What Happens to Workers at Firms that Automate?*

James Bessen — *Boston University*
Maarten Goos — *Utrecht University*
Anna Salomons — *Utrecht University*
Wiljan van den Berge — *CPB*

July 2020

Abstract

We provide the first estimates of the impact of firms' decisions to automate on individual worker outcomes by combining Dutch micro-data with a direct measure of automation expenditures covering firms in all private non-financial sectors over 2000-2016. Using a novel differences-in-differences event-study design leveraging lumpy investment in automation technology, we find that automation by firms increases the probability of incumbent workers separating from their employers. Workers experience a 5-year cumulative wage income loss of 10 percent of one year's earnings, driven by decreases in days worked. This loss is only partially offset by various benefits systems, and older workers are more likely to enter early retirement. We document that these impacts are pervasive across private non-financial sectors of the economy, though income losses are larger for older workers, lower-paid workers, and workers employed at smaller firms. In contrast, no such losses are found for firms' investments in computers.

Keywords: Firm-level automation, Worker displacement

JEL: J23, J31, J62, J63, O33

*Helpful comments by Daron Acemoglu, David Autor, Peter Blair, Egbert Jongen, Guy Michaels, Pascual Restrepo, Bas van der Klaauw, and participants of seminars and conferences at ASSA 2020 annual meeting, CPB, the European Central Bank, Erasmus University Rotterdam, Groningen University, IZA, Maastricht University, NBER, TASKS V Bonn, Utrecht School of Economics, and University of Sussex are gratefully acknowledged. We are grateful to Pascual Restrepo for providing us with automation-related product codes for imports. James Bessen thanks Google.org for financial support.

1 Introduction

Advancing technologies are increasingly able to fully or partially automate job tasks. These technologies range from robotics to machine learning and other forms of artificial intelligence, and are being adopted across many sectors of the economy. Applications include selecting job applicants for interviews, picking orders in a warehouse, interpreting X-rays to diagnose disease, and automated customer service. These developments have raised concern that many workers are being displaced because their firms adopt automation technology (Eurobarometer 2017; Pew 2017).

An emerging literature studies firms adopting automation technology. The literature focuses predominantly on robotics, and measures adoption through imports, surveys, or electricity usage (Dinlersoz and Wolf, 2018; Cheng et al., 2019; Dixon et al., 2019; Humlum, 2019; Acemoglu et al., 2020; Aghion et al., 2020; Bonfiglioli et al., 2020; Koch et al., 2020).[1] Compared to non-adopters, firms adopting automation technology are generally found to experience faster employment, revenue, and productivity growth, and either declining or stable labor shares. However, Bonfiglioli et al. (2020) find that, when using variation in adoption intensity, instrumenting robot adoption, or using an event study methodology, employment in adopting firms declines.

There are as yet no empirical studies that directly examine what happens to individual workers when their firm decides to automate.[2] Firm-level impacts do not inform about the adjustment process that workers go through: even positive aggregate changes may be accompanied by worker churn, and economic consequences for individual workers directly affected by automation. Studying this adjustment process is critical to our understanding of how labor markets are impacted by the automation of work. Any such adjustments are also of first-order importance for policymakers aiming to diminish adverse impacts out of distributional concerns.

In studying the impacts of firms' automation activity on individual workers, our paper makes three contributions to the literature. First, we measure automation at the firm-level across all private non-financial sectors. We use data from an annual Dutch firm survey on automation costs. These automation costs refer to an official bookkeeping entry on firms' profit and loss ac-

[1] An related literature studies aggregate adoption of robotics, using industry and/or regional variation (Graetz and Michaels, 2018; Aghion et al., 2019; Dauth et al., 2019; Acemoglu and Restrepo, 2020).

[2] Existing studies on worker adjustments have used more aggregate sources of variation and do not always focus on causal effects. In particular, Cortés (2016) finds that workers switching out of routine-intense occupations experience faster wage growth relative to those who stay; Dauth et al. (2019) correlate regional variation in robot exposure with worker outcomes; while Edin et al. (2019) show that workers have worse labor market outcomes when their occupation is experiencing long-term decline.

count, and are defined as expenditures on third-party automation services provided by specialist integrators across a wide range of automation technologies. This means our measure complements the literature which has so far focused on the adoption of machinery (and particularly robots) by manufacturing firms.

Second, we develop a novel empirical approach based on evidence that firms' automation expenditures occur in discrete episodes of lumpy investment. In particular, we leverage spikes in automation cost shares to define automation events. We exploit the *timing* of these events in a difference-in-differences event-study methodology for identifying causal effects on firms' incumbent workers. This approach is made possible by the relatively high frequency of automation events in our data. The advantage of this approach is that it allows for more credible identification of worker-level effects: this is important because the literature has shown that automating and non-automating firms are on different employment trends, a finding we replicate in our data.

Third, by linking our annual firm survey to administrative firm and worker data, we can follow individual workers over time and measure a rich set of outcomes for workers in the years surrounding an automation event. These outcomes include annual wage income, daily wages, firm separation, days spent in non-employment, self-employment, early retirement, unemployment benefits and welfare receipts. To our knowledge, our paper provides the first estimates of the impacts of firm-level automation on these outcomes. Our data cover the years 2000-2016 and we observe 35,567 firms, employing close to 5 million unique workers per year on average.

We find that automation at the firm increases the probability of incumbent workers (with at least three years of firm tenure) separating from their employer. On average, these workers experience a 5-year cumulative wage income loss of 10% of an annual wage. These losses are driven by increases in non-employment: we do not find evidence of wage scarring for workers impacted by automation. Lost wage earnings from non-employment spells are only partially offset by various benefits systems and workers are more likely to enter early retirement. Wage losses are pervasive across different sectors of the economy, but largest for older workers, workers in lower (age-specific) wage quartiles, and those employed in small firms. Finally, we show that, unlike firm investments in automation, firm investments in computers do not result in displacement effects.

This paper is structured as follows. In section 2, we introduce our data, and compare our measure of automation costs to an import-based measure. Section 3 explains how the lumpiness of firms' automation expenditures is used to identify automation events, and documents these events in our data. Section 4 analyzes associations between automation and firm-level employment and wage outcomes, based on Bessen et al. (2020). Section 5 reports our main results on the effects of automation on worker-level outcomes using our difference-in-differences event study design. In section 6 we directly compare the worker-level impacts of automation to those of computerization. The final section concludes.

2 Data

We use Dutch data provided by Statistics Netherlands. In particular, we link an annual firm survey to administrative firm and worker databases covering the universe of firms and workers in the Netherlands. The firm survey is called "Production Statistics" and includes a direct question on automation costs – it covers all non-financial private firms with more than 50 employees, and samples a subset of smaller non-financial private firms.[3] This survey can be matched to administrative company and worker records.

Our data cover the years 2000-2016, and we retain 35,567 unique firms with at least 3 years of automation cost data – together, these firms employ around 5 million unique workers annually on average. We remove firms where Statistics Netherlands indicate that the data are (partly) imputed. We further remove workers enrolled in full-time studies, and those earning either less than 5,000 euros per year or less than 10 euros per day, as well as workers earning more than half a million euros per year or more than 2,000 euros on average per day. For workers observed in multiple jobs simultaneously, we only retain the job providing the main source of income in each year. We use a worker's total annual gross earnings in all jobs as the main measure of wage income. Since we observe the number of days but not hours worked per year, we use daily wages as a measure of wage rates.

We further observe workers' gender, age, and nationality. A downside to these data is that we neither observe workers' occupations nor their level of education: the former is unavailable entirely, whereas the latter is only defined for a small and selected subset of workers (with availability skewed toward younger and high-educated workers). We further match worker-level data to administrative records on receipts from unemployment, welfare, disability, and retirement benefits. We can track workers across firms on a daily basis, allowing us to construct indicators

[3]Firms are legally obliged to respond to the survey when sampled. However, the sampling design implies our data underrepresent smaller firms: we also examine effect heterogeneity across firm size classes to consider how this sample selection affects our overall findings.

Table 1. Automation costs

	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 × years		238,623			163,810	
N with 0 costs		31%			0%	

Notes: Automation cost level and per worker are reported in 2015 euros, automation cost share is calculated as a percentage of total costs, excluding automation costs. The number of observations is the number of firms times the number of years.

for firm separation and days spent in non-employment.

An important advantage of our data is the availability of a direct measure of automation at the firm level. In particular, automation costs are an official bookkeeping entry defined as expenditures on third-party automation services. While the disadvantage of this measure is that we do not know the exact automation technology being used by the firm, it does capture all automation technologies rather than focusing on a single one, and we measure it at the level of the firm rather than the industry, and across all private non-financial sectors. From discussions with company representatives and automation service providers, we know that these expenditures are related to automation technologies such as self-service check-outs, warehouse and storage systems, data-driven decision making, or automated customer service.

Table 1 shows summary statistics on annual automation costs for firms, both in levels, per worker, and as a percentage of total costs (excluding automation costs). This highlights several things. First, almost one-third of firm-year observations has zero automation expenditures. Second, the average automation cost share is 0.44%, corresponding to an outlay of around 200,000 euros annually, or 1,000 euros per worker. Third, this distribution is highly right-skewed as the median is only 0.16% – this skewness persists even when removing observations with zero automation costs.

Table 2 further shows how these automation costs and cost shares differ by broad (one-digit) sector. Our comprehensive measure of automation technologies indicates that all sectors have automation expenditures, though there is substantial variation at the firm level both between

Table 2. Automation costs by sector

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 N firms is 35,567; Total N firms × years is 238,623.

and within each of these sectors. Average expenditures at the sectoral level range from 245 to 1,789 euros per worker. The highest mean automation expenditures per worker are observed in Information & communication, followed by Professional, scientific & technical activities, Wholesale & retail, and Manufacturing. Conversely, Accommodation & food serving has the lowest expenditure per worker, followed by Construction, Administrative & support activities, and Transportation & storage. However, there is much variation between firms in the same sector, as shown by the standard deviations of the automation cost share in total (other) costs. While we do not use either this sectoral or between-firm variation in our empirical identification strategy, we will consider effect heterogeneity across sectors since the nature of automation technologies may be sector-specific.

Table 3 reports the same statistics but separately by firm size class, grouped into 6 classes used by Statistics Netherlands: the smallest firms have up to 19 employees whereas the largest have more than 500. Unsurprisingly, automation cost levels rise with firm size: firms with fewer than 20 employees spend around 12,000 euros annually on automation services, whereas the largest firms spend close to 3.2 million euros. Less obviously, this table also reveals that automation cost shares increase with firm size, particularly at the very top: the smallest firms have average automation cost shares of 0.40%[4], whereas firms with between 20 to 200 employees have a cost share of around 0.44%. This increases to 0.51% for firms between 200 and 500 workers, and 0.76% for firms with more than 500 workers. This is consistent with the literature's findings that more productive and therefore larger firms are more likely to automate (e.g. Bonfiglioli et al. 2020, Koch et al. 2020, Humlum 2019). However, there is substantial variation within size

[4]The relatively high expenditure per worker for the smallest firm size class is driven by a small number of one-person firms with high automation expenditures – when we eliminate the top-1% of observations in terms of automation cost per worker, outlays per worker are monotonically rising in firm size.

Table 3. Automation costs by firm size class

Firm size class	Total cost Mean	Cost per worker Mean	Cost per worker SD	Cost share (%) Mean	Cost share (%) SD	Nr of obs *Firms*	Nr of obs *Firms × yrs*
1-19 employees	12,270	921	14,571	0.40	1.30	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
≥500 employees	3,161,867	2,124	14,294	0.76	1.60	1,109	10,481

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

classes, also.[5]

We also observe average imports of robots and other automation technology for a subset of firms in our data[6]. Automation imports are defined as imports of intermediates classified by Acemoglu and Restrepo (2019) as automatically controlled machines, automatic transfer machines, automatic welding machines, numerically controlled machines, and (other) industrial robots. Net imports are defined as imports minus re-exports. Table 4 compares these two measures as a share of total operating cost for the overlapping subsample at the sector level, 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. However, automation imports and automation expenditures are correlated at the firm-level, as shown in Table 5 where firm-level automation expenditures are regressed onto (net) automation imports while controlling for firms' total operating cost. This comparison validates our measure as being correlated to the import measures often used in the literature, while also showing its broader coverage across sectors and firms.

[5] Appendix A.1 further illustrates how the distribution of automation cost expenditures per worker and automation cost shares change over time. It shows that mean automation costs per worker and mean automation cost shares have increased over time, and that there is a fanning out of these distributions with automation costs rising faster for higher percentiles.

[6] Consistent import data only start in 2010, so we construct firm-level averages and remove firms which cease operations before 2009 when comparing our automation cost data to automation imports. Details on import data can be found in Appendix A.2.

Table 4. Comparing automation costs to automation imports by sector

Sector	**Mean share in total costs of automation:** Expenditures	Imports	Net imports
Manufacturing	0.35	0.08	0.04
Construction	0.19	0.00	0.00
Wholesale & retail trade	0.30	0.06	0.05
Transportation & storage	0.35	0.12	0.09
Accommodation & food serving	0.27	0.00	0.00
Information & communication	0.81	0.00	0.00
Prof'l, scientific, & technical activities	1.00	0.01	0.01
Administrative & support activities	0.44	0.00	0.00

Notes: Total N firms is 30,280. Net automation imports are defined as imports minus re-exports.

Table 5. Comparing automation costs to automation imports at the firm level

Dependent variable: Automation costs, IHS transformed	(1)	(2)
Automation imports (IHS transformed)	0.028*** (0.008)	
Net automation imports (IHS transformed)		0.022** (0.008)
Log total costs	1.164*** (0.013)	1.169*** (0.013)

Notes: 30,280 observations. Automation costs, imports, and net imports are transformed using the inverse hyperbolic sine. Net automation imports are defined as imports minus re-exports.

3 Automation events

3.1 Defining automation events

We begin by defining automation cost spikes as follows. Firm j has an automation cost spike in year τ if its real automation costs $AC_{j\tau}$ relative to real total operating costs (excluding automation costs) averaged across all years t, $\overline{TC}_j$, are at least thrice the average firm-level cost share excluding year τ:

$$spike_{j\tau} = \mathbb{1}\left\{\frac{AC_{j,t=\tau}}{\overline{TC}_j} \geq 3 \times \frac{\overline{AC}_{j,t\neq\tau}}{\overline{TC}_j}\right\}, \tag{1}$$

where $\mathbb{1}\{\ldots\}$ denotes the indicator function. As such, a firm that has automation costs around one percent of all other operating costs for year $t \neq \tau$ will be classified as having an automation spike in $t = \tau$ if its automation costs in τ exceed three percent of average operating costs over years t. Finally, we define an automation event when the firm has its largest increase in automation cost share that qualifies as a spike.[7]

Rather than using the size of the increase in the automation cost share, we prefer to use this measure of automation events for a number of reasons. First, there may be measurement error in the survey variable making it more difficult to measure the exact size of the increase. Second, we use the automation costs survey variable to flag automation events, but other (indirect) costs may be incurred which are not directly surveyed: as such, our baseline approach identifies automation events without taking a strong stance on their exact size.

The existence of automation events is also highlighted in Humlum (2019) and would be consistent with a literature on lumpy investment (Haltiwanger et al. 1999; Doms and Dunne 1998). In fact, such events occur when the investment is irreversible and there are important indivisibilities. Under uncertainty, irreversibility creates an option value to waiting (Pindyck 1991; Nilsen and Schiantarelli 2003), whereas indivisibilities can arise from fixed adjustment costs (Rothschild 1971). Taken together, this implies investment occurs in relatively infrequent episodes of disproportionately large quantities. It is plausible that investments in automation meet these two criteria: major automation investments likely both include substantial irreversible investments (for example in terms of worker training or from developing custom software) as well as involve fixed adjustment costs from reorganizing production processes.

3.2 Summary statistics of automation events

Table 6 shows that out of the total number of 35,567 firms with at least 3 years of automation cost data, there are 10,422 firms that have at least one spike in their automation cost share between 2000 and 2016. That is, 29% of the firms in our sample spike at least once over the 17 years of observation.[8] Out of the firms that have at least one automation cost spike, the large majority spikes only once over 2000-2016, although some spike twice and up to five times at most.

Figure 1 shows the evolution of automation costs shares around automation events. This is constructed by redefining time t as event time: the difference between the actual calendar year ($year$) and the calendar year of the spike (τ), i.e. $t \equiv year - \tau$, such that all automation events line up in $t = 0$. Figure 1 uses the sample of 10,422 firms, each with their automation event in

[7] Note that firms without an automation event do not necessarily have zero automation costs: it is just that their automation expenditures do not fluctuate much as a percentage of total costs, implying they do not undergo automation events as we define them.

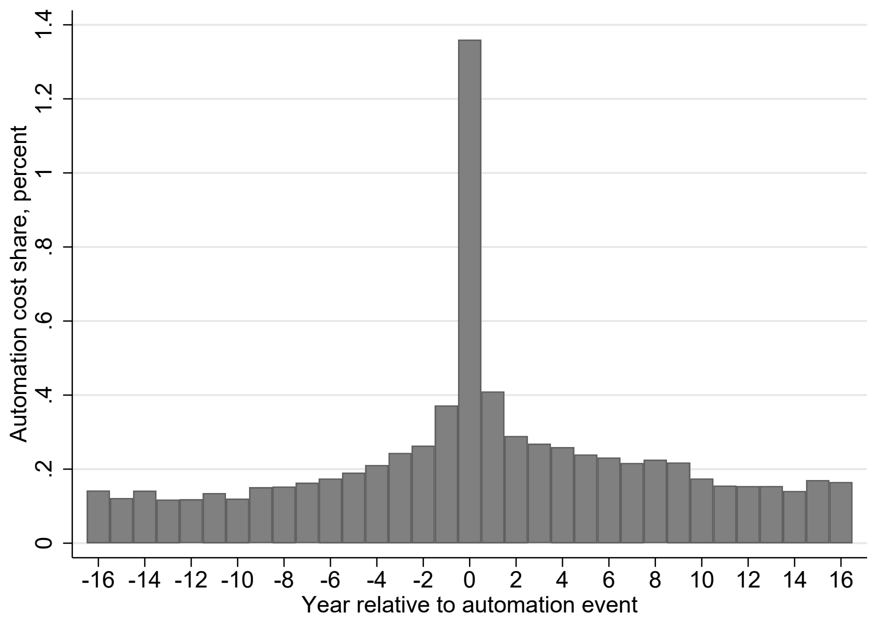
[8] We also find that this percentage varies relatively little across sectors, ranging from 25% in Construction to 39% in Information & communication.

Table 6. Spike frequency

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.

Figure 1. Automation cost shares around automation events



Notes: N=10,422 in $t = 0$.

$t = 0$.[9] Figure 1 shows a clear one-period increase in the automation cost share when a firm has its automation event.

[9]The number of observations for other years depends on the calendar years in which automation events occur, and on how often firms enter in the Production Statistics survey. For example, if an automation event occurs in the first calendar year of data, there are no observations for $t < 0$; if it took place in the last calendar year, there are no observations for $t > 0$.

4 Firm-level analyses

Following the recent literature examining changes in firm-level outcomes from firm-level automation that was summarized in the introduction, this section briefly discusses how our automation events correlate with firm-level outcomes, in particular firm-level employment and the firm's average daily wage. Section 4.1 compares firms with an automation event to firms without an automation event. Section 4.2 presents an automation event-study only using the sample of firms with an automation event.

4.1 Comparing automating to non-automating firms

We first ask how firms with an automation event differ from those without an automation event: as outcomes, we consider growth in firm-level employment and in the firm's average daily wage. In particular, we estimate variants of the following model:

$$\Delta \ln Y_{jyear} = \beta \times A_j + D_{year} + \gamma \times X_j + \varepsilon_{jyear}, \tag{2}$$

where the dependent variables are annual log changes in firm-level employment and the average daily wage for firm j in calendar year $year$. A_j is a dummy for the firm having an automation event over the 2000-2016 period, D_{year} are calendar year fixed-effects, and X_j additional controls consisting of two-digit sector dummies and baseline firm-level characteristics.[10] The term ε_{jyear} is an error term and standard errors are clustered at the firm-level.

The coefficient of interest, β, tells us whether automating firms experience different employment and mean daily wage trajectories. Since the existing literature focuses on manufacturing firms only, we additionally interact this with a manufacturing dummy to test whether this association is significantly different for automating firms in the manufacturing sector.

Table 7 shows that automating firms have 1.8 to 2.1% higher employment growth but not higher daily wage growth compared to non-automating firms. Further, these associations are not significantly different for manufacturing firms compared to non-manufacturing ones, as seen from the economically small and statistically insignificant interaction term coefficients. In sum, these results show that, on average, automating firms have faster employment growth[11] but not

[10] Appendix A.3 estimates a linear probability model where the dependent variable is a dummy for the firm having an automation event over 2000-2016, showing that larger firms and firms paying higher wages are more likely to automate. We include initial-year values for these variables as additional controls to capture convergence dynamics.

[11] Appendix A.4 shows similar results when using a balanced panel of firms.

Table 7. Firm-level outcomes for automating vs non-automating firms

	Δ log employment		Δ log mean daily wage	
	(1)	(2)	(3)	(4)
Automate	0.018**	0.021***	0.000	0.000
	(0.008)	(0.007)	(0.003)	(0.003)
Automate × manufacturing	-0.004	-0.009	-0.004	-0.001
	(0.012)	(0.011)	(0.005)	(0.004)
Additional controls	no	yes	no	yes

Notes: N = 167,490 firm-year observations, where 10,422 out of 35,567 unique firms automate. All models include calendar year fixed-effects and a dummy for the manufacturing sector. Additional controls are two-digit sector dummies 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.

faster wage growth across both manufacturing and non-manufacturing sectors. Appendix A.5 shows that similar effects are found for importers of automation technology compared to non-importers, although the employment growth differences are larger, reflecting that automation importers are a positively selected subset of the automating firms we identify.

4.2 Event-study only using automating firms

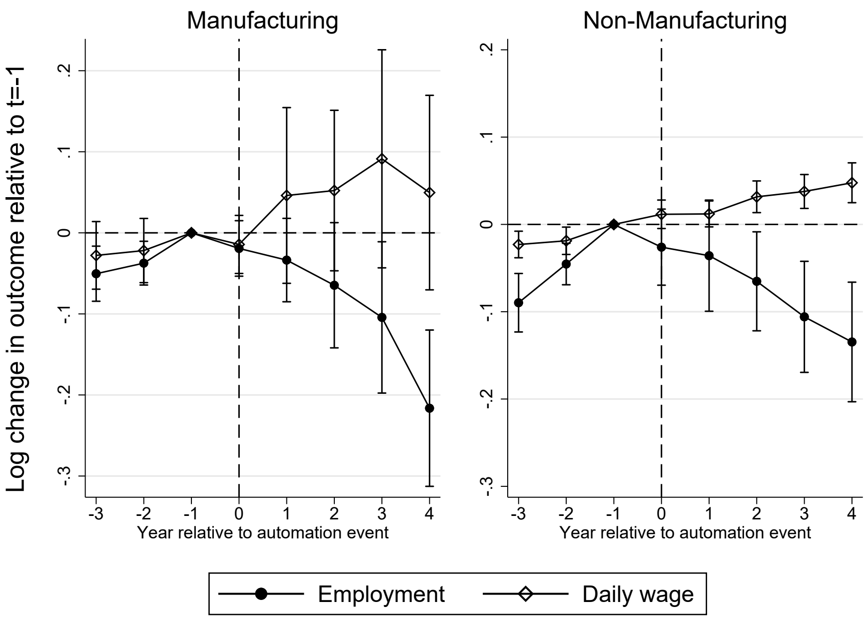
The employment growth for automating firms relative to non-automating ones does not rule out that automation at the firm level can be labor-saving when it occurs; such labor-saving effects would matter for individual workers employed in these automating firms. We therefore also consider the evolution of firm outcomes around an automation event, looking at the sub-sample of automating firms only. More formally, consider the following event-study equation:

$$\ln Y_{jt} = \sum_{t\neq -1; t=-3}^{4} \beta_t \times I_t + \gamma \times X_{jt} + \varepsilon_{jt}, \tag{3}$$

where j indexes firms, and t is defined as event time, i.e. calendar year ($year$) minus the calendar year in which the firm has an automation event (τ), i.e. $t \equiv year - \tau$. I_t are leads and lags for firms having an automation event in $t = 0$, with $t = -1$ as the reference category. Y_{jt} is j's employment or average daily wage, and X_{jt} are a set of controls, specifically calendar year and firm fixed-effects. We use all firms that have an automation event in $t = 0$ and that have no other spikes in automation cost shares in the event window $t \in [-3, 4]$.

Results are shown in Figure 2. Here, we see that employment appears to grow slower and

Figure 2. Firm-level outcomes for automating firms using event timing



Notes: All automating firms that exist in all years $t \in [-3, 4]$. N = 2,600 for manufacturing and N = 16,888 for non-manufacturing. Both models are weighted by firm-level employment size in $t = -1$. Standard errors are clustered at the firm-level. Whiskers reflect 95% confidence intervals.

even to contract following an automation event. Although these effects are quantitatively larger (albeit more imprecisely estimated) in manufacturing, they are observed in non-manufacturing firms as well. In contrast to employment growth, daily wages continue to grow during these events.

It is important to note that the firm-level estimates in Table 7 and Figure 2 are purely descriptive. However, these results highlight that to identify the causal impact of automation on firms' workers, as is this paper's aim, differences in employment trajectories between automating and non-automating firms need to be accounted for. Our identification approach is explained in the next section.

5 The impact of automation on individual workers

This section presents our main results: the impact of firm-level automation on individual workers. Section 5.1 outlines our difference-in-differences event-study design for identification of causal effects. In section 5.2 we present our main results, and in section 5.3 we test the robustness of these results.

5.1 Difference-in-differences event-study design

As before, we define t as event time, i.e. calendar year ($year$) minus the calendar year in which the firm has an automation event (τ) or $t \equiv year - \tau$, such that all automation events are lined up in $t = 0$. We then define the group of treated incumbent workers as those with 3 or more years of firm tenure in $t-1$ in treatment group firms, i.e. firms with an automation event in $t = 0$ and no other automation cost share spikes in the event window $t \in [-3, 4]$. Treated workers are further divided into cohorts by the calendar year in which their firm automates. Specifically, given that our sample covers calendar years 2000 to 2016, the earliest cohort of treated workers are those employed between 2000 and 2002 at a firm that automates in 2003. Similarly, the last cohort of treated workers are those employed between 2008 and 2010 in firms that automate in 2011.

For each cohort of treated incumbents, we then define a control group of incumbent workers with at least 3 years of firm tenure in $t-1$ and who are, in $t-1$, employed in control group firms, i.e. firms with their automation event in $t+5$ or later and without spikes in automation cost shares in the event window $t \in [-3, 4]$.[12] For example, the control group for the earliest cohort of treated workers are workers employed between 2000 and 2002 at a firm that has its automation event in 2008 or later. Similarly, the control group workers for the last cohort of treated workers are those employed between 2008 and 2010 at the same firm that automates in 2016.[13]

We use a difference-in-differences specification for each cohort of treatment and control group workers, with the data stacked across cohorts:

$$Y_{ijt} = \alpha + \beta \times treat_i + \sum_{t \neq -1; t=-3}^{4} \gamma_t \times I_t + \sum_{t \neq -1; t=-3}^{4} \delta_t \times I_t \times treat_i + \lambda X_{ijt} + \varepsilon_{ijt}, \qquad (4)$$

where i indexes individuals, j firms, and $t \in [-3, 4]$ is event time. Y_{ijt} is an individual-level outcome for worker i who must be employed at firm j in $t \in [-3, -1]$. Firm j can be a treatment group firm if j has an automation event in $t = 0$, or can be a control group firm if it has an automation event in $t+5$ or later.

Turning to the right-hand side of equation (4), $treat_i$ is a treatment indicator for worker

[12] We only require control group workers to be at a firm j that automates at $t+5$ or later to stay at firm j from $t = -3$ until $t = -1$. Hence, they do not have to be employed at firm j when firm j actually automates in year $t+5$ or later.

[13] Appendix A.6 provides more details on our sample construction.

i if her firm j has an automation event at $t = 0$. Further, I_t are event-time indicators, with $t = -1$ as the reference category. Lastly, X_{ijt} are controls: these are a set of worker (age and age squared, gender, and nationality) and firm (sector and firm-level employment at $t = -1$) characteristics as well as year fixed-effects. In our baseline specification, we replace $\beta \times treat_i$ with individual fixed-effects which absorbs non-time varying controls (gender, nationality, firm sector and employment at $t = -1$).[14] We cluster standard errors at the level where treatment occurs: that is, all workers employed at the same firm in $t - 1$ are one cluster.

In equation (4), the parameters of interest are δ_t for $t \geq 0$: these estimate the period $t \geq 0$ treatment effect relative to pre-treatment period $t = -1$ (given that I_{-1} is the reference category). For example, if automation leads to an immediate decrease in wage income that equals 1% of annual labor earnings in $t = -1$, we have that $\delta_0 = -0.01$. Similarly, if automation leads to an annual wage income loss in $t = 4$ that equals 3% of annual wage income in $t = -1$, we have that $\delta_4 = -0.03$. The figures in the next section plot estimates of δ_t for $t \in [-3, 4]$ for several worker-level outcomes: annual wage earnings, the hazard rate of leaving the firm, annual days in non-employment, annual benefit income and the probability to retire early or to become self-employed.

Estimates of δ_t can be interpreted as causal effects under the identifying assumption of parallel trends in the absence of automation events. Our empirical approach directly supports this assumption in several ways.

First, our specification strictly exploits differences in event timing rather than also using event incidence for identification. Only exploiting event timing across automating firms is important if firms without automation events are on different labor demand trajectories: the outcomes for workers employed at firms without an automation event are not an appropriate counterfactual. Effectively, we are matching workers on the firm-level outcome of having an automation event at some point in time. Our use of timing differences across firms is in the spirit of a recent literature exploiting event timing differences in other contexts (see e.g. Duggan et al. 2016; Fadlon and Nielsen 2017; Miller 2017; Lafortune et al. 2018).[15]

Second, our specification only considers incumbent workers, defined as those with at least 3 years of firm tenure in $t = -1$. On average across firms in our data 64% of workers are incumbents (where the median is 70%). This captures workers who have a stable working

[14] Except when we estimate the hazard rate of a worker leaving the firm.

[15] In support of our approach, Appendix A.8 provides some evidence that the timing of automation events cannot be easily predicted from observed firm-level characteristics in our data.

relation with their firm.[16] This is important because identification requires that workers are not selected into the firm in anticipation of an automation event occurring in the near future. This reasoning is similar to the focus on incumbent workers in the mass lay-off literature (e.g. see Jacobson et al. 1993; Couch and Placzek 2010; Davis and Von Wachter 2011).

Third, we further strengthen the assumption of parallel trends by matching on worker and firm observables to ensure that $\delta_t = 0$ for all $t < 0$ (Azoulay et al. 2010). In our baseline specification, 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 percentiles of real wage income, as well as separate bins for the 99th and 99.9th percentiles, in each of the three pre-treatment years $t = -3, -2, -1$. 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.

Finally, parallel trends in the absence of treatment requires that the results we find are not driven by concurrent events unrelated to automation. In section 5.3 below, we therefore check the robustness of our results by eliminating firms with other events occurring inside the event-window, including take-overs, acquisitions, firm splits, and restructuring.

5.2 What happens to workers at firms that automate?

We now turn to our main findings. First, we discuss the impact of automation on incumbents' average annual wage income and its components: the probability to leave the firm, days in non-employment, and daily wages conditional on being employed. We then discuss the impact of automation on benefit income, on the probability to retire early, and on the probability to become self-employed. Finally, we discuss effect heterogeneity across worker characteristics.

[16] Dutch labor law during almost our entire data period ensures that temporary contracts are of a maximum duration of 3 years, implying that workers with 3 years of tenure are very likely to have permanent contracts.

5.2.1 Annual wage earnings, firm separation, non-employment and daily wages

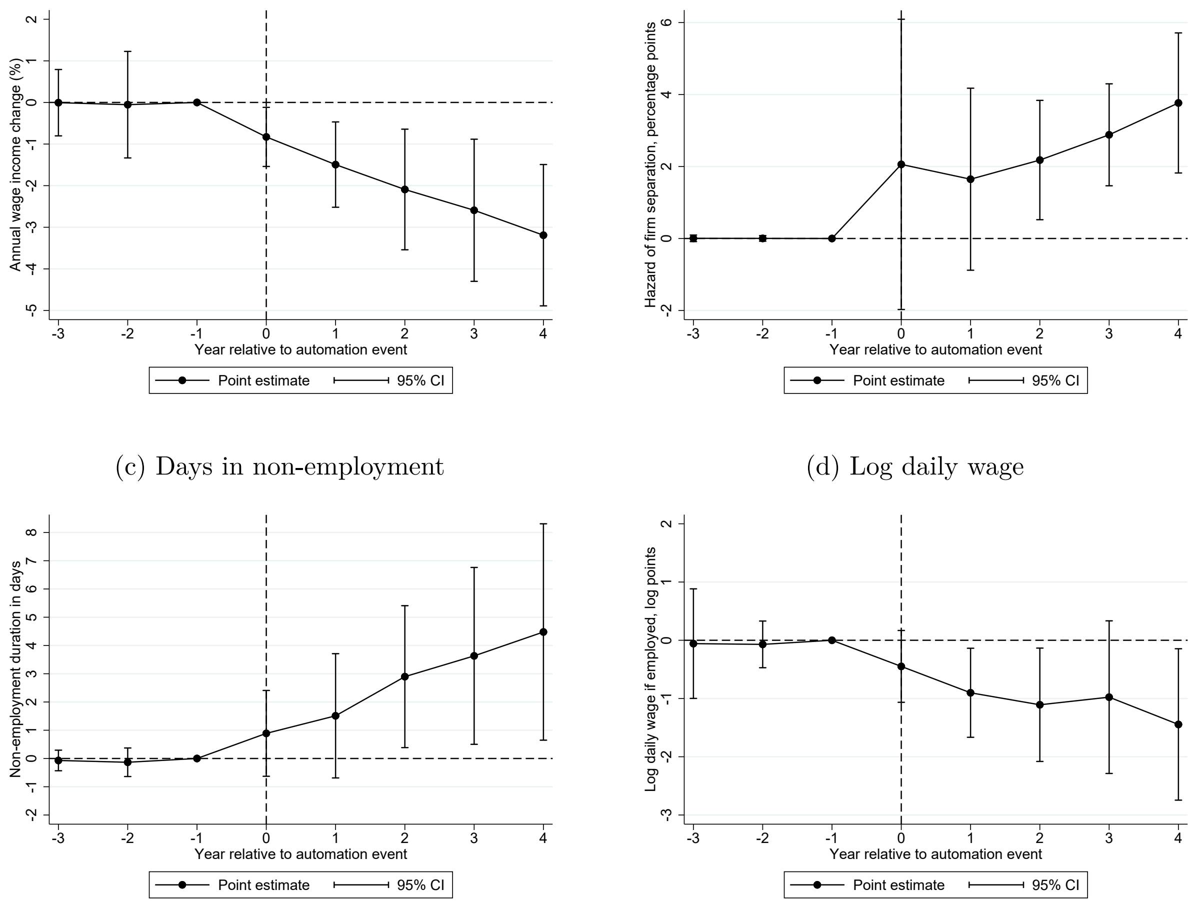
We begin by estimating the impact of firm-level automation on individual workers' real annual wage earnings, scaled by wage earnings levels in $t = -1$ to obtain relative impacts. These estimates of parameters δ_t in equation (4) are shown in panel (a) in Figure 3, multiplied by 100 to capture percentage changes in annual real wage earnings relative to their $t = -1$ level.

The estimates highlight that incumbent workers lose wage income as a result of the automation event. Indeed, the average incumbent worker loses about 1% of annual wage earnings in $t = 0$; 1.5% (of annual wage earnings in $t = -1$) in $t = 1$; 2% in $t = 2$; 2.5% in $t = 3$; and 3% in $t = 4$. Overall, automation decreases annual wage earnings for an incumbent worker by a cumulative 10% (=1%+1.5%+2%+2.5%+3%) of her annual wage earnings in $t = -1$ after 5 years. Given that annual wage earnings grow by 1.6% annually on average, this reflects a non-negligible loss compared to usual wage earnings trajectories. In euros, this 10% annual earnings loss corresponds to a cumulative earnings loss of around 4,000 euros for the average incumbent worker over the 5 years following her firm's decision to automate.

These losses in annual earnings from work may be driven by changes in days worked following firm separation, changes in daily wages if employed, or a combination of both. To estimate the importance of firm separation, panel (b) in Figure 3 presents estimates from equation (4) where the dependent variable is the worker's hazard rate of separating from her pre-treatment employer. All coefficients have been multiplied by 100, such that the effects are in percentage points. The panel shows that automation leads to some incumbent workers leaving the firm: after 5 years, incumbent workers at automating firms have a statistically significant 3.9 percentage point higher hazard rate of firm exit. The average hazard rate for treatment group workers in $t = 4$ is 12.5%, which is 41% higher compared to the corresponding hazard among control group workers of 8.8%.

It is noteworthy that worker displacement does not occur instantly: rather, displacement effects arise over time. There are various (and non-mutually exclusive) possible explanations for this. For one, these patterns are consistent with incumbent workers having open-ended contracts, making it costly to fire them. Further, these gradual changes could in part also result from a time delay in the effective implementation of automation technologies relative to the cost outlay, or because it takes time for workers and firms to learn about changes to their match quality under the new technology. Gradual displacement following an automation event is in

Figure 3. What happens to workers at firms that automate?



(a) Annual wage income

(b) Hazard rate of leaving the firm

(c) Days in non-employment

(d) Log daily wage

Notes: N=8,142,568 for annual real wage income and days in non-employment, N=7,469,702 for hazard of firm separation, N=7,918,209 for log daily wages. Whiskers represent 95 percent confidence intervals.

contrast to the well-documented phenomenon of mass lay-off events, where at least 30% of the firm's incumbent workforce is laid-off at once.[17]

Although our results so far show that automation leads to an increase in firm separation, this need not translate to losses in annual wage income if displaced workers find re-employment quickly (and at similar wage rates): we now turn to impacts on non-employment. Results are shown in panel (c) in Figure 3, where we define the dependent variable in equation (4) as the annual number of days spent in non-employment. Starting in the automation event year ($t = 0$), non-employed days for treated workers gradually increase relative to control group workers. In particular, non-employment increases by 1 day in the automation event year (although this estimate is not statistically significant), which increases to around 4.5 days annually after 5 years, with a total cumulative increase in non-employment of around 13 days compared to the

[17]See Davis and Von Wachter (2011) for an overview.

control group. By comparison, in the event year, matched control group incumbents spend around 5.7 days in non-employment on average, suggesting automation produces an average increase of 22% in non-employment days in the automation year itself. The cumulative five-year impact also corresponds to a 22% average increase relative to the five-year cumulative non-employment duration (82 days) experienced by control group incumbents.

Finally, we do not find strong evidence to support that automation affects incumbents' wages conditional on employment, as shown in the final panel of Figure 3. Recall that we do not observe daily hours worked in our data: changes in daily wages can therefore result from changes in hourly wages and/or changes in daily hours worked. The absence of strong daily wage effects implies that the decrease in annual wage income for incumbent workers when their firm automates is largely driven by the observed rise in non-employment spells. This absence of wage scarring effects is in contrast to displacement effects from mass lay-offs or firm closures Jacobson et al. 1993; Couch and Placzek 2010; Davis and Von Wachter 2011, which have been found in the Netherlands as well.[18]

5.2.2 Annual benefit income, early retirement and self-employment

Panel (a) in Figure 4 considers the impact of automation on incumbent workers' real annual benefit income (in euros), comprised of unemployment benefits, disability benefits, and welfare payments. We find that incumbent workers receive more benefit income following an automation event: after 5 years, the cumulative total amount received is around 600 euros on average. Given that the average annual wage income loss cumulates to about 4,000 euros after 5 years (see above), this implies that only 15% of the wage earnings loss from automation is offset by benefit payments. This finding is comparable to that for other worker displacement events, where typically only a small part of the average negative impact on earnings is compensated by social security (Hardoy and Schøne 2014).

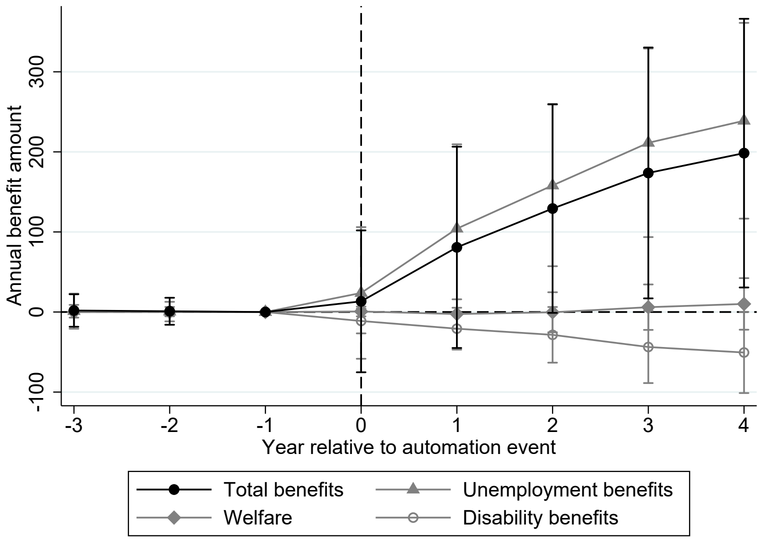
Panel (a) in Figure 4 further shows that all of the additional benefit payments arise from unemployment insurance: this is expected, as unemployment benefit eligibility is very high among workers with at least three years of firm tenure.[19] Consistent with high unemployment benefit eligibility, we do not see any increase in welfare payments. Lastly, disability benefits

[18]See Deelen et al. 2018; Mooi-Reci and Ganzeboom 2015 who find evidence of wage scarring after mass lay-offs in the Dutch context, using the same administrative data as we do here.

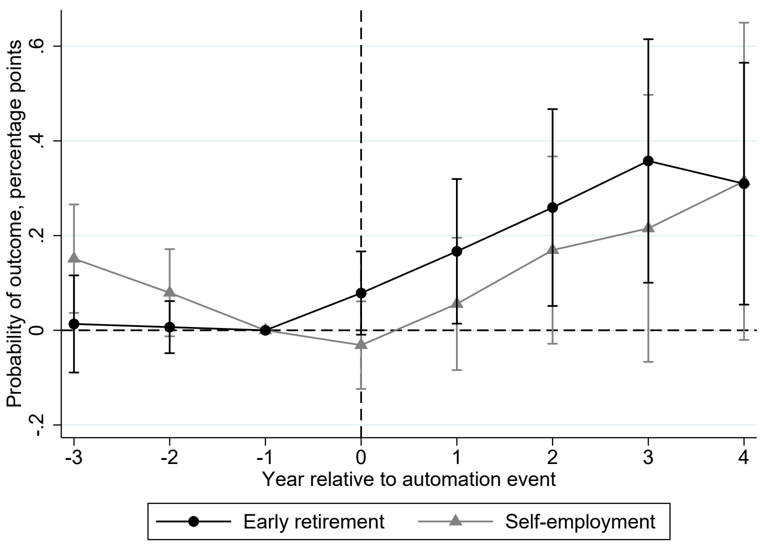
[19]From 2000 to 2015, eligible workers in the Netherlands were entitled to up to 38 months of unemployment benefits following job loss. Since 2016, maximum eligibility is 24 months.

Figure 4. Annual benefit income and early retirement

(a) Annual benefit income



(b) Early retirement and self-employment



Notes: N=8,142,568. Whiskers represent 95 percent confidence intervals.

are actually slightly decreasing over time. Since all incumbents were previously employed, this implies some were receiving benefits for a partial disability prior to the automation event, but they no longer receive these benefits with their new employer.

Panel (b) in Figure 4 examines whether automation also has an effect on early retirement, defined as the receipt of retirement benefits prior to reaching the legal retirement age. In particular, 5 years after the automation event, treated incumbent workers are 0.3 percentage points more likely to be observed in early retirement. While this effect might seem small in size, the average probability of early retirement among control-group incumbents in $t = 4$ is around 1.7%. As such, the treatment effect represents an 18% increase in the incidence of early retirement. Besides early retirement, the figure also examines the possibility that displaced workers enter self-employment. Estimates suggest that an incumbent's probability of becoming self-employed has increased by 0.3 percentage points after 5 years. However, this is only a 6% increase relative to probability of 5.2% among the control group workers in $t = 4$. This means self-employment is unlikely to be an important compensating income source.

5.2.3 Effect heterogeneity

We briefly discuss effect heterogeneity by worker characteristics, and compare our results for incumbent workers to estimates for recent hires: these results are laid out in more detail in Appendix A.9.

First, we consider effect heterogeneity by incumbent worker characteristics: age, gender, (initial) firm sector and employment size, and a worker's wage quartile in her age group as a proxy for skill. Overall, we find that effects are quite pervasive across different worker groups: this implies our results are not driven by a specific subsample of workers. While we do not always have enough statistical power to distinguish effects, older workers, workers employed in smaller firms, and lower-paid workers (for the same age) seem to experience somewhat larger wage income losses from automation than their counterparts.

Second, we compare our results for incumbent workers to estimates for recent hires, defined as those with less than three years of firm tenure prior to the automation event. Unlike for incumbent workers, we find no income losses from automation for the average recent hire, nor any statistically significant increase in firm separation and non-employment duration. This could be the case because 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.

5.3 Robustness tests

Our estimates can be interpreted as causal effects only under the identifying assumption of parallel trends in the absence of automation events. A simple falsification test is to see whether $\delta_t = 0$ for $t < 0$, as seems to be the case in Figures 3 and 4. To go further, we can additionally match workers on their firms' pre-treatment employment trends. This implies we now ensure that treated and control workers are not only employed at firms that experience an automation event at some point in time, but where pre-treatment employment growth is similar.

Another important concern is that automation events could coincide with other firm-level events. Our data include administrative information on several other important firm-level events, namely mergers, take-overs, acquisitions, firm splits, and restructuring.[20] As a second robustness check, we therefore eliminate firms that experience such events anywhere in the event window.

As a third robustness check, we remove outlier firms in terms of employment changes (those experiencing an employment change exceeding 90 percent in any one year), both in the event window and outside of it. The removal of these outliers is intended to capture any firm-level events which are not formally documented in our administrative records.

Fourth, we also remove firms where there was a new worker among the firm's top-decile annual wage income earners[21] in the three years prior to the automation event. This is intended to capture automation events coinciding with managerial change, which may bring changes in personnel policy unrelated to automation.

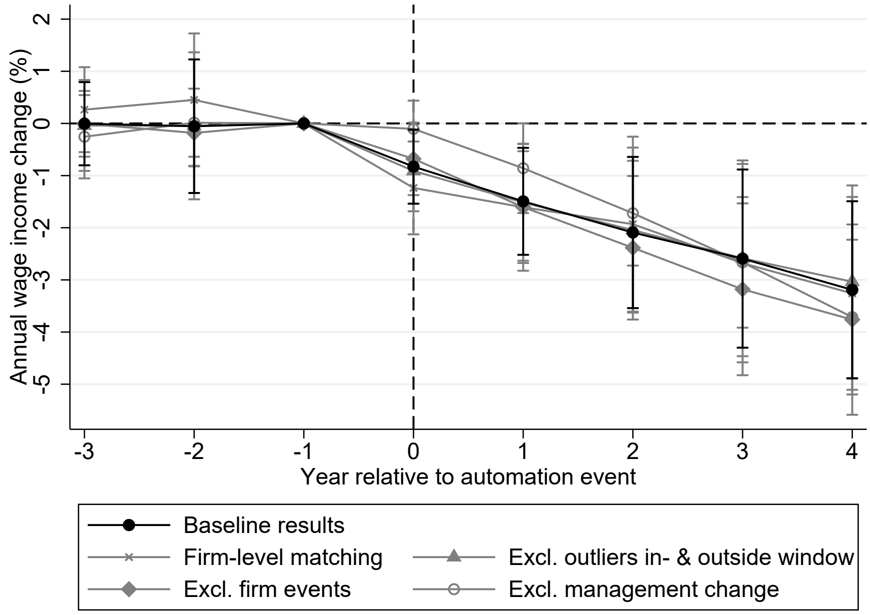
Figure 5 summarizes our estimates for relative annual wage income for all four robustness checks pertaining to firm-level pre-trends and events (along with our baseline estimates from panel (a) in Figure 3). Estimates are very similar across the board, though effects are somewhat smaller when eliminating firms with (suspected) management change: this suggests that automation may sometimes be the result of a new manager changing business practices. Overall, however, our findings are very robust, showing that firm-level events other than automation are unlikely to be the driving force behind the worker impacts we find.

Appendix A.10 reports results from three further robustness tests: 1) using alternative definitions of automation events, 2) changing the model specification, and 3) performing a randomization test as first introduced by Fisher (1935). Also here, we find that our results are robust.

[20]We additionally observe firm births and deaths, but these are already excluded since we consider a balanced sample of firms over the event window: we do allow firm births in the first year of observation, however.

[21]Conditional on this worker earning at least 150 euros a day, i.e. 40,000 euros a year.

Figure 5. Robustness tests



6 Automation versus computerization events

We have found that automation displaces incumbent workers: this raises the question whether this effect is specific to automation technologies or occurs with investment in new technology more generally.

Statistics Netherlands also conducts a separate and partially overlapping firm survey on investments, including computer investments.[22] This item is called ‘computers’ or ‘computers and other hardware’ (depending on the year) and consistently defined as follows: “All data-processing electronic equipment insofar as they can be freely programmed by the user, including all supporting appliances. Do not include software.”. All investment within the company counts towards the expenditures, also if the equipment is second-hand, or leased or rented, or produced within the company. It excludes investments in plants that are located abroad or resulting from take-overs of other organizations whose operations are continued without change.

In this section, we analyze the effects of computer investments in a similar way to that of automation events, and directly contrast it to the impacts of automation in the part of the sample where we have overlapping data. This serves two purposes. First, we can consider to what extent spikes in automation costs have different effects on workers than do spikes in

[22]Investments in software and in communication equipment are only measured from 2012 onwards, so we only consider computer investments. In 2012, software investments are of a similar magnitude as computer investments.

Table 8. 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	263,000	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 × yrs	171,797		171,797	
N firms × yrs with 0 costs	48,220		61,773	

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

computer investment. Second, if automation cost expenditures and computer investments are correlated at the firm level, we can remove firms which have computer investment spikes within our estimation window to rule out that our automation event is partially capturing investment in computers. Conversely, we will also estimate the effects of computer investments in isolation, that is, excluding any events where automation events occur within the estimation window.

We first show some summary statistics on computer investments (section 6.1) before turning to the comparison between automation and computerization (section 6.2). Throughout, we consider the overlapping sample of firms where we observe both automation events and at least three years of computer investment data. This means our dataset consists of 25,107 instead of 35,567 firms, and is more skewed towards larger firms as these are most likely to be sampled in both surveys.

6.1 Defining computerization events

Table 8 informs on the distribution of automation costs and computer investment across firms and years.[23] Automation costs are higher than computer investments across the distribution, both in total and per worker. Of course, it should be noted that both can come with other unmeasured correlated costs, such as software for computers, and machinery for automation.

In order to compare automation to computerization events, we construct computer investment spikes in the same way we have for automation, but using computer investment per

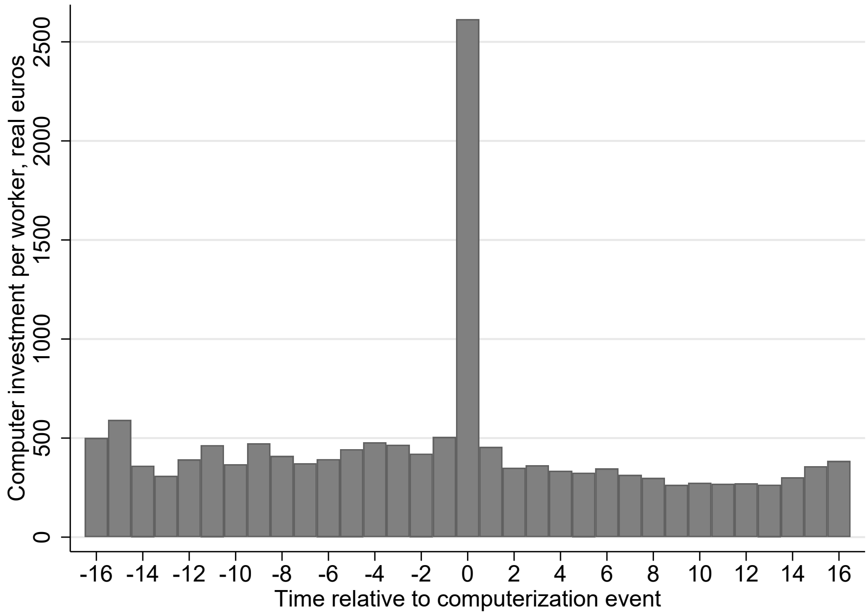
[23] Appendix A.11 provides further details of computer investments by sector and firm size.

Table 9. Automation costs and computer investments spikes

	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.

Figure 6. Computer investment per worker around computerization events



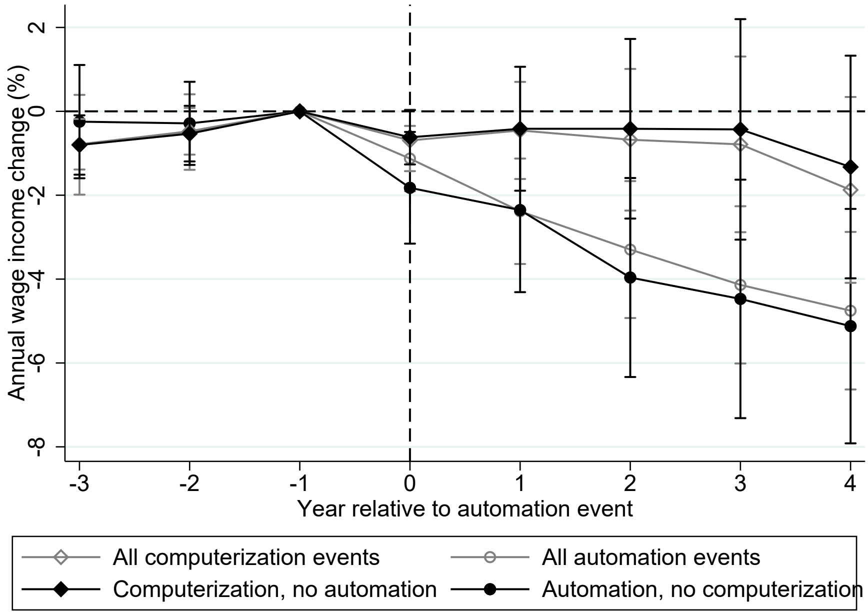
Notes: Overlapping sample of firms with a computerization event. N=13,081.

worker.[24] We use the same threshold, assigning firms a computer investment spike if their computer investment per worker exceeds three times their usual level.

The resulting distribution of computer investment spikes is reported in Table 9. Compared to automation spikes, computer investment spikes are more frequent. However, similar to Figure 1 for automation events, Figure 6 shows a clear one-period increase in computer investments per worker when a firm has its computerization event: in the event year, treated firms spend around 2,500 euros per worker, compared to around 400 euros in the years before and after.

[24]We use computer investments per worker because, unlike automation expenditures, computer investments are not part of total costs; and because total investments cannot serve as a denominator because they are inconsistently defined over our sample period.

Figure 7. Relative annual wage earnings effects of automation and computerization



Notes: All estimates are for the overlapping sample where we observe data on both automation costs and computer investments. N=10,217,088 for all computerization; N=7,783,929 for all computerization excluding automation; N=8,110,456 for all automation; and N=4,632,880 for automation excluding computerization.

6.2 Comparing effects of automation to computerization events

After restricting the overlapping sample further to firms that exist in all years in their computerization event window (as we also did for automation events), we construct four different datasets. First, we consider automation events and computerization events in isolation: that is, we identify treated and control group workers for one type of event while ignoring the other. This allows us to estimate our difference-in-differences event-study for automation and computerization separately. However, these two events are correlated across firms over time: that is, firms that have recently had one type of event are more likely to also experience the other sometime soon – sometimes even in the same year. This implies any estimated impact of automation may be contaminated by computerization, and vice versa. We therefore construct two additional samples of events which occur in isolation: that is, we only retain those automation (computerization) events where there is no computerization (automation) event occurring in the estimation window for either treated or control group firms. For each of the four samples, we then estimate equation (4) and report results in Figure 7.

This comparison leads to several findings. First and foremost, computerization does not lead to wage earnings losses for incumbent workers: estimates are small and never statistically significant. This is in contrast to automation, which does lead to income losses for the average incumbent worker. Further, the income losses of automation are (slightly) larger when removing concurrent computerization, and the effects of computerization on income are (slightly) smaller when removing concurrent automation. Consistent with these results, we do not find any increase in firm separation or non-employment duration for workers impacted by computerization. This means that automation is a more labor-displacing force than computerization from the perspective of a firm's incumbent workers.

7 Conclusion

We provide the first estimate of the impacts of firms' automation activities on individual workers, using firm-level data on automation expenditures across all non-financial private sectors in the Netherlands over 2000-2016. Leveraging a novel differences-in-differences event-study design, we show that automation at the firm significantly increases incumbent workers' hazard of separating from their employers. This finding of course does not imply that automation destroys jobs on net in automating firms or in the economy as a whole. However, we show that automation can be a disruptive process for firms' incumbent workers, leading to job churn and non-negligible adjustment costs. Specifically, on average, these workers experience a 5-year cumulative wage income loss of 10 percent of one year's earnings, driven by decreases in days worked. While we do not find evidence of substantial wage scarring, wage income losses are only partially offset by various benefits systems, and older workers are more likely to enter early retirement. We document that these impacts are quite pervasive across different sectors of the economy, though income losses are larger for older workers, lower-paid workers, and workers employed at smaller firms. In contrast, we do not find evidence that incumbent workers face similar job and income losses from firms' investments in computer technology. This suggests that, from the perspective of incumbent workers, automation is (currently) a more labor-displacing force.

References

Acemoglu, D., Lelarge, C., and Restrepo, P. (2020). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment. *American Economic Association Papers and Proceedings*, 110:383–388.

Acemoglu, D. and Restrepo, P. (2019). Demographics and Automation.

Acemoglu, D. and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, 128(6):2188–2244.

Aghion, P., Antonin, C., and Bunel, S. (2019). Artficial Intelligence, Growth and Employment: The Role of Policy. *Economics and Statistics*, 510-511-512:149–164.

Aghion, P., Antonin, C., Bunel, S., and Jaravel, X. (2020). What Are the Labor and Product Market Effects of Automation? New Evidence from France.

Azoulay, P., Graff Zivin, J. S., and Wang, J. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, 125(2):549–589.

Bessen, J., Goos, M., Salomons, A., and van den Berge, W. (2020). Firm-Level Automation: Evidence from the Netherlands. *American Economic Association Papers & Proceedings*, Forthcoming.

Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). CEM: Coarsened Exact Matching in Stata. *Stata Journal*, 9(4):524–546.

Bonfiglioli, A., Crinò, R., Fadinger, H., and Gancia, G. (2020). Robot Imports and Firm-Level Outcomes.

Brier, G. W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, 78:1–3.

Cheng, H., Jia, R., Li, D., and Li, H. (2019). The Rise of Robots in China. *Journal of Economic Perspectives*, 33(2):71–88.

Cortés, G. (2016). Where Have the Middle-Wage Workers Gone? A Study of Polarization using Panel Data". *Journal of Labor Economics*, forthcoming.

Couch, K. A. and Placzek, D. W. (2010). Earnings Losses of Displaced Workers Revisited. *American Economic Review*, 100(1):572–89.

Dauth, W., Findeisen, S., Südekum, J., and Wößner, N. (2019). The Adjustment of Labor Markets to Robots.

Davis, S. J. and Von Wachter, T. (2011). Recessions and the Costs of Job Loss. *Brookings Papers on Economic Activity*, 42(2):1–71.

Deelen, A., de Graaf-Zijl, M., and van den Berge, W. (2018). Labour Market Effects of Job Displacement for Prime-age and Older Workers. *IZA Journal of Labor Economics*, 7(1):3.

Dinlersoz, E. and Wolf, Z. (2018). Automation, Labor Share, and Productivity: Plant-Level Evidence from U.S. Manufacturing. Working Papers 18-39, Center for Economic Studies, U.S. Census Bureau.

Dixon, J., Hong, B., and Wu, L. (2019). The Employment Consequences of Robots: Firm-Level Evidence. Ssrn discussion paper, SSRN.

Doms, M. E. and Dunne, T. (1998). Capital Adjustment Patterns in Manufacturing Plants. *Review of Economic Dynamics*, 1(2):409–429.

Duggan, M., Garthwaite, C., and Goyal, A. (2016). The Market Impacts of Pharmaceutical Product Patents in Developing Countries: Evidence from India. *American Economic Review*, 106(1):99–135.

Edin, P.-A., Evans, T., Graetz, G., Hernnäs, S., and Michaels, G. (2019). Individual Consequences of Occupational Decline. Working paper.

Eurobarometer (2017). Attitudes Towards the Impact of Digitisation and Automation on Daily Life. Technical report, European Commission, Directorate-General for Communication, Special Eurobarometer 460.

Fadlon, I. and Nielsen, T. H. (2017). Family Health Behaviors. Working Paper 24042, National Bureau of Economic Research.

Fisher, S. R. A. (1935). *The Design of Experiments*. Macmillan.

Graetz, G. and Michaels, G. (2018). Robots at Work. *Review of Economics and Statistics*, forthcoming.

Haltiwanger, J., Cooper, R., and Power, L. (1999). Machine Replacement and the Business Cycle: Lumps and Bumps. *American Economic Review*, 89(4):921–946.

Hardoy, I. and Schøne, P. (2014). Displacement and Household Adaptation: Insured by the Spouse or the State? *Journal of Population Economics*, 27(3):683–703.

Humlum, A. (2019). Robot Adoption and Labor Market Dynamics. Technical report, Mimeo Princeton University.

Iacus, S. M., King, G., and Porro, G. (2012). Causal Inference Without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20(1):1–24.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings Losses of Displaced Workers. *The American Economic Review*, pages 685–709.

Kennedy, P. E. (1995). Randomization Tests in Econometrics. *Journal of Business & Economic Statistics*, 13(1):85–94.

Koch, M., Manuylov, I., and Smolka, M. (2020). Microeconomic implications of robot adoption: A firm-level perspective.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (2018). School Finance Reform and the Distribution of Student Achievement. *American Economic Journal: Applied Economics*, 10(2):1–26.

Miller, C. (2017). The Persistent Effect of Temporary Affirmative Action. *American Economic Journal: Applied Economics*, 9(3):152–90.

Mooi-Reci, I. and Ganzeboom, H. B. (2015). Unemployment Scarring by Gender: Human Capital Depreciation or Stigmatization? Longitudinal Evidence from the Netherlands, 1980–2000. *Social Science Research*, 52:642–658.

Nilsen, O. A. and Schiantarelli, F. (2003). Zeros and Lumps in Investment: Empirical Evidence on Irreversibilities and Nonconvexities. *The Review of Economics and Statistics*, 85(4):1021–1037.

Pew (2017). Automation in Everyday Life. Technical report, Pew Research Center.

Pindyck, R. (1991). Irreversibility, Uncertainty, and Investment. *Journal of Economic Literature*, 29(3):1110–48.

Rothschild, M. (1971). On the Cost of Adjustment. *The Quarterly Journal of Economics*, 85(4):605–622.

Young, A. (2018). Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results. *The Quarterly Journal of Economics*, 134(2):557–598.

A Appendices

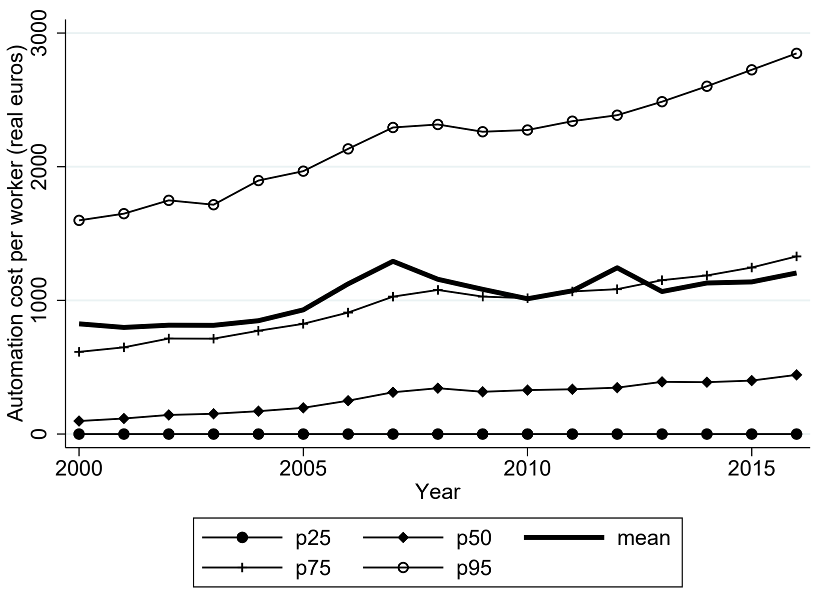
Appendices occur in the order as they are referred to in the main text and are for online publication only.

A.1 Automation costs over time

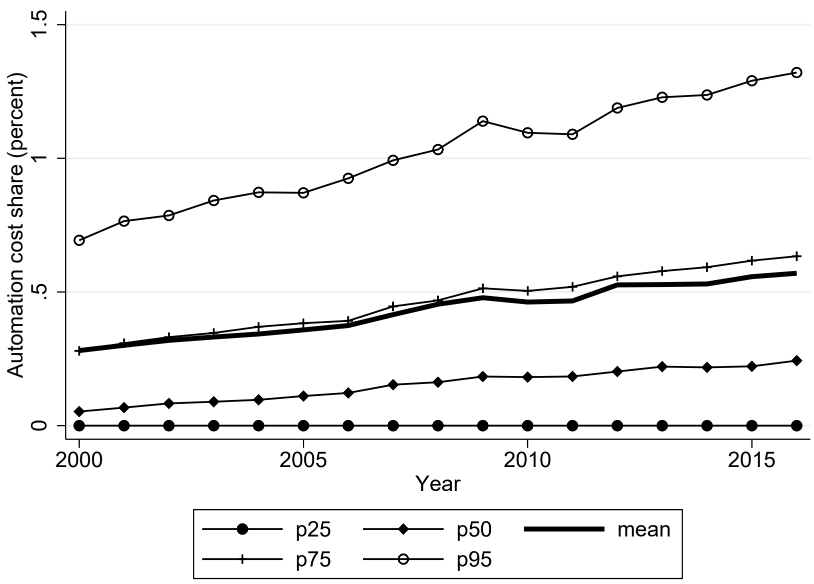
Figure A.1 shows how the distribution of automation costs per worker (top panel) and automation cost shares (bottom panel) have changed over time. Mean automation costs per worker and the mean automation cost share are rising over 2000-2016. Furthermore, besides increases in means, there is a fanning 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 Automation import data

We obtain data on firms' imports and re-exports of intermediates from Statistics Netherlands: unlike our automation cost measure, which starts in 1999, we can only identify these trade variables from 2010 onwards. Following the literature, we construct a firm-level indicator for being an importer (or net importer) of automation machinery using CN-2018 product codes. We deflate values to 2015 constant euros.

We follow the categorization of Acemoglu and Restrepo (2019) 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

A.3 Predicting automation events

Table A.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 we observe having automation events are different from those that have not. In particular, the probability of having an automation event is higher for firms with younger workers and with a higher fraction of men, firms that pay higher wages, lager firms, and firms in Information & communication, Professional, scientific & technological activities, Transportation & storage, and Transportation & storage.

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

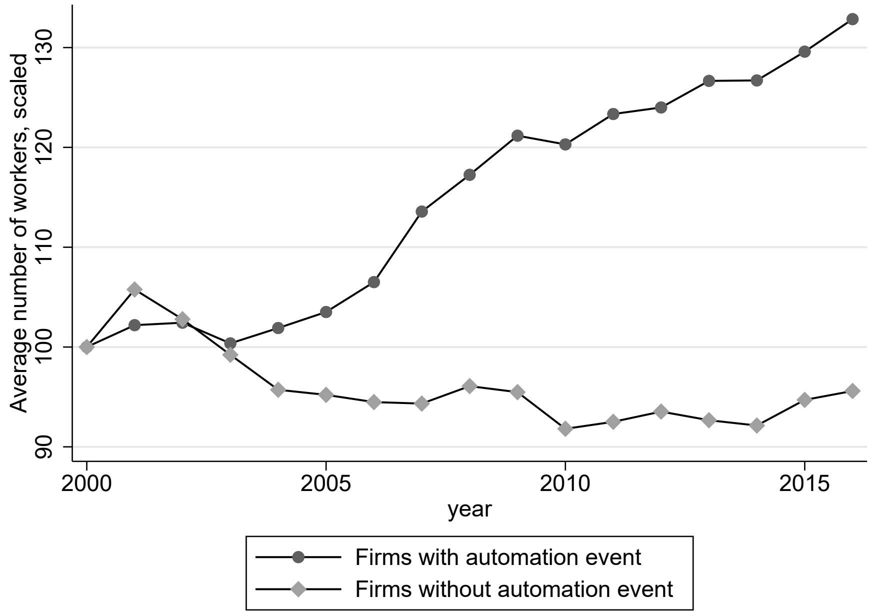
Mean worker age	-0.0027***	Manufacturing	*reference*
	(0.0001)		
Share of women	-0.0320***	Construction	-0.0261***
	(0.0035)		(0.0026)
Mean real annual wage / 1,000	0.0007***	Wholesale & retail trade	-0.0042*
	(0.0000)		(0.0022)
1-19 employees	*reference*	Transportation & storage	0.0312***
			(0.0028)
20-49 employees	0.0470***	Accommodation & food serving	-0.0303***
	(0.0016)		(0.0043)
50-99 employees	0.0589***	Information & communication	0.1116***
	(0.0021)		(0.0032)
100-199 employees	0.0547***	Prof'l, scientific, & techn'l act's	0.0438***
	(0.0026)		(0.0028)
200-499 employees	0.0636***	Administrative & support act's	0.0204***
	(0.0033)		(0.0028)
≥500 employees	0.0411***	Constant	0.3621***
	(0.0041)		(0.0059)

Notes: 35,564 observations, each observation is a unique firm. The dependent variable is having an automation spike at any point in the sample. Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01.

A.4 Employment growth for automating and non-automating firms: balanced panel

Figure A.2 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 A.2. 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 and 623 for firms without an automation event.

A.5 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 firms in our main sample where we can construct (time-invariant) importer information. First, Table A.2 shows that firms with (net) automation imports are more likely to have automation events: this further establishes the relationship between these measures. However, while over 30% of firms in this sample have an automation event, less than 10% are importers.

Further, Table A.3 shows that, for both firm-level measures of automation, automating firms have faster employment but not wage growth: however, employment growth differences are larger when comparing importers to non-importers.

A.6 Sample construction

For each calendar year *year* we define a set of potential treatment and control group automation events as follows. Potential treatment events for *year* are defined as a firm having an automation event in *year*. *year* has to lie 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. Automation events are excluded if the firm also has another spike in automation cost shares in $t = [-3, 4]$ (with t

Table A.2. Firm-level correlation between automation events and automation imports

Dependent variable: dummy for firm having an automation event				
	(1)	(2)	(3)	(4)
Importer	0.028**	0.035***		
	(0.009)	(0.010)		
Net importer			0.026**	0.033***
			(0.009)	(0.010)
Controls	No	Yes	No	Yes

Notes: N = 30,280 firm observations, where 31% of firms have automation events, and 9.5% (9.2%) are (net) importers. Controls are log total costs and industry fixed effects. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

Table A.3. Firm-level outcomes for firms with automation events and automation imports

	Δ log employment for firms with automation:			Δ log mean daily wage for firms with automation:		
	Events	*Imports*	*Net imports*	*Events*	*Imports*	*Net imports*
	(1)	(2)	(3)	(4)	(5)	(6)
Automating	0.0146***	0.0243***	0.0249***	0.000	0.003	0.003
	(0.005)	(0.007)	(0.007)	(0.002)	(0.003)	(0.003)

Notes: N = 152,006 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.

defined as event time, i.e. $t \equiv year - \tau$ for a treatment group firm that has an automation event in calendar year τ). This gives us 2,436 potential treatment group events. Potential control group events for *year* are defined as firms that have an automation event in year $year + 5$ or later. Hence, these events have to occur between 2008 and 2016. Furthermore, the event is excluded if the control group firm has another spike in automation cost shares in $t = [-3, 4]$ (with t still defined by treatment group firms that have their automation event in $t = 0$). This gives us 21,482 potential control group events.

Columns (1) and (2) in Table A.4 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 an 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 an automation event in 2011 can serve as a control group event for treatment events in 2003, 2004, 2005, or 2006. For our 21,482 potential control events, 20,746 involve a firm that is involved in more than one potential control event, while 736 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 2,436 potential treated events, 1,016 involve a firm that is also involved in at least one potential control event in another year and 1,420 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 $t = -1$. This leaves us with 2,429 potential treatment events merged to 124,146 incumbent workers and 21,308 potential control events merged to 1,156,627 incumbent workers.[25]

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 percentiles of real wage income, as well as separate bins for the 99th and 99.9th percentiles, in each of the three pre-treatment years $t = -3, -2, -1$. 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 30,247 strata for incumbent workers, and in doing so can match 98% of treated incumbents (using 93% of control group incumbents).

After matching, our sample contains $1,017,821$ unique incumbent workers in treatment and control groups. Of those incumbent workers, $100,318$ are treated and $917,503$. Our estimation sample of firms for identifying these treated and control group workers contains $6,940$ unique firms, all of which experience an automation event at some point over the period. As indicated in columns (3) and (4) of Table A.4, workers employed at $2,417$ of those firms are treated, and workers employed at $4,523$ firms serve as controls at least once.

[25] Appendix A.7 below provides further summary statistics of our worker data.

Table A.4. Number of treatment and control events at the firm level by calendar year

	Potential events		Events after matching	
Calendar year	Control	Treatment	Control	Treatment
2003	3,477	199	3,392	199
2004	3,257	197	3,190	195
2005	2,955	202	2,907	200
2006	2,716	222	2,671	222
2007	2,442	316	2,379	313
2008	2,195	310	2,146	307
2009	1,916	360	1,881	358
2010	1,536	315	1,509	312
2011	988	315	974	311
Total	**21,482**	**2,436**	**21,049**	**2,417**
Unique firms involved	4,573	2,436	4,523	2,417
Unique firms only used once	736	2,436	751	2,417

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

A.7 Summary statistics for workers

Table A.5 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 $100,318 + 917,503 = 1,017,821$ observations for incumbents: given our observation window of 8 years ($t = -3$ through $t = 4$) this adds up to the $1,017,821 \times 8 = 8,142,568$ incumbent worker observations used in our regressions.

A.8 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 A.6 are constructed as follows. We draw a 10 percent random sample without replacement from the sample of 10,422 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

Table A.5. Descriptives for incumbent workers

	(1) *Full sample*	(2) *Treated*	(3) *Control*
Annual wage income	40893.37	38955.52	39005.63
	(26823.89)	(25491.86)	(25594.48)
Daily wage if employed	162.08	149.24	149.35
	(100.82)	(97.66)	(97.84)
Annual non-employment duration (in days)	10.32	1.44	1.31
	(48.37)	(14.28)	(13.57)
Hazard of leaving the firm	0.04	0.00	0.00
	(0.20)	(0.00)	(0.00)
Total benefits	441.33	0.00	0.00
	(2975.92)	(0.00)	(0.00)
Probability of entering early retirement	0.01	0.00	0.00
	(0.09)	(0.00)	(0.00)
Probability of becoming self-employed	0.03	0.03	0.03
	(0.18)	(0.16)	(0.17)
Share female	0.25	0.33	0.31
	(0.43)	(0.47)	(0.46)
Foreign born or foreign-born parents	0.15	0.17	0.16
	(0.36)	(0.38)	(0.37)
Age	42.55	40.73	40.61
	(10.26)	(10.20)	(10.08)
Calendar year	2006.91	2006.42	2006.42
	(3.37)	(2.39)	(2.39)
Manufacturing	0.37	0.23	0.23
	(0.48)	(0.42)	(0.42)
Construction	0.12	0.08	0.08
	(0.32)	(0.28)	(0.28)
Wholesale and retail trade	0.19	0.25	0.25
	(0.39)	(0.44)	(0.44)
Transportation and storage	0.09	0.09	0.09
	(0.28)	(0.28)	(0.28)
Accommodation and food serving	0.02	0.02	0.02
	(0.13)	(0.15)	(0.15)
Information and communication	0.06	0.05	0.05
	(0.23)	(0.21)	(0.21)
Professional, scientific, and technical activities	0.08	0.10	0.10
	(0.28)	(0.30)	(0.30)
Administrative and support activities	0.09	0.18	0.18
	(0.29)	(0.38)	(0.38)
0-19 employees	0.06	0.07	0.06
	(0.23)	(0.25)	(0.24)
20-49 employees	0.14	0.16	0.16
	(0.35)	(0.37)	(0.36)
50-99 employees	0.12	0.12	0.13
	(0.32)	(0.33)	(0.34)
100-199 employees	0.12	0.12	0.13
	(0.33)	(0.32)	(0.34)
200-499 employees	0.15	0.14	0.15
	(0.36)	(0.35)	(0.36)
$\geq$500 employees	0.42	0.39	0.37
	(0.49)	(0.49)	(0.48)
Observations	875,1328	100,318	917,503

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 $t = -1$, where weights are obtained from coarsened exact matching as described in Appendix A.6. Standard deviations in parentheses.

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

Table A.6. Brier skill scores for predicting automation event timing

Sample	N	Brier skill score
1	140,524	0.042
2	140,165	0.040
3	139,544	0.041
4	140,484	0.038
5	139,424	0.052
6	138,904	0.039
7	140,642	0.041
8	139,838	0.040
9	139,409	0.044
10	140,314	0.041

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 A.6 shows that these improvements are low, ranging between 3.8 and 5.2%, suggesting that the timing of automation events is essentially random with respect to firms' observed characteristics.

A.9 Additional results

Here, we consider effect heterogeneity by incumbent characteristics, and we compare our results for incumbent workers to estimates for recent hires. 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.

A.9.1 Effect heterogeneity

This section considers 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:

$$y_{ijt} = \alpha + \beta D_i + \gamma post_t + \delta_0 \times treat_i \times post_{it} + \sum_k \left[\delta_k \times treat_i \times post_t \times z_{ki}\right] + \lambda X_{ijt} + \varepsilon_{ijt}, \quad (5)$$

where, as before, i indexes workers, j firms, and t 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, $post_t$ is a dummy variable indicating the post-treatment period (i.e. $t \geq 0$). Further, z_{ki} is a dimension of worker heterogeneity, such as gender, age in the year before automation, or age-specific wage rank, containing $k+1$ categories. 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 (5), 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 A.7 summarizes how average post-treatment effects for annual wage income differ across workers of different ages, gender, and their (initial) firms' sector and employment size. First, we find that workers over the age of 50 are most negatively affected by automation events: while differences with younger age groups are not always statistically significant, the point estimates suggest all other groups experience somewhat smaller income losses. This is not because older workers leave the automating firm at higher rates, but rather, because they experience larger increases in non-employment duration. Unsurprisingly (and not reported here), the early retirement effects we found are entirely driven by the oldest workers. Taken together, older workers appear to face higher adjustment costs from automation than do younger ones.

Column 3 shows effect heterogeneity by firm size: this is important in our setting because the automation cost survey over-represents large firms – while these of course employ the majority of workers, it could still bias the found worker-level effect of automation events by including too low a number of workers experiencing such events in small firms. For this reason, it is reassuring that displacement effects are found across the firm size distribution. However, losses are higher for workers employed at smaller firms, which implies we would probably find somewhat higher average wage losses from automation if our data were more representative in terms of firm size. Lastly, although not reported here, we find that firm separation increases as a result of automation across all firm sizes, but most strongly so for the largest firms – the fact that this does not translate to larger wage losses for these workers suggests they have better outside options. We do not find any statistically significant differences in impact by gender.

In column 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 $t = -1$. 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 quite 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 almost half of all incumbents in our sample (26% and 23%, 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.

Unfortunately, our data do not contain any occupation information, and only contain education information for a small and selected subsample of workers. Instead, we obtain a measure of workers' skill level by calculating each worker's wage rank by age in $t = -1$. 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.[26]

Results are reported in the first column of Table A.8: workers in the highest 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.

[26] As an alternative skill measure, we calculate residual wage quartiles (by first regressing worker wages in $t = -1$ onto a set of observables and their interactions): results (not reported here) are similar.

Table A.7. Relative wage earnings effects by incumbents' characteristics

(1) Age		(3) Gender	
Age 50+ (ref)	-3.06***	Male (ref)	-1.54***
	(1.15)		(0.57)
Deviations from reference group for:		*Deviations from reference group for:*	
Age <30	0.99	Female	-1.17
	(4.48)		(0.98)
Age 30-39	1.16	(4) Sector	
	(0.92)	Manufacturing (ref)	-1.82*
Age 40-49	1.72*		(1.00)
	(0.92)	*Deviations from reference group for:*	
(2) Firm size		Construction	1.01
500+ employees (ref)	-1.15		(1.75)
	(1.22)	Wholesale & retail trade	-2.20
Deviations from reference group for:			(1.46)
200-499 employees	0.35	Transportation & storage	0.49
	(1.66)		(1.80)
100-199 employees	-2.80*	Accommodation & food serving	4.53*
	(1.64)		(2.54)
50-99 employees	-0.24	Information & communication	0.29
	(1.47)		(1.62)
20-49 employees	-2.39*	Prof'l, scientific, & techn'l act's	-0.33
	(1.35)		(1.88)
1-19 employees	-2.68*	Administrative & support act's	0.88
	(1.42)		(1.95)

Notes: Estimates from four separate models, N=8,142,568 for each model. All coefficients are average annual effects over the post-treatment period ($t = 0$ through $t = 4$); coefficients have been multiplied by 100. *p<0.10, **p<0.05, ***p<0.01.

However, the lowest-paid workers (i.e. those in the bottom quartile) do experience the largest wage earnings losses, followed by those in the in second lowest earnings quartile.

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 1 matter for the average worker's exposure to displacement from automation, we are also interested in which workers are displaced within firms. Therefore, the second column in Table A.8 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.[27] If anything, this reveals that the highest-paid workers by age *within* firms appear to lose more wage income than do lower quartiles, although these differences are not statistically significant. However, we should be careful about drawing strong conclusions from these results since they

[27] 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 A.8. Relative wage earnings effects by incumbents' wage quartile

(1) Overall age-specific wage quartile		(2) Within-firm age-specific wage quartile	
Top quartile (ref)	-1.33	Top quartile (ref)	-1.51
	(0.84)		(1.11)
Deviations from reference group for:		*Deviations from reference group for:*	
Second quartile	0.05	Second quartile	-0.74
	(0.73)		(0.65)
Third quartile	-0.96	Third quartile	0.12
	(0.85)		(0.81)
Bottom quartile	-1.46	Bottom quartile	0.79
	(1.67)		(3.14)

Notes: The two models are estimated separately. 8,142,568 observations for column (1); 5,894,240 observations for column (2). All coefficients are average annual effects over the post-treatment period ($t = 0$ through $t = 4$); coefficients have been multiplied by 100. *p<0.10, **p<0.05, ***p<0.01.

may be capturing other factors than pure worker skill, such as the quality of the worker-firm match.

To conclude, although we detect some effect heterogeneity, our findings for incumbents are not driven by workers in a small subset of sectors, firm sizes, or age groups. Further, the income losses found for incumbent workers are not seen among recent hires, and automation affects incumbent workers from all ranks of the "skill" distribution.

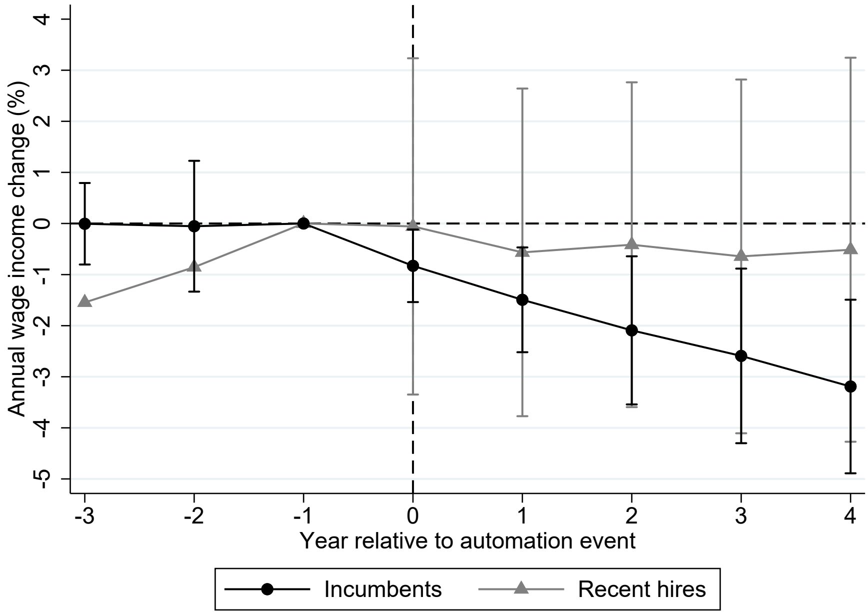
A.9.2 Incumbent workers versus recent hires

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 $t = -1$ but not in $t = -3$ – 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, and matching individual workers on pre-event trends in non-employment duration.[28] After matching, our

[28] In particular, we estimate a linear trend in non-employment duration for individual recent hires before treatment, and match treated and control group recent hires using four bins of this trend: up to the 10th percentile,

Figure A.3. Relative annual wage income effects for incumbents versus recent hires



Notes: N=3,161,056 for recent hires and N=8,142,568 for incumbents. Whiskers represent 95 percent confidence intervals.

sample contains 404,796 unique recent hires (78,282 of whom are treated): given our observation window of 8 years ($t = -3$ through $t = 4$) this results in 3,161,056 observations.

Unlike for incumbent workers, we find no income losses from automation for recent hires, as shown in Figure A.3. 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 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 do not find any statistically significant increase in firm separation for these workers, and differences in non-employment duration with the control group are very close to zero over the entire pre- and post-treatment period. Lastly, we find small positive (albeit statistically insignificant) wage effects for this group.

the 10th percentile to the median, the median to the 90th percentile and higher than the 90th percentile. Together with the other matching variables, we obtain 82,942 strata for recent hires, and can match 95% of treated recent hires (using 65% of control group recent hires).

A.10 Further robustness tests

For succinctness, we only show robustness tests for relative annual wage earnings.

A.10.1 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 A.4.

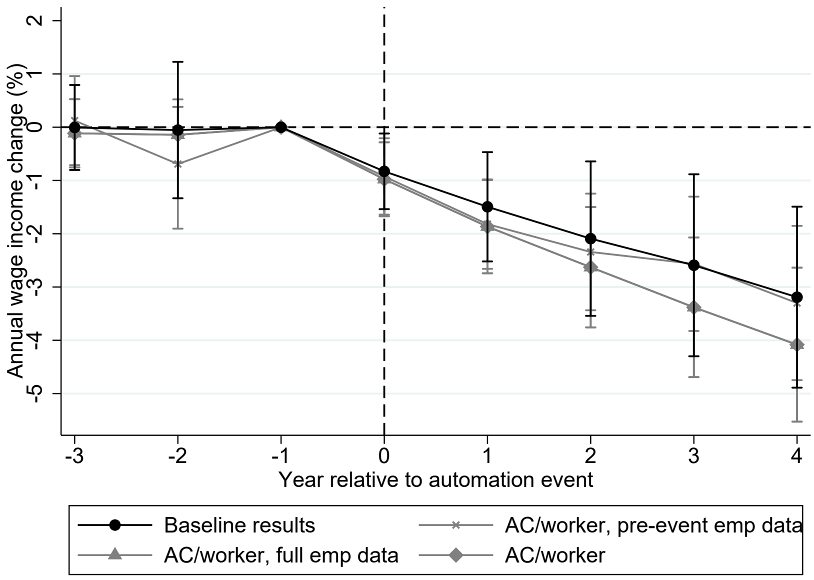
Further, we show that results are robust to varying the spike threshold from two to four times the average automation costs (our baseline is thrice the average automation costs). Panel (b) in Figure A.4 reveals that estimated effect sizes are somewhat larger the higher the threshold, as expected, but 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.

A.10.2 Changes in model specification

Here, we change our model specification in a number of ways. In particular, compared to our baseline estimates, Figure A.5 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 A.4. Robustness to different definitions of automation events

(a) Automation costs per worker



(b) Changes in spike threshold for automation cost shares

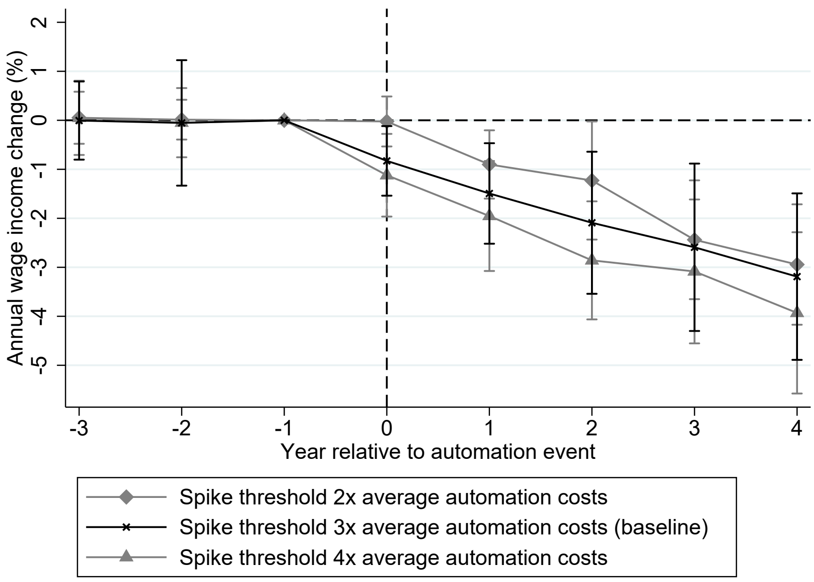
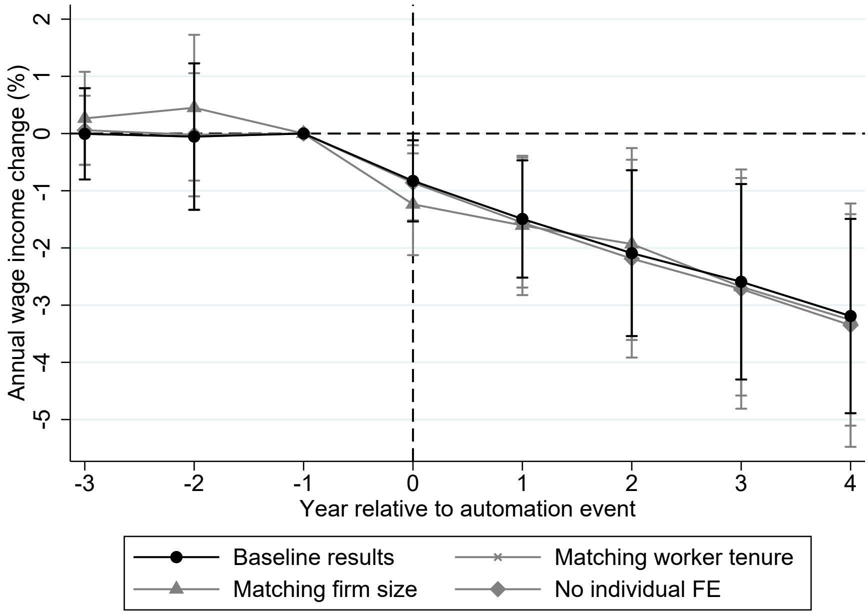


Figure A.5. Robustness to changes in model specification



A.10.3 Randomization test

We subject our results to a randomization test as first introduced by Fisher (1935).[29] To do this, we take our sample of 35,567 firms, randomly draw firms with replacement, and then for each of these firms randomly assign a year to have a placebo automation event.[30] 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.

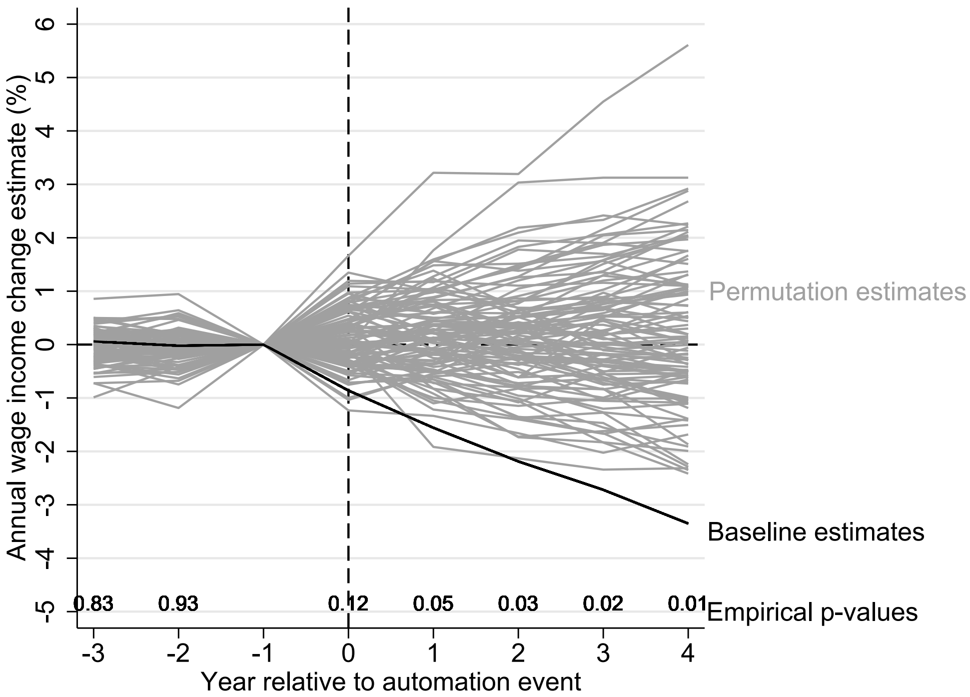
Results are shown in Figure A.6: each gray line presents a set of placebo (dynamic) treatment estimates, whereas the black line presents our actual treatment estimates. The graph also shows probability values calculated using the rank of the absolute value of our estimated coefficient among the 100 permutated estimates.[31] Something at least as extreme as our treatment estimate is unlikely to occur by chance: from the first post-event year onwards, the probability is very close to zero. All in all, this increases confidence that our estimates are not a statistical false positive.

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

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

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

Figure A.6. Robustness to different definitions of automation events



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

A.11 Computer investments by sector and firm size

Tables A.9 and A.10 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 A.9. 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	1,086	402	2.7	5,185	40,841
Construction	535	233	2.3	2,816	18,271
Wholesale & retail trade	1,261	594	2.1	7,228	50,395
Transportation & storage	997	497	2.0	2,282	15,850
Accommodation & food serving	280	165	1.7	742	4,460
Information & communication	2,212	2,547	0.9	1,573	9,851
Prof'l, scientific, & technical activities	1,387	963	1.4	2,360	14,738
Administrative & support activities	936	407	2.3	2,921	17,391

Notes: Overlapping sample: N=25,107.

Table A.10. 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	2,435	1,194	2.0	2,248	11,340
20-49 employees	928	593	1.6	10,493	66,433
50-99 employees	912	498	1.8	5,852	41,532
100-199 employees	1,028	572	1.8	3,413	26,519
200-499 employees	1,314	621	2.1	1,936	16,202
≥500 employees	1,794	695	2.6	1,165	9,771

Notes: Overlapping sample: N=25,107.