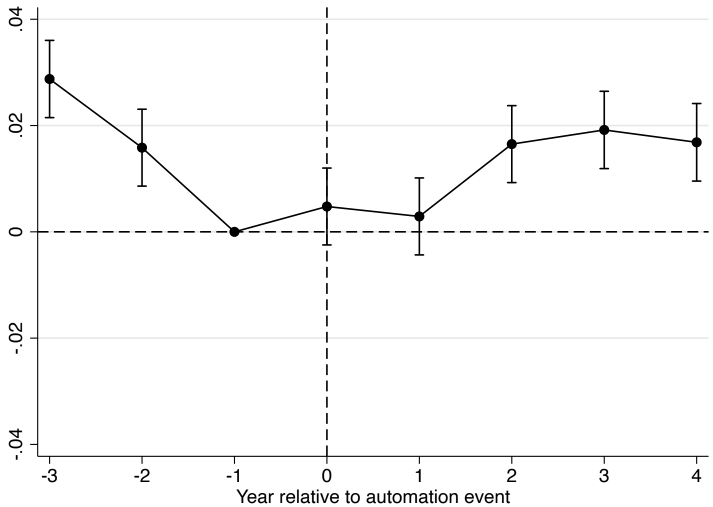
(a) Log employment (baseline)



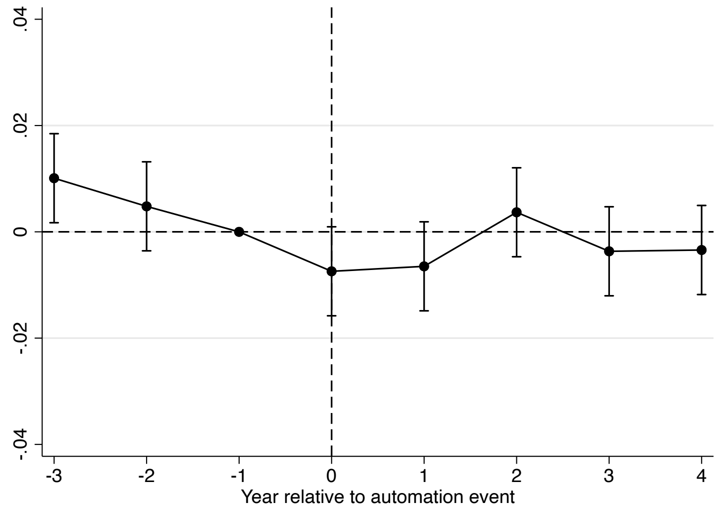
(b) Log employment (using non-automating firms)



(c) Log daily wage (baseline)



(d) Log daily wage (using non-automating firms)



Notes: $N = 865,848$ for the sample using non-automating firms as control group and $N = 170,022$ for the sample using later-automating firms as control group. Whiskers represent 95 percent confidence intervals.

D.4 Comparison to import-based automation measure

Here, we compare firms with and without automation events to importers and non-importers of automation technology: this is done for the subset of $30,291$ firms in our main sample where we can construct time-invariant importer information.

Table D.2 shows that firms that import automation technology are substantially larger than firms that do not import such technology, consistent with findings for other countries (e.g. see Bonfiglioli et al. 2021 and Humlum 2021). In particular, importing firms are around 131% ($exp(0.839) - 1$) larger than non-importers, while firms with automation cost spikes are only around 8.5% larger than firms without such spikes.

Table D.3 shows that, for all firm-level measures of automation, automating firms also have faster employment and wage bill but not daily wage growth: however, employment growth and wage bill differences are substantially larger when comparing automation importers to non-importers. Specifically, (net) importers have around 3.7% faster wage bill growth compared to non-(net-)importers, whereas firms with automation cost spikes have 1.2% faster wage bill growth compared to firms without such events.

Table D.4 considers whether automation events are correlated with firms importing automation technology at the firm level. We find that firms with non-zero mean (net) automation imports are more likely to have automation events, implying that some firms with automation events are also importers of automation technology. However, while over 30% of firms in this sample have an automation event, only around 8% are importers. Further, Table D.5 shows that the correlation between automation events and occurrence of automation imports is negligible *within* firms: this mirrors our finding for the within-firm correlation between automation cost shares and automation import values reported in Appendix Table A.5.

Lastly, we estimate the impact of automation cost spikes on firm-level employment, wage, and wagebill growth for the sample of firms with automation events where we observe import data, distinguishing between impacts for all firms and for the subsample of firms which also import automation technology. Figure D.3 shows difference-in-differences estimates as in equation 3. This shows that automation importers experience employment

Table D.2. Correlation between firm size and automation type

Dependent variable: Log firm-level number of employees

	Automation cost spike		Automation imports	
	(1)	(2)	(3)	(4)
Automating	0.078***	0.085***	0.856***	0.839***
	(0.013)	(0.013)	(0.022)	(0.022)
Sector fixed effects	No	Yes	No	Yes

Notes: N = 30,291 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.

Table D.3. Employment, wage, and wagebill growth for firms with automation cost spikes and non-zero automation imports

Dependent variable:	Δ log employment	Δ log mean daily wage	Δ log wage bill
	(1)	(2)	(3)
Cost spikes	0.012**	0.000	0.012**
	(0.006)	(0.002)	(0.005)
	(4)	(5)	(6)
Imports	0.0311***	0.006*	0.037***
	(0.008)	(0.004)	(0.006)
	(7)	(8)	(9)
Net imports	0.030***	0.006*	0.036***
	(0.008)	(0.003)	(0.006)

Notes: N = 152,550 firm-year observations. All models include calendar year fixed effects, and initial-year values for log employment and log mean daily wage. All models are weighted by the inverse of the number of firm-level observations multiplied by baseline firm-level employment size. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

Table D.4. Firm-level correlation between automation events and automation imports

Dependent variable: Dummy for firm having an automation cost spike				
	(1)	(2)	(3)	(4)
Importer	0.023*	0.029**		
	(0.010)	(0.011)		
Net importer			0.023*	0.029**
			(0.010)	(0.011)
Controls	No	Yes	No	Yes

Notes: N = 30,291 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.

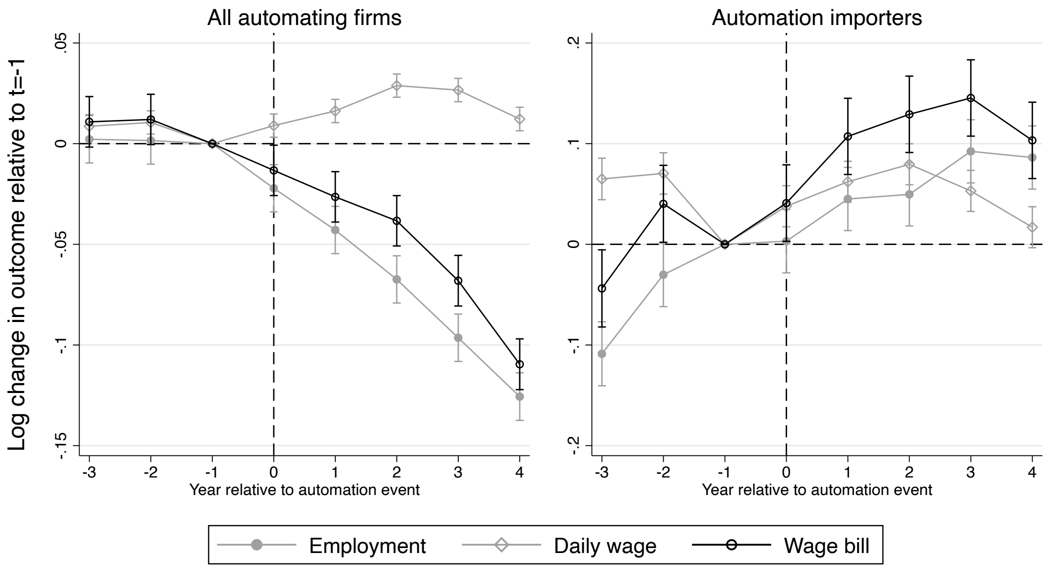
Table D.5. Within-firm correlation between automation events and automation imports

Dependent variable: Dummy for firm having an automation cost spike				
	(1)	(2)	(3)	(4)
Importer	0.005	0.002	0.003	0.003
	(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 = 110,805 firm-year observations. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

and wage growth around automation events.

Figure D.3. Firm-level outcomes for automating firms, difference-in-differences using event timing



Notes: All automating firms that exist in all years $\tau \in \{-3, .., 4\}$. Both models are weighted by firm-level employment size in $\tau = -1$. Standard errors are clustered at the firm-level. Whiskers reflect 95% confidence intervals.

E Worker-level analyses

E.1 Sample construction

For each calendar year y we define a set of potential treatment and control group automation events as follows. Potential treatment events are defined as a firm having its first automation event in treatment year c. c is between 2003 and 2011, so that for each automation event we at least have a window of three years before and five years after the event. This gives us 3,004 potential treatment group events. Potential control group events for c are defined as firms that have their first automation event in year $c + 5$ or later. Hence, these events have to occur between 2008 and 2016. This gives us 21,289 potential control group events.

Columns (1) and (2) in Table E.1 show the number of potential treatment and control events per calendar year. Note that our procedure implies that multiple control group events can involve the same firm, but for different calendar years. It is also possible that one treatment group event and one or more control group events involve the same firm in different calendar years. For example, a firm that has its first automation event in 2010 can be a potential treatment event in 2010, but also serve as a potential control event for treatment events in 2003, 2004, or 2005. Similarly, a firm having its first automation event in 2011 can serve as a control group event for treatment events in 2003, 2004, 2005, or 2006. For our 21,289 potential control events, 20,572 involve a firm that is involved in more than one potential control event, while 717 events involve a firm that is involved in only one potential control event. Firms with potential control events are on average involved in 4.7 potential control events, with a maximum of 9 events. For our 3,004 potential treated events, 1,288 involve a firm that is also involved in at least one potential control event in another year and 1,716 involve a firm that is not involved in a potential control event.

We then merge our firm-level data to worker data and keep only events for which we can find at least one incumbent worker who is between 18 and 65 years old at $\tau = -1$. This leaves us with 2,995 potential treatment events merged to 192,755 incumbent workers and 21,115 potential control events merged to 1,132,190 incumbent workers. We then apply some sample selections to drop outliers. In particular, we drop workers with yearly earnings more than 500,000 euros in any one year or an average daily wage above 2,000 euros. We also exclude students. Finally, we require incumbents to not receive any benefits in the three years before treatment. This leaves us with 997,057 potential control workers and 162,493 potential treated workers.[66]

Finally, we match treated and control group workers on pre-treatment annual real wage income, separately by sector and calendar year. While the match is exact for calendar year and sector, we use coarsened exact matching (CEM, see Iacus et al. 2012; Blackwell et al. 2009) for pre-treatment income. To this end, we construct separate strata for each 10 deciles of real wage income, as well as separate bins for the 99th and 99.9th percentiles, in each of the three pre-treatment years $\tau = -3, -2, -1$. We then match treated workers to control group workers for each of these income bins, while additionally requiring them to be observed in the same calendar year, and work in the same sector one year prior to treatment. We include calendar year and sector matching to ensure we are not capturing sector-specific business cycle effects, or other unobserved time-varying shocks affecting workers based on their original sector of employment. As such, each treated worker is matched to a set of controls from the same calendar and sector and belongs to the same pre-treatment earnings percentile bin. This procedure results in 29,224 strata for incumbent workers, and in doing so can match 98.7% of treated incumbents (using 94.2% of control group incumbents).

After matching, our sample contains 1,098,924 incumbent workers in treatment and control groups. Of those incumbent workers, 160,419 are treated and 938,505 are controls. Our estimation sample of firms for identifying these treated and control group workers contains 6,179 unique firms, all of which experience an automation event at some point over the period. As indicated in columns (3) and (4) of Table E.1, workers employed at 2,981 firms are treated, and workers employed at 4,464 firms serve as controls at least once. This entails that 1,266 firms who serve as control events in one year,

[66]Appendix E.2 below provides further summary statistics of our worker data.

Table E.1. Number of treatment and control events at the firm level by calendar year

	Potential events:		Events after matching:	
	Control	Treatment	Control	Treatment
Year 2003	3,474	224	3,399	223
Year 2004	3,245	242	3,185	240
Year 2005	2,936	237	2,890	235
Year 2006	2,688	299	2,651	300
Year 2007	2,415	380	2,350	379
Year 2008	2,167	394	2,125	392
Year 2009	1,887	418	1,853	414
Year 2010	1,510	406	1,480	401
Year 2011	967	401	951	397
Total	21,289	3,004	20,884	2,981
Unique firms involved	4,512	3,004	4,464	2,981
Unique firms only used once	717	1,716	734	1,715

Notes: Table shows the number of potential treatment and control events, and the number of events remaining after matching, for each calendar year.

also serve as treatment event in an earlier year.

E.2 Summary statistics for workers

Table E.2 provides summary statistics on our sample of incumbent workers across all years. Column 1 shows descriptives before matching, and columns 2 and 3 show descriptives for our matched sample of incumbent workers (both treated and control). Note that we have $160,419 + 938,505 = 1,098,924$ observations for incumbents: given our observation window of 8 years ($\tau = -3$ through $\tau = 4$) this adds up to the $1,098,924 \times 8 = 8,791,392$ incumbent worker observations used in our regressions. Note that column 1 has more observations as this also includes the small fraction of workers not used as a treated worker (because no match could be found for them) or not used as a control group worker.

Table E.2. Descriptives for incumbent workers

	(1) *Full sample*	(2) *Treated workers*	(3) *Control workers*
Annual wage income	40244.15	39708.60	39780.08
	(27344.07)	(26616.67)	(26866.99)
Daily wage if employed	117.41	110.35	110.33
	(75.09)	(76.21)	(75.98)
Annual non-employment duration (in days)	22.22	5.43	5.19
	(81.06)	(32.46)	(31.82)
Hazard of leaving the firm	0.05	0.00	0.00
	(0.21)	(0.00)	(0.00)
Total benefits	443.74	0.00	0.00
	(2992.76)	(0.00)	(0.00)
Probability of entering early retirement	0.01	0.00	0.00
	(0.11)	(0.00)	(0.00)
Probability of becoming self-employed	0.02	0.01	0.01
	(0.13)	(0.09)	(0.11)
Share female	0.26	0.35	0.32
	(0.44)	(0.48)	(0.47)
Foreign born or foreign-born parents	0.16	0.18	0.16
	(0.36)	(0.38)	(0.37)
Age	42.11	40.24	40.11
	(10.20)	(9.99)	(9.92)
Calendar year	2006.88	2006.14	2006.14
	(3.36)	(2.35)	(2.35)
Manufacturing	0.36	0.20	0.20
	(0.48)	(0.40)	(0.40)
Construction	0.11	0.07	0.07
	(0.32)	(0.25)	(0.25)
Wholesale & retail trade	0.19	0.33	0.33
	(0.40)	(0.47)	(0.47)
Transportation & storage	0.09	0.08	0.08
	(0.28)	(0.28)	(0.28)
Accommodation & food serving	0.02	0.02	0.02
	(0.13)	(0.12)	(0.12)
Information & communication	0.06	0.08	0.08
	(0.23)	(0.27)	(0.27)
Prof'l, scientific, & technical activities	0.08	0.09	0.09
	(0.27)	(0.29)	(0.29)
Administrative & support activities	0.09	0.13	0.13
	(0.29)	(0.33)	(0.33)
0–19 employees	0.05	0.05	0.07
	(0.22)	(0.21)	(0.26)
20–49 employees	0.14	0.13	0.17
	(0.35)	(0.34)	(0.37)
50–99 employees	0.11	0.11	0.13
	(0.32)	(0.31)	(0.34)
100–199 employees	0.12	0.11	0.13
	(0.33)	(0.32)	(0.34)
200–499 employees	0.15	0.14	0.16
	(0.36)	(0.35)	(0.36)
$\geq$500 employees	0.43	0.46	0.34
	(0.49)	(0.50)	(0.47)
Missing education	0.71	0.71	0.71
	(0.45)	(0.45)	(0.45)
Low education	0.06	0.05	0.05
	(0.23)	(0.23)	(0.22)
Middle education	0.12	0.12	0.12
	(0.32)	(0.32)	(0.32)
High education	0.12	0.12	0.12
	(0.32)	(0.32)	(0.33)
N	9,276,400	160,419	938,505

Notes: Column 1 shows unweighted means for all incumbent worker-year observations. Columns 2 and 3 show weighted means for the full regression sample at $\tau = -1$, where weights are obtained from coarsened exact matching as described in Appendix E.1. Standard deviations in parentheses.

E.3 Predicting automation event timing

To test whether the timing of automation events is random, one can try to predict the timing of automation events based on observable characteristics of automating firms. In particular, using Brier (1950) skill scores, we can test whether a predictive model with observables performs better than a random prediction where we uniformly distribute automation events across years where the automating firms are observed.

Specifically, Brier (1950) skill scores for the ten k-folded samples reported in Table E.3 are constructed as follows. We draw a 10 percent random sample without replacement from the sample of 10,425 automating firms, and do this ten times: these are the test samples. The remaining 90 percent of observations for each of these test samples constitute the ten training samples. We then estimate a logit model with firm fixed effects and time-varying observables (firm average log yearly and daily wages, log total wage bill, log number of workers, log average worker age, log average worker tenure at the firm, share female and a full set of interactions) for each training sample and predict the probability of having a spike in a year for each corresponding test sample, assuming that each firm will have exactly one spike. We also calculate the spike probability by year per firm from random prediction, simply as one over the number of years the firm is observed. For the model-based and random predictions in each of the ten test samples, we calculate the Brier score, defined as the mean squared difference between the prediction and the actual outcome. Lastly, we obtain the Brier skill score as $1 - \frac{Brier_{model}}{Brier_{random}}$, reflecting the percent prediction improvement of the model relative to random prediction. Table E.3 shows that these improvements are low, ranging between 2.6 and 3.4%, suggesting that the timing of automation events is essentially random with respect to firms' observed characteristics.

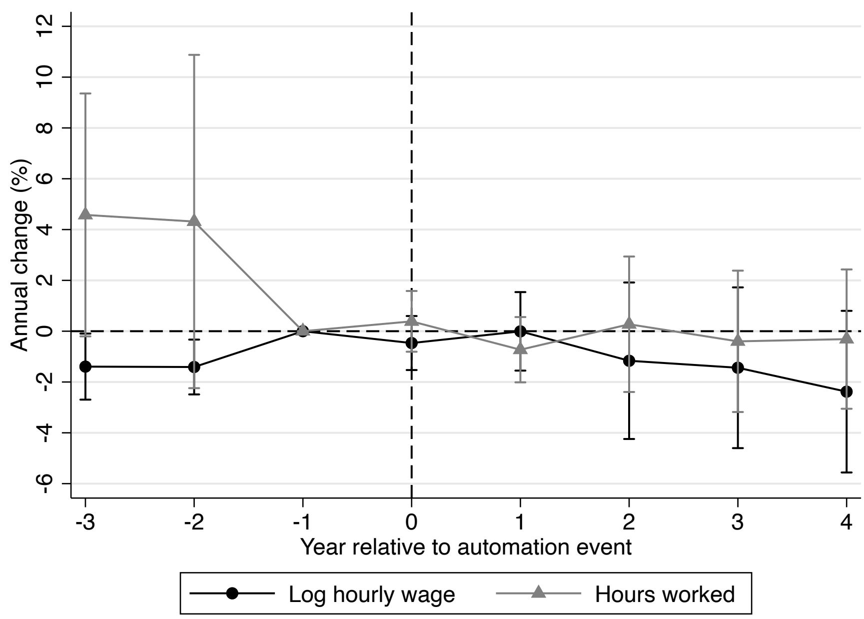
Table E.3. Brier skill scores for predicting automation event timing

Sample	N	Brier skill score
1	127,485	0.029
2	127,463	0.027
3	126,753	0.027
4	127,708	0.026
5	126,890	0.030
6	126,328	0.028
7	127,921	0.034
8	127,145	0.033
9	126,676	0.033
10	127,475	0.032

E.4 Effects on hourly wages

Figure E.1 shows effects on incumbents log hourly (rather than daily) wages and relative hours worked, both of which which we observe only for 2006 onward. In line with the impacts on daily wages we find no statistically significant impacts on hourly wages.

Figure E.1. Impact of automation on incumbents' log hourly wages and hours worked



Notes: N=2,042,874 for hourly wages and N=2,128,936 for hours worked. Whiskers represent 95 percent confidence intervals.

E.5 Effect heterogeneity

Here, we consider effect heterogeneity by incumbent characteristics. For succinctness, we only show estimates for relative annual wage earnings, as this is the summary measure capturing all other impacts. Any noteworthy differences in results for other worker-level outcomes are described where relevant.

We consider how incumbent workers with different characteristics fare after an automation event. For each of the groups considered here, we contrast the effect against the same group at the control firm by using an interaction term – this results in a decomposition of the mean effects found in the main text. In particular, we estimate the following model:

$$y_{ijt} = \alpha + \beta D_i + \gamma post_\tau + \delta_0 \times treat_i \times post_{i\tau} + \sum_k \left[\delta_k \times treat_i \times post_\tau \times z_{ki}\right] + \lambda X_{ijt} + \varepsilon_{ijt}, \tag{E.1}$$

where, as before, i indexes workers, j firms, t calendar time, and τ time relative to the automation event. For succinctness, we estimate the average annual effect over the entire post-treatment period rather than reporting the year-by-year coefficients. As such, $post_\tau$ is a dummy variable indicating the post-treatment period (i.e. $\tau \geq 0$). Further, z_{ki} is a dimension of worker heterogeneity, such as gender, age, or sector, containing $k + 1$ categories– all time-varying characteristics are measured in the year before automation. In addition to the controls included in equation (4), X_{ijt} also contains z_{ki} as well as the interaction terms $z_{ki} \times treat_i$ and $z_{ki} \times post_i$. In equation (E.1), δ_0 gives the estimated treatment effect for the reference group, and δ_k the deviation from that effect for category k of worker characteristic z_i. βD_i capture worker fixed effects, and standard errors are clustered at the treatment level as before.

Table E.4 summarizes how average post-treatment effects for annual wage income differ across sectors, and for workers of different genders, with different contract types, and in different age-specific wage quartiles. Results by firm size, worker age, and eduation level are reported in the main text.

In column 1 of Table E.4, we consider to what extent the impacts of automation differ depending on which sector the worker's firm belongs to: that is, our treatment effect is interacted with workers' sector of employment in $\tau = -1$. For this model, Manufacturing is the reference category. Note that sectoral differences may exist for various reasons. First, automation technologies may be sector-specific, and differ in terms of how much they displace labor. For example, it is possible that industrial or warehouse robots are more labor-replacing than automated check-out systems. Second, the workers employed in these different industries may have different characteristics (including unobservable ones), making the impacts differ. Third, to the extent that skills are industry-specific, sectoral labor market conditions matter: displacement would be more costly in sectors with an excess supply of workers. While we cannot distinguish between these different explanations, it is still important to consider whether our results are driven by displacement effects in a subset of sectors, or whether the found impacts are pervasive. Our finding here is that automation leads to wage income losses that are very pervasive across sectors: this highlights that robotics is likely not the only automation technology displacing workers from their jobs. The exception is Accommodation & food serving, where no income losses (nor increases in firm separation) are detected. However, Accommodation & food serving is also a sector with one of the lowest automation expenditures per worker, as well as contributing only 2% of the sample of incumbent workers. On the other hand, incumbent workers in Wholesale & retail and Manufacturing do experience earnings losses – together, these two sectors employ over half of all incumbents in our sample (33% and 20%, respectively). We find that automation leads to increased firm separation rates for all sectors except Accommodation & food serving and Construction. All in all, we find that automation events originating in different sectors have qualitatively similar impacts on workers.

Similarly, we do not find any statistically significant differences in impact by gender (column 2). If anything, the coefficient suggests wage losses are larger for female workers, which would be consistent with recent work from the displacement literature showing that job loss leads to larger losses for women (Illing et al., 2021).

While we also do not find statistically significant differences by workers' contract type (column 3), but the estimates suggest losses are larger for workers with flexible contracts as opposed to open-ended contracts. This could reflect differences in employment protection.

Unfortunately, our data do not contain any occupation information, and only limited education information. For this reason, we obtain an alternative measure of workers' skill level by calculating each worker's wage rank by age in $\tau = -1$. We then group workers into quartiles based on this rank. For example, the top-quartile workers in this measure are those who earn in the top 25% of earnings across the sample for workers of their age in the year before the automation event.

Results are reported in the fourth column of Table E.4: workers in the lowest age-specific wage quartile are used as the reference category. We do not detect any statistically significant differences: that is, workers across all wage quartiles experience displacement from automation. However, the lowest-paid workers (i.e. those in the bottom two quartiles) do experience the largest wage earnings losses, compared to those paid above the median wage (the top two quartiles).

Differences in losses across the wage distribution may of course be partially driven by differences in the firms where automation spikes occur: lower losses for one "skill" group may be offset by higher exposure to automation events in our sample. While the estimates in column (4) matter for the average worker's exposure to displacement from automation, we are also interested in which workers are displaced within firms. Therefore, the fifth column in Table E.4 reports estimates by workers' age-specific *within-firm* wage quartile. That is, the bottom quartile reflects incumbents who are in the lowest 25 percent of their firm's wage distribution for their age.[67] If anything, this reveals that the medium-paid workers by age *within* firms appear to lose more wage income than do workers in the top and bottom quartiles, although these differences are not statistically significant. Overall, these results are consistent with our findings for education level reported in the main

[67]Note that these quartiles cannot be calculated for the smallest firms: however, all previous findings are very similar in this subsample, suggesting that this is not driving the results.

Table E.4. Annual wage income effects by incumbents' characteristics

(1) Sector	
Manufacturing (reference)	-1.66**
	(0.83)
Deviations from reference group for:	
Construction	0.25
	(1.52)
Wholesale & retail trade	-0.58
	(1.13)
Transportation & storage	1.41
	(1.52)
Accommodation & food serving	2.89**
	(1.43)
Information & communication	-1.05
	(1.55)
Prof'l, scientific, & technical activities	-0.95
	(1.53)
Administrative & support activities	-1.09
	(2.46)

(2) Gender	
Male (reference)	-1.55***
	(0.56)
Deviation from reference group for:	
Female	-0.88
	(0.73)

(3) Contract type	
Open-ended contract (reference)	-1.74***
	(0.44)
Deviation from reference group for:	
Flexible contract	-2.34
	(3.11)

(4) Overall age-specific wage quartile	
Bottom quartile (reference)	-2.10*
	(1.24)
Deviations from reference group for:	
Second quartile	-0.06
	(1.20)
Third quartile	0.44
	(1.24)
Top quartile	0.14

(5) Within-firm age-specific wage quartile	
Bottom quartile (reference)	-1.38
	(1.78)
Deviations from reference group for:	
Second quartile	-0.86
	(2.12)
Third quartile	-1.02
	(2.22)
Top quartile	-0.26
	(1.77)

Notes: Estimates from five separate models, N=8,791,392 for models (1) through (4); N=6,418,104 for model (5). All coefficients are average annual effects over the post-treatment period ($\tau = 0$ through $\tau = 4$): estimates have been multiplied by 100 to reflect percentages. *p<0.10, **p<0.05, ***p<0.01.

text. However, we should be careful about drawing strong conclusions from columns (4) and (5) of Table E.4 since they may be capturing other factors than pure worker skill, such as the quality of the worker-firm match.

Lastly, we study effect heterogeneity by restricting our data to the subset of incumbents working in automating firms that import automation technology. Table E.5 shows that earnings declines are not found in this selected subsample. This is similar to our findings for the largest firms, and highlights that effects for importers may reflect the fact that these are large and high-productivity firms.

Table E.5. Average annual impacts for incumbents in importing firms

	Annual wage income	Firm separation hazard	Days in non-employment	Log daily wage
Automation event impact	-0.47	0.23	2.35	0.38
	(1.33)	(1.07)	(2.01)	(1.39)
N	1,418,320	1,339,577	1,418,320	1,374,858

Notes: Importing firms are those who import automation technology worth at least 10,000 euros.

E.6 Incumbent workers versus recent hires

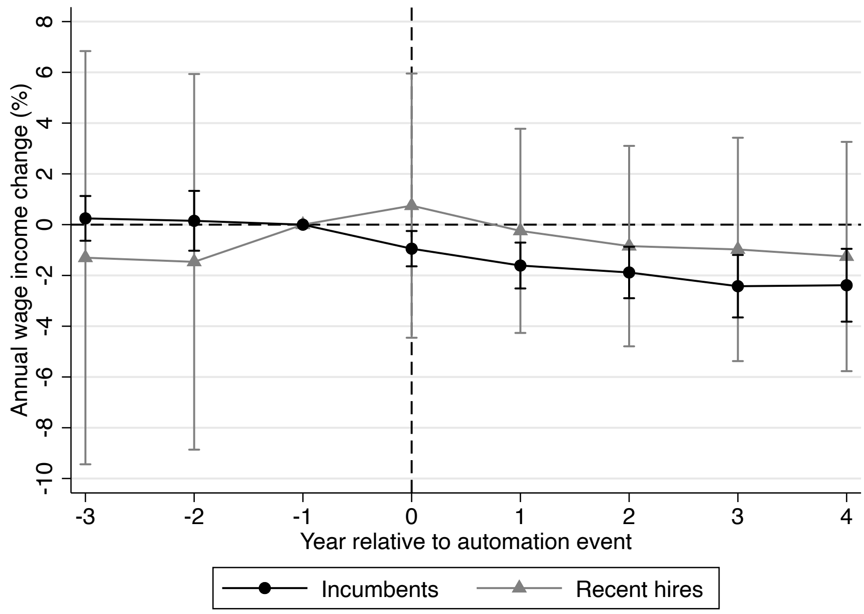
Our identification strategy for the impacts of automation is to consider individual workers who have a pre-existing working relationship with the firm, as evidenced by at least three years of firm tenure. Here we estimate our models for a second group of workers: those with less than three years of firm tenure prior to the automation event. Compared to incumbent workers, these workers are employed at a firm in $\tau = -1$ but not in $\tau = -3$ – we therefore refer to them as recent hires. This worker group is more likely to hold temporary contracts, which could imply different treatment effects. However, causal identification of the treatment effect for recent hires could prove more difficult as they may have been hired in anticipation of the automation event. We therefore analyze them separately, and put more stock in our results for incumbent workers.

We estimate equation (4) for recent hires in the same way we have for incumbents, while additionally creating a zero income bin when matching on pre-event income.[68] After matching, our sample contains $314,484$ unique recent hires ($63,178$ of whom are treated): given our observation window of 8 years ($\tau = -3$ through $\tau = 4$) this results in $2,515,872$ observations.

We find income losses from automation for recent hires that are only half the size of those of incumbents, as shown in Figure E.2. Moreover, relative to recent hires in the control group, point estimates are not significant – hence, recent hires do not have different annual wage earnings as a result of automation. This could be the case because

[68] We obtain 30,679 strata for recent hires, and can match 96.1% of treated recent hires (using 80.0% of control group recent hires).

Figure E.2. Relative annual wage income effects for incumbents versus recent hires



Notes: N=8,791,392 for incumbents and N=2,515,872 for recent hires. Whiskers represent 95 percent confidence intervals.

recent hires have built up less firm-specific human capital, and therefore are more able to adapt to new job tasks either within the same firm or when moving to a new employer. However, it may also be the case that recent hires do not lose income because these workers are in part hired in anticipation of the automation event – in this case their outcomes are endogenous to the event. Consistent with new hires being better matched (or able to adjust) to their firms' new technologies, we find small positive (albeit statistically insignificant) wage effects for this group, on the order of 1.0–1.4%.

E.7 Robustness tests

E.7.1 Constructing placebo events

In this subsection we report descriptive statistics for our placebo analysis using spikes in other material fixed assets. The results of this analysis are reported in section E.4b.

Table E.6 shows the distributions of automation costs and investments in other material assets in the overlapping sample of firms, both in real euros and in real euros per worker. Table E.7 shows the frequency of both types of spikes in this sample, where we consider spikes in other material fixed assets placebo spikes. Lastly, Figure E.3 shows the evolution of investments in material fixed assets around placebo spikes. Results for worker impacts are shown in the main text.

Table E.6. Automation costs and other material asset investments distributions

	Automation cost		**Other material fixed assets**	
	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	1,213	23
p75	75,758	1,043	34,277	456
p90	263,129	2,373	180,821	1,684
p95	620,724	3,839	473,242	3,344
mean	271,888	1,125	181,772	1,067
mean excl. zeros	377,964	1,564	349,980	2,054
N firms × years	171,875		171,875	
N firms × years with 0 costs	48,237		82,607	

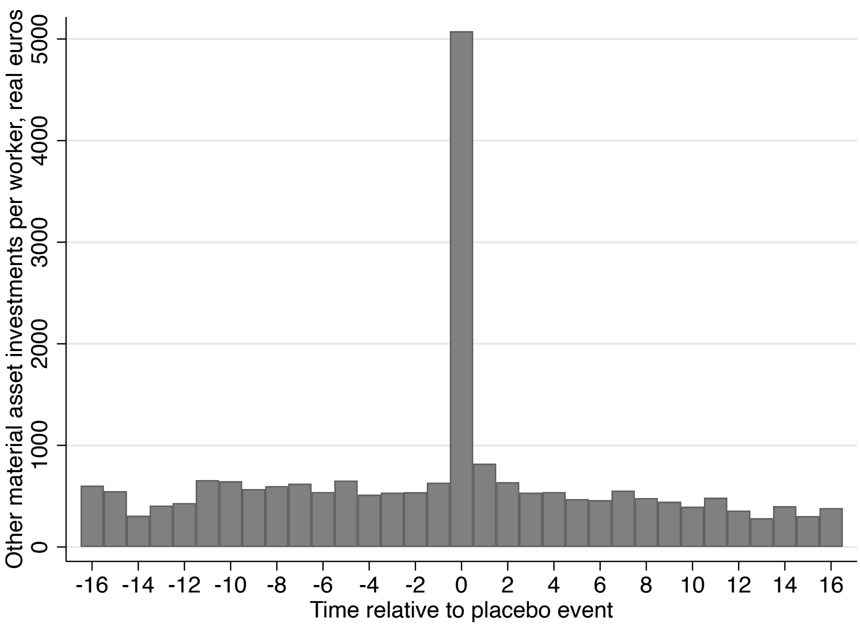
Notes: All numbers are in 2015 euros.

Table E.7. Automation and placebo events

Nr of spikes	**Percentage of firms with event type:** Automation	Placebo (other material fixed assets)
0	71.9	44.6
1	22.5	42.0
2	4.8	11.8
3	0.7	1.5
≥4	0.1	0.1

Notes: Overlapping sample of firms, N = 25,103.

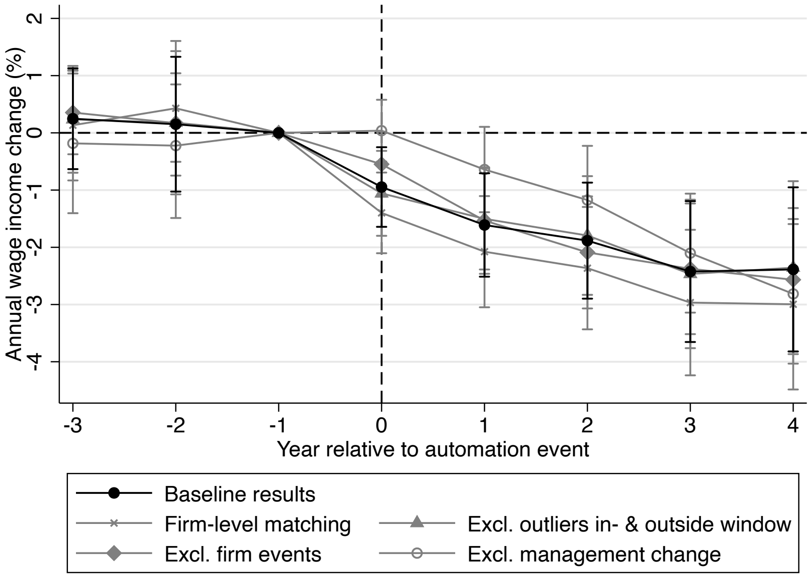
Figure E.3. Placebo events: Spikes in other material fixed assets



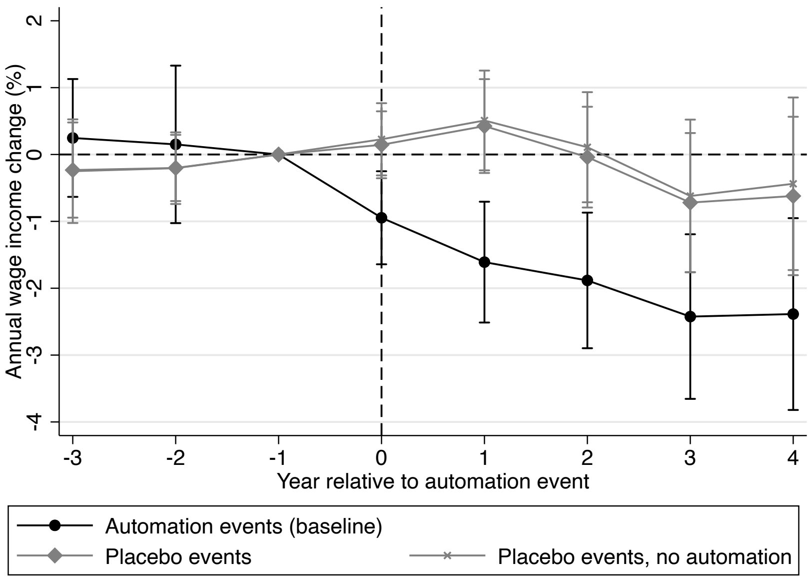
E.7.2 Robustness to other firm events and placebo events

Figure E.4. Robustness tests

(a) Robustness to other firm events



(b) Comparison to placebo events



E.7.3 Alternative definitions of automation events

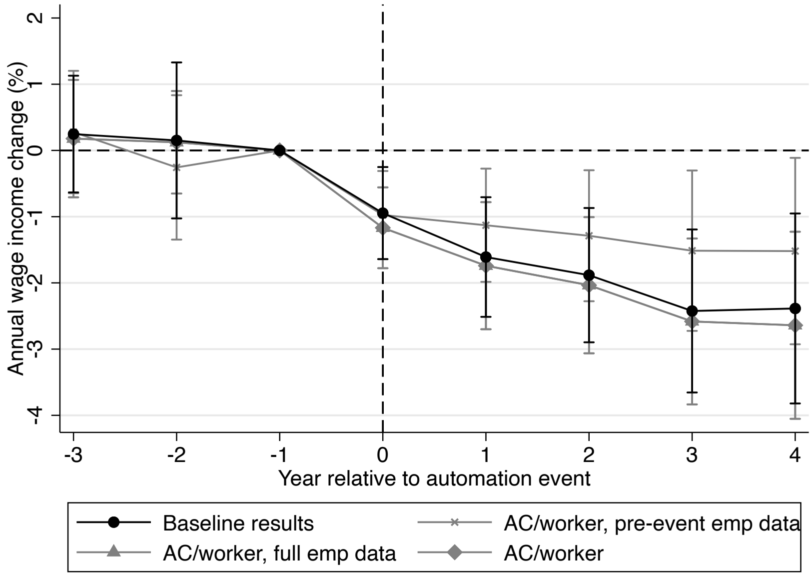
Rather than using automation cost shares (i.e. automation costs in total costs), we can construct automation events from sharp increases in automation outlays per worker.

This is more in the spirit of a literature studying the impact of increasing the number of robots per worker. Within this event definition, we then also vary the point(s) in time where we measure employment – either for the years where we have data on total costs (“AC/worker”); or for the full set of years (“AC/worker, full emp data”); or only for the years pre-dating the candidate automation event (“AC/worker, pre-event emp data”). All variations produce similar results to our baseline estimates, as seen in panel (a) of Figure E.5.

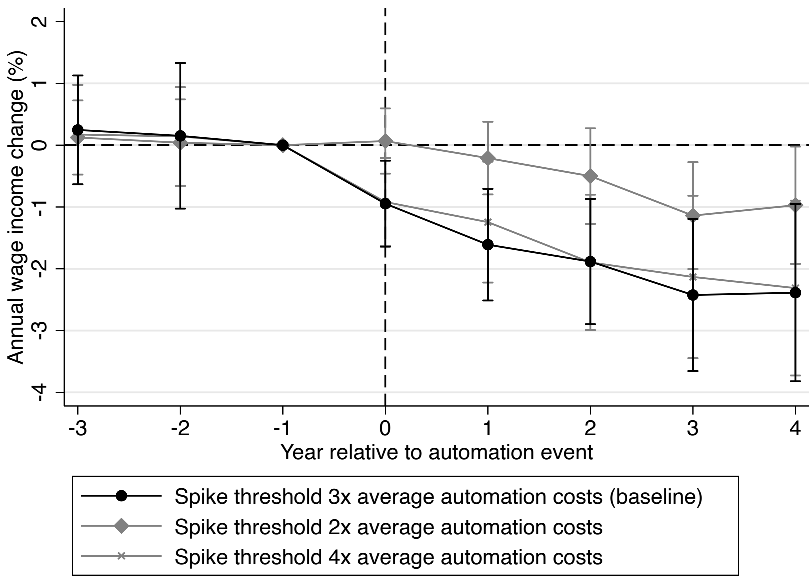
Further, we show that results are robust to varying the spike threshold from two to four times the average automation costs (our baseline is thrice the average automation costs). Panel (b) in Figure E.5 reveals that estimated effect sizes are somewhat larger the higher the threshold, as expected, but these differences are not statistically significant. This highlights that our results are not driven by the specific spike size cut-off we employ in our baseline estimates.

Figure E.5. Robustness to different definitions of automation events

(a) Automation costs per worker



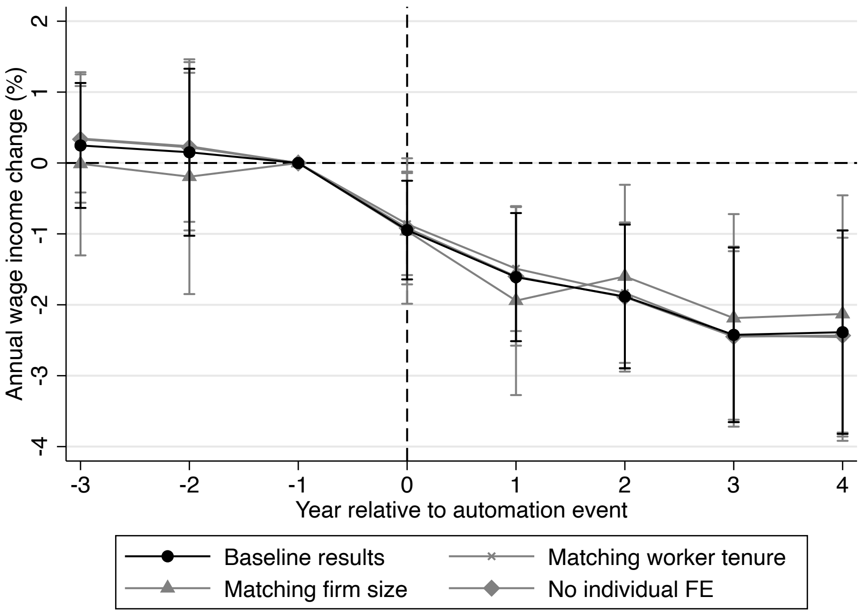
(b) Changes in spike threshold for automation cost shares



E.7.4 Changes in model specification

Here, we change our model specification in a number of ways. In particular, compared to our baseline estimates, Figure E.6 shows results when additionally matching workers on their firm tenure in years (that is, beyond the three years of firm tenure that all treated and control group workers have); additionally matching workers on firm size; and when removing individual fixed effects from the model (these are then replaced by dummies for worker gender and nationality, as well as fixed effects for firm size categories, and for firm sector). Although estimates without individual fixed effects are a little less precise, results are extremely robust to these changes in specification.

Figure E.6. Robustness to changes in model specification



E.7.5 Randomization test

We subject our results to a randomization test as first introduced by Fisher (1935).[69] To do this, we take our sample of 35,580 firms, randomly draw firms with replacement, and then for each of these firms randomly assign a year to have a placebo automation event.[70] We then construct treated and control firms based on these placebo events. We repeat this procedure 100 times, where each permutation sample contains the same number of treated and control firms we have in our actual estimation sample.

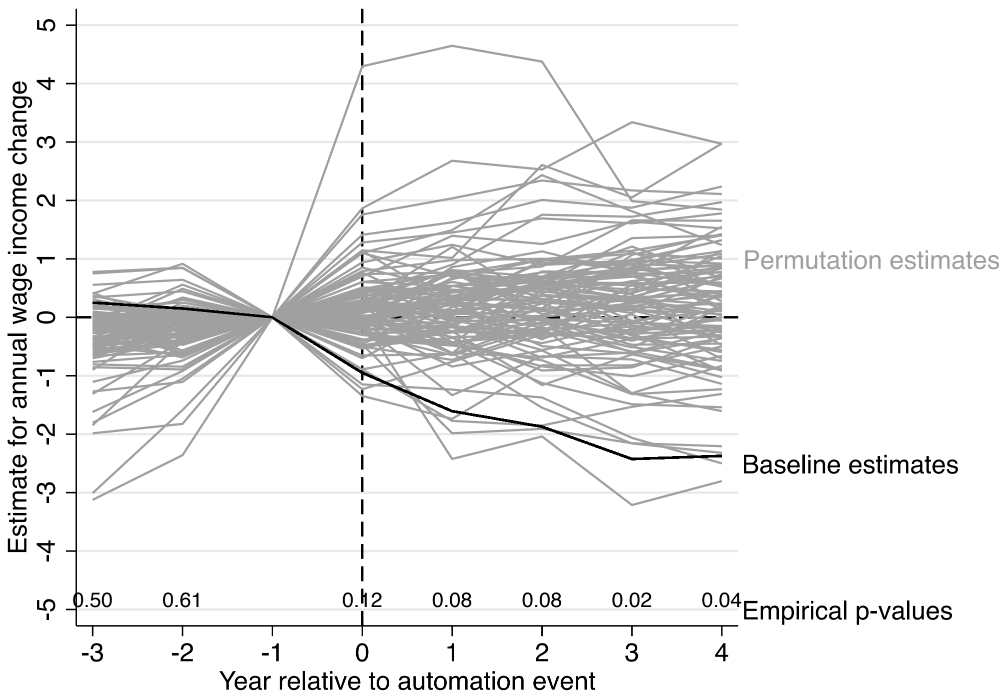
Results are shown in Figure E.7: each gray line presents a set of placebo (dynamic) treatment estimates, whereas the black line presents our actual treatment estimates. The

[69]Also see Kennedy (1995) for an overview and Young (2018) for a recent application and evaluation of the value of these tests.

[70]Note that this permutates both the assignment of treatment to firms, and their timing across years, since both are part of our empirical procedure.

graph also shows probability values calculated using the rank of the absolute value of our estimated coefficient among the 100 permutated estimates.[71] This shows that something at least as extreme as our treatment estimate is unlikely to occur by chance, increasing confidence that our estimates are not a statistical false positive.

Figure E.7. Randomization test



Notes: 100 permutations. The numbers printed at the bottom of the graph are probability values for the treatment estimates, based on the randomization test.

[71]Results are very similar when using t-statistics rather than coefficient estimates to calculate probability values.

F Computer investments

Table F.1 shows the distribution of automation costs and computer investment across firms and years, highlighting automation costs are higher than computer investments.

Tables F.3 and F.4 compare automation and computer investments per worker across firms by sector and firm size. As expected, Information and communication has the highest computer investment per worker, followed by Professional, scientific & technical activities. Accomodation & food serving and Construction have the lowest computer investment per worker. When considering the relative importance of automation and computer technology, Manufacturing is the most automation-intense compared to other sectors, whereas Information & communication is the most computer-intense. Like for automation, we generally see higher computer investment per worker for larger than smaller firms, but the pattern is less dramatic.

Table F.2 shows the distribution of computer investment spikes: while more firms have computer investment spikes than automation spikes, Figure F.1 shows that computer investment spikes are also characterized by a large one-time increase in investment.

Table F.1. Automation costs and computer investments distributions

	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	263,129	2,373	123,000	1,229
p95	620,724	3,839	273,263	2,039
mean	271,884	1,125	109,390	615
mean excl. zeros	377,959	1,564	170,810	960
N firms $\times$ years	171,878		171,878	
N firms $\times$ years with 0 costs	48,238		61,804	

Notes: All numbers are in 2015 euros. The number of observations is the number of firms times the number of years.

Table F.2. Automation and computerization events

	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 or 5	0.1	0.1

Notes: Overlapping sample of firms, N = 25,118.

Table F.3. Automation costs and computer investments by sector

Sector	Autom. cost per worker	Comp. inv. per worker	Ratio autom. to comp.	Nr of obs	
				Firms	*Firms × yrs*
Manufacturing	1,088	403	2.70	5,191	40,887
Construction	543	234	2.32	2,814	18,248
Wholesale & retail trade	1,257	594	2.12	7,230	50,471
Transportation & storage	999	496	2.01	2,283	15,868
Accommodation & food serving	279	165	1.69	742	4,462
Information & communication	2,214	2,713	0.82	1,563	9,762
Prof'l, scientific, & technical activities	1,381	844	1.64	2,376	14,830
Administrative & support activities	941	423	2.22	2,919	17,350

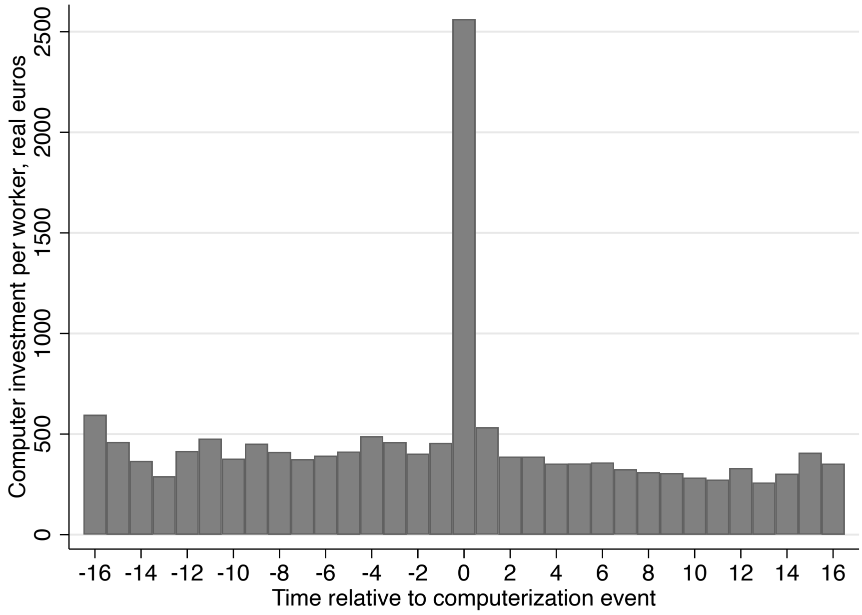
Notes: Overlapping sample, total number of firms is 25,118.

Table F.4. Automation costs and computer investments by firm size

Firm size	Autom. cost per worker	Comp. inv. per worker	Ratio autom. to comp.	Nr of obs	
				Firms	*Firms × yrs*
1–19 employees	2,433	1,193	2.04	2,260	11,352
20–49 employees	928	593	1.56	10,459	66,448
50–99 employees	914	497	1.84	5,873	41,560
100–199 employees	1,029	572	1.80	3,430	26,529
200–499 employees	1,314	621	2.12	1,929	16,218
≥500 employees	1,794	695	2.58	1,167	9,771

Notes: Overlapping sample, total number of firms is 25,118.

Figure F.1. Computer investment per worker around computerization events



Notes: Overlapping sample of firms with a computerization event. $N = 13,079$ in event year.