

What Happens to Workers at Firms that Automate?*

James Bessen
Boston University

Maarten Goos
Utrecht University

Anna Salomons
Utrecht University

Wiljan van den Berge
Utrecht University

November 2022

Abstract

We estimate the impact of firm-level automation on individual worker outcomes by combining Dutch micro-data with a direct measure of automation expenditures covering all private non-financial sector firms. Using a novel difference-in-differences event-study design leveraging lumpy investment, we find that automation increases the probability of incumbent workers separating from their employers. Workers experience a 5-year cumulative wage income loss of 9 percent of one year's earnings, driven by decreases in days worked. These adverse impacts of automation are larger in smaller firms, and for older and middle-educated workers. By contrast, no such losses are found for firms' investments in computers.

Keywords: Firm-level automation, Worker displacement

JEL: D24, J23, J62, O33

*Helpful comments by Daron Acemoglu, David Autor, Peter Blair, Brian Jacob, Egbert Jongen, Guy Michaels, Pascual Restrepo, Robert Seamans, Bas van der Klaauw, three anonymous reviewers, 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 Productivity Lunch, NBER Summer Institute, TASKS V Bonn, Tinbergen Institute, 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. Goos, Salomons, and Van den Berge thank Instituut Gak 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 concerns 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 or other manufacturing technology, and measures adoption through imports, surveys, or electricity usage (Doms et al., 1997; Dunne and Troske, 2005; Dinlersoz and Wolf, 2018; Cheng et al., 2019; Dixon et al., 2019; Koch et al., 2021; Acemoglu et al., 2020; Aghion et al., 2020; Humlum, 2021; Bonfiglioli et al., 2021; Hirvonen et al., 2021).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. (2021) find that, when accounting for firm-level product demand shocks that generate a positive correlation between robot imports and employment, exogenous exposure to automation imports leads to a decline in firm-level employment.

There are as yet no empirical studies that directly examine what happens to individual workers when their firm decides to automate.² However, studying this adjustment process

¹A related literature studies aggregate adoption of robotics, using industry and/or regional variation (Graetz and Michaels, 2018; Aghion et al., 2019; Dauth et al., 2021; Acemoglu and Restrepo, 2020).

²An exception is Humlum (2021) who finds that a firm's wage bill declines for production workers and increases for tech workers after automation. Other related studies on worker adjustments have used more aggregate sources of variation and do not always focus on causal effects. In particular, Cortes (2016) finds that workers switching out of routine-intensive occupations experience faster wage growth relative to those who stay; Dauth

is critical to our understanding of how labor markets are impacted by the automation of work. While the theoretical literature models automation as displacing humans from certain tasks (see Acemoglu and Restrepo 2018), this need not translate to worker-level displacement if productivity increases are sufficiently strong. Further, firm- and sector-level impacts do not inform about the adjustment process that workers go through: even positive aggregate changes may be accompanied by worker churn (including across firms within sectors) and economic consequences for individual workers directly affected by automation. 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 account, 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.

et al. (2021) 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.

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 35,580 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 9% of an annual wage. These losses are driven by increases in non-employment: we do not find evidence of wage scarring on average 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. Turning to effect heterogeneity, we find that these adverse impacts from automation are larger in smaller firms and for older and middle-educated workers. 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 other measures of firm-level technology use. 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 the relationship between automation and firm-level employment and wage outcomes. 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

2.1 Administrative data on workers and firms

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 35,580 unique firms with at least 3 years of automation cost data – together, these firms employ around 5 million unique workers annually on average. 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 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 when matching workers to firms.

We further observe workers’ gender, age, and nationality. A downside to these data is that we do not observe workers’ occupations, and only have information on the level of education 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.

³Firms 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.

We can track workers across firms on a daily basis, allowing us to construct indicators for firm separation and days spent in non-employment.

2.2 Automation costs

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. Further, we measure it at the level of the firm rather than the industry, and across all private non-financial sectors.

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 211,000 euros annually, or 1,045 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.

The top panel of 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 and within each of these sectors. Average expenditures at the sectoral level range from 244 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 2 also 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 around 0.41%, whereas firms with between 100 to 200 employees have a cost share of 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. However, there is substantial variation within size classes, also.⁴

2.3 What do automation costs capture?

Since our automation cost measure does not specify the automation technology being implemented, we compare firm-level automation costs with other, specific, technologies obtained from a firm-level survey provided by Statistics Netherlands. These measures are self-reported indicators (unlike automation costs, which are part of firms’ official bookkeeping), and the questions differ by survey year.⁵

We obtain correlations between firms’ (standardized) automation cost shares and their self-reported implementation with self- or externally developed process, product, and

⁴Appendix A.1 further illustrates how the distribution of automation cost expenditures per worker and automation cost shares change over time. It shows that both the means and variances of automation costs per worker and automation cost shares have increased over time.

⁵We take firms’ most recent reported technology use.

organizational innovations (all three measured as dummies), controlling for sector fixed effects and firm size.⁶ Automation cost shares are positively correlated with all three self-reported innovation types, indicating that automating firms are more innovative across the board. However, the coefficient for process innovation is more than twice as large as those for product and organizational innovations, highlighting that this innovation margin is differentially more important for firms with high automation cost shares. Firms with a one standard deviation higher automation cost share are 20.5 percentage points more likely to report implementing process innovations. This suggests that automation cost shares differentially measure process automation technology rather than innovations more broadly.

We also estimate the same models for a host of self-reported uses of specific technologies, all of which are given as dummy variables in the data.⁷ First of all, this shows that automation cost shares are positively associated with the general indicator of using electronic data suited to automated processing. Automation cost shares are also positively correlated with several well-known technologies, such as Customer Relationship Management (CRM) software used for inventory and distribution and Enterprise Resource Planning software. Further, value chain integration itself is not predictive of automation cost shares, but the use of automated records in such integration is.

Further, automation cost shares are correlated with the use of big data analysis, and with the use of cloud services for both customer management and accounting and financial management. Firms with high automation cost shares are more likely to receive their own orders through Electronic Data Interchange (EDI) but less likely to order through EDI at other firms— the latter would of course require these trading partners to have EDI technology. Finally, automation cost shares are higher for firms using sales software.

Importantly, not all technologies are associated with firms' automation cost shares: the use of radio-frequency identification, Local Area Networks, or using internet for financial transactions or worker training are not predictive of a firm's automation cost share.

⁶Results are reported in Appendix Table A.1.

⁷Results are reported in Appendix Table A.2.

These results can be seen as a type of placebo result, since we would not expect these specific technologies to be associated with substantial automation at the firm. All in all, these descriptives show that firms' automation costs are most strongly related to process innovations, and meaningfully associated with the usage of various specific technologies, particularly those which make use of electronic data suited to automated processing.⁸

Another measure of firm-level technology adoption used in the literature are firm imports of robots and other automating technologies. We compare this to our automation cost measure by leveraging data on automation imports, which are available from 2010 onward.⁹ Automation imports are defined as imports of intermediates classified by Acemoglu and Restrepo (2021) as automatically controlled machines, automatic transfer machines, automatic welding machines, numerically controlled machines, and (other) industrial robots. We find that average automation expenditures are substantially higher than average automation imports since few firms are importers, and that automation costs are observed across a wide range of sectors whereas automation imports are largely limited to firms in Manufacturing, Wholesale & retail, and Transportation & storage.

Automation imports and automation expenditures are correlated between, but not within firms: that is, firms are not more likely to have higher automation costs when they (net) import more automation technology. This comparison shows that the firm-level import measures used in the literature likely capture a different dimension of technology adoption. One reason could be that our measure of automation costs captures that the firm relies on an integrator to install the machinery whereas in the case of automation imports the firm is purchasing the automation machinery directly. However, a limitation

⁸Appendix A.3 further reports the sectoral prevalence of the technologies which are significantly positively correlated with automation cost shares: this gives insight into which specific automation technologies are being used in these industries. The main finding is that different sectors do implement these technologies at different rates, but the overall use of data for automated processing is relatively common in all sectors, reflecting a general characteristic of automation.

⁹Results are reported in Appendix A.4.

to our comparison is that Dutch firms also export automation technology at high rates—in fact, the Netherlands is a net exporter of automation technology. This could indicate that, in the Netherlands, firm-level automation imports need not correspond to actual deployment of the imported technology within the firm.

3 Automation events

3.1 Defining automation events

We identify automation events using what we call automation spikes. We define an automation spike 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 :

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

where $\mathbb{1}\{\dots\}$ 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 as the year when the firm has its first increase in automation cost share that qualifies as a spike.^{10,11}

¹⁰Firms without an automation event do not necessarily have zero automation costs, but their automation expenditures do not vary much relative to total costs, implying they do not undergo automation events as we define them. Firms without such events do have lower automation costs also on average, however.

¹¹Note that automation costs capture the use of third-party automation services: this raises the question whether these events capture outsourcing at the firm-level. However, for outsourcing, we would expect to see a sustained increase of costs for the firm (reflecting a flow of wages to these outsourced workers) rather than a temporary spike.

We develop a theoretical model to show that the existence of automation events is consistent with monopolistic competition between profit-maximizing firms that choose their optimal timing to incur a firm-specific fixed cost of automation.¹² Assuming technological progress such that a firm’s technology ages and its relative output price increases if it does not automate, the model predicts spikes in automation cost shares, i.e. a one-period increase in automation expenditures relative to other costs preceded and followed by periods in which the firm does not automate. Moreover, firms will automate at different points in time if they have different fixed costs of automation.

3.2 Summary statistics of automation events

Out of the total number of 35,580 firms with at least 3 years of automation cost data, there are 10,425 firms that have at least one spike in their automation cost share between 2000 and 2016.¹³ That is, 29% of the firms in our sample spike at least once over the 17 years of observation.¹⁴ 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 τ as event time: the difference between the actual calendar year and the calendar year of the spike, such that all automation events line up in $\tau = 0$. Figure 1 uses the sample of 10,425 firms, showing a clear one-period increase

¹²This model is related to Bonfiglioli et al. (2021) and Humlum (2021) and extends task-based frameworks of automation in Acemoglu and Restrepo (2018, 2019) by endogenizing firm-level automation as a lumpy investment following Haltiwanger et al. (1999). It is provided in Appendix B.

¹³Shown in Appendix Table C.1.

¹⁴We also find that this percentage varies relatively little across sectors (ranging from 25% in Construction to 39% in Information & communication) and firm sizes, although the smallest firms are least likely to have automation events. See Appendix C.2 for details.

in the automation cost share when a firm has its automation event.¹⁵

We also assess the ‘lumpiness’ of automation costs. We find that while only 5% of firm-year observations are automation spikes, these make up 11% of total investment in automation across all firms that we observe at least three times in the automation cost data. Further, among firms that have an automation cost spike, on average 54% of all observed automation costs occurs during such spikes (with a median of 52%): if automation costs were evenly distributed within firms over time, any one year would only correspond to 13% of the total outlay. This suggests that while firms also have outlays on automation that do not fall within our strict automation event definition, automation costs are lumpy.

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, showing that these firms are on substantially different labor demand trajectories. Section 4.2 therefore presents an analysis of the impacts of automation 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:

$$\Delta \ln Y_{jt} = \beta \times A_j + D_t + \gamma \times X_j + \varepsilon_{jt}, \quad (2)$$

¹⁵Appendix C.3 shows this increase is driven by automation costs rising, not total costs decreasing.

where the dependent variables are annual log changes in employment, in the average daily wage, and in the wage bill for firm j in calendar year t . A_j is a dummy for the firm having an automation event over the 2000–2016 period, D_t are calendar year fixed effects, and X_j additional controls consisting of two-digit sector dummies and baseline firm-level characteristics.¹⁶ The term ε_{jt} 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, mean daily wage, and wage bill trajectories.

Table 3 shows that automating firms have 1.5 to 1.9% higher employment growth but not higher daily wage growth compared to non-automating firms. Automating firms’ wage bills therefore grow faster compared to non-automating firms—Dunne and Troske (2005) find similar patterns for U.S. manufacturing firms that are adopting information technology. In Bessen et al. (2020), we show that these associations are not significantly different for manufacturing firms compared to non-manufacturing ones. In sum, these results show that, on average, automating firms have faster employment growth¹⁷ but not faster wage growth.¹⁸

4.2 Difference-in-differences 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

¹⁶Appendix D.1 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.

¹⁷Appendix D.2 shows similar results when using a balanced panel of firms.

¹⁸Appendix D.4 shows that similar effects are found for importers of automation technology compared to non-importers, although the employment growth and wage bill differences are more than twice as large, reflecting that automation importers are much more positively selected.

firms. We therefore also consider the evolution of firm outcomes around an automation event, looking at the sub-sample of automating firms only and exploiting the timing of automation events for identification.

We set up our data in a stacked difference-in-differences design as in e.g. Cengiz et al. (2019).¹⁹ More specifically, we create separate datasets for each cohort of firms that have their first automation event in year c , with $c \in \{2003, \dots, 2011\}$. As before, τ is event time, i.e. calendar year t minus the calendar year c in which the firm has an automation event ($\tau \equiv t - c$). In each dataset we keep $\tau \in \{-3, \dots, 4\}$ as our event window. Then for each dataset, we add observations for the same calendar years on all firms that have an automation event in year $c + 5$ or later as controls. For example, our first cohort of treated firms are firms that have their first automation event in 2003. The event window surrounding this event contains the calendar years from 2000 to 2007. All control firms for the 2003-cohort are those firms that we observe over the same event window (2000 to 2007), and that have their first automation event later than 2007. These firms are “clean controls” in the sense that they do not have an automation event during the event window. We repeat this procedure for each cohort of firms. Finally, we stack the cohort-specific datasets so that they line up in terms of event time $\tau \in \{-3, \dots, 4\}$.²⁰

The by now well-documented problems related to two-way fixed effects in an event-study design arise from staggered treatment timing (de Chaisemartin and D’Haultfoeuille, 2020; Sun and Abraham, 2021; Goodman-Bacon, 2021; Callaway and Sant’Anna, 2020). By creating a balanced panel in event time, we have effectively eliminated the staggered timing in the data, and hence do not suffer from the issues that staggered timing may create. Most notably, we do not use “already-treated” units as control units. Baker et al. (2021) show that a stacked difference-in-differences setup recovers the true treatment

¹⁹For other recent papers using this setup, see for example Goldschmidt and Schmieder (2017); Deshpande and Li (2019); Clemens and Strain (2021); Baker et al. (2021).

²⁰To ensure that our results are not driven by selective survival of firms, we also ensure that our treatment firms survive until at least 5 years after the automation event. This does not affect our results.

effects in the case of staggered timing, just as the Callaway and Sant’Anna (2020) and Sun and Abraham (2021) approaches do.

Using this stacked dataset, we estimate the following difference-in-differences specification:

$$\ln Y_{jt} = \alpha + \sum_{\tau \neq -1; \tau = -3}^4 \beta_{\tau} \times I_{\tau} + \sum_{\tau \neq -1; \tau = -3}^4 \delta_{\tau} \times I_{\tau} \times \text{treat}_j + \eta_j + \theta_t + \varepsilon_{jt}, \quad (3)$$

where j indexes firms by cohort.²¹ I_t are leads and lags in event time, with $\tau = -1$ as the reference category. Treatment is indicated by the dummy variable treat_j which turns on for firms that have an automation event in the window. Y_{jt} is j ’s employment or average daily wage and η_j are a set of cohort by firm fixed effects²² and θ_t a set of calendar year fixed effects.²³

We require two assumptions for a causal interpretation of δ_{τ} (e.g. Callaway and Sant’Anna, 2020; Borusyak et al., 2021).²⁴ First, we need that treated and control firms follow parallel trends in absence of treatment. We provide evidence for this assumption by showing that pre-event trends are mostly similar for firms that have an automation event now compared to those that have an automation event later. In addition, as we saw above, firms that do not have an automation event are on different trends than firms that do have an automation event. By relying only on differences in event timing, rather than also on event incidence, we can be more confident that we are comparing firms on similar trends.²⁵ Second, we need that firms do not anticipate an automation event, i.e. there should not be an effect of treatment in the future on current outcomes (e.g. Ab-

²¹This is because individual firms can appear multiple times in the stacked data.

²²We account for the fact that control firms can appear multiple times in the stacked dataset by including cohort by firm fixed effects (Baker et al., 2021).

²³Adding a full set of cohort by year fixed effects does not affect our findings.

²⁴Also see Appendix B for a more formal discussion in the context of our theoretical model.

²⁵Appendix D.3 shows results using non-automating firms as control group: we find diverging pre-trends in employment in this comparison. This is consistent with the

bring and Van Den Berg, 2003). This “no-anticipation” assumption is more difficult to maintain at the firm than worker level, because firms that decide to automate are more able to anticipate their own decision and this might affect other decisions they make in anticipation of the automation event. For this reason we will interpret these firm-level results as descriptive.

Results are shown in Figure 2 for two groups of firms: those with fewer than 500 workers, and firms with 500 workers or more (where firm size is determined at $\tau = -1$). First, we find that in firms with fewer than 500 workers, employment contracts following an automation event by about 20% compared to firms that have an automation event later. In contrast to employment, daily wages remain stable around these events. In the largest firms, wages actually grow by about 5% relative to the control group, but there is no significant change in employment. All in all, wage bills fall after the automation event in the smaller but not the larger firms. Bessen et al. (2020) show that while these effects are quantitatively larger (albeit more imprecisely estimated) in manufacturing firms, they are observed in non-manufacturing firms as well.

We also find that employment, wage, and wage bill trajectories for the subsample of automation importers are similar to those of the largest firms in Figure 2, with even more sizable positive impacts.²⁶ This is because there is a strong positive correlation between being an importer of automation technology and firm size²⁷, suggesting that using imports to measure automation involves important selection effects, as also argued by Bonfiglioli et al. (2021). Our main analyses, by contrast, rely on our measures of automation costs that are observed across different firm sizes.

evidence presented in section 4.1 and suggests that non-automating firms are not a good control group for automating firms in our case.

²⁶See Appendix D.4 for details.

²⁷Firms with automation cost spikes employ around 8.5% more workers than do firms without such spikes; by contrast, firms with automation imports are 131% larger than non-importers.

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 main results. Section 5.3 discusses effect heterogeneity, and in section 5.4 we present robustness tests.

5.1 Difference-in-differences event-study design

For the worker-level analysis we exploit our linked employer-employee data and follow a similar stacked difference-in-differences setup as in our firm-level analysis. Again, we create separate datasets for each cohort of firms that have an automation event in year c . We subsequently merge all workers who are employed at this firm in $c - 1$. These are our sets of treated workers. Similar to the firm-level analysis, we keep observations for these workers over the event window $\tau \in \{-3, \dots, 4\}$, with τ defined as event time ($\tau \equiv t - c$). Then, for each cohort c and the same set of calendar years, we add all workers employed in $c - 1$ at firms that have their first automation event in year $c + 5$ or later as control workers. For example, we observe the cohort of workers treated in 2005 over the event window 2002 to 2009. We pick our control workers such that we observe them over the same window from 2002 to 2009, and that in 2004 they are employed at a firm that has an automation event later than 2009.²⁸ We repeat this procedure for each cohort $c \in \{2003, \dots, 2011\}$ and stack the resulting datasets so that they line up in event time. We restrict our baseline analysis to incumbent workers: those workers with at least 3 years of firm tenure in $c - 1$.²⁹

²⁸We only require control group workers to be at a firm j in year $c - 1$. Hence, control group workers do not have to be employed at firm j when firm j actually automates in year $c + 5$ or later.

²⁹Appendices E.1 and E.2 provide further details on sample construction and summary statistics for workers.

We use the following difference-in-differences specification:

$$Y_{ijt} = \alpha + \beta \times treat_i + \sum_{\tau \neq -1; \tau = -3}^4 \gamma_\tau \times I_\tau + \sum_{\tau \neq -1; \tau = -3}^4 \delta_\tau \times I_\tau \times treat_i + \lambda X_{ijt} + \varepsilon_{ijt}, \quad (4)$$

where i indexes individuals by cohort, j firms by cohort, and $\tau \in \{-3, \dots, 4\}$ is event time.³⁰ Y_{ijt} is an individual-level outcome for worker i who must be employed at firm j in $\tau \in \{-3, \dots, -1\}$. Firm j can be a treatment group firm if j has an automation event in $\tau = 0$, 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 i if her firm j has an automation event at $\tau = 0$. Further, I_t are event-time indicators, with $\tau = -1$ 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 $\tau = -1$) characteristics as well as year fixed effects. In our baseline specification, we replace $\beta \times treat_i$ with individual by cohort³¹ fixed effects which also absorb the non-time varying controls in X (gender, nationality, firm sector and employment at $\tau = -1$).³² 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.³³

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 $\tau = -1$ (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 $\tau = -1$, we have that $\delta_0 = -0.01$. Similarly, if automation leads to an annual wage income loss in $\tau = 4$ that equals 3% of annual wage income in $\tau = -1$, we have that $\delta_4 = -0.03$. The figures in the next section

³⁰Index i is for an individual worker within a cohort c since individual workers can appear multiple times in the data. The same holds for any firm j .

³¹We account for the fact that individual control workers can appear multiple times in the stacked dataset by using cohort by individual fixed effects (Baker et al., 2021).

³²Except when we estimate the hazard rate of a worker leaving the firm.

³³Adding a full set of cohort by year fixed effects does not affect our findings.

plot estimates of δ_τ for $\tau \in \{-3, \dots, 4\}$ 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_τ can be interpreted as causal effects under the identifying assumptions of (i) parallel trends in the absence of automation events, and (ii) no anticipation of the automation event by workers. Our empirical approach directly supports these assumptions 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, as we have shown in the previous section: 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 2021; Miller 2017; Lafortune et al. 2018).³⁴

Second, our specification only considers incumbent workers, defined as those with at least 3 years of firm tenure in $\tau = -1$. On average across firms in our data, 64% of workers are incumbents (where the median is 68%). This captures workers who have a stable working relation with their firm.³⁵ 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. In this respect our worker-level analysis differs most clearly from the firm-level analysis: the no anticipation assumption is much more likely to hold for incumbent workers who are in a stable working relationships with the firm. This

³⁴In support of our approach, Appendix E.3 shows that the timing of automation events cannot be easily predicted from observed firm-level characteristics in our data.

³⁵Dutch labor law during almost our entire data period ensures temporary contracts are of a maximum duration of 3 years, implying that workers with 3 years of tenure are likely to have open-ended contracts.

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), where such lay-off events are assumed not to be anticipated for the firms' incumbent workers.

Third, we further strengthen the assumption of parallel trends by matching on worker and firm observables to ensure that $\delta_\tau = 0$ for all $\tau < 0$ (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 deciles of real wage income, as well as separate bins for the 99th and 99.9th percentiles, in each of the three pre-treatment years $\tau = -3, -2, -1$. We then match treated workers to control group workers for each of these income bins, while additionally requiring them to be observed in the same calendar year and work in the same sector one year prior to treatment. We include calendar year and sector matching to ensure we are not capturing sector-specific business cycle effects, or other unobserved time-varying shocks affecting workers based on their original sector of employment. As such, each treated worker is matched to a set of controls from the same calendar year and sector and belongs to the same pre-treatment earnings 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 5.4 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.

5.2.1 Annual wage income, 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 $\tau = -1$ 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 $\tau = -1$ 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 $\tau = 0$; 1.6% (of annual wage earnings in $\tau = -1$) in $\tau = 1$; 1.9% in $\tau = 2$; 2.4% in $\tau = 3$; and 2.4% in $\tau = 4$. Overall, automation decreases annual wage earnings for an incumbent worker by 9.3% ($=1+1.6+1.9+2.4+2.4$) cumulatively of her annual wage earnings in $\tau = -1$ 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 9% annual earnings loss corresponds to a cumulative real earnings loss of around 3,700 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 study 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 1.7 percentage point higher hazard rate of firm exit. This is a non-negligible increase, given that the corresponding hazard among control group workers is 8.4%.

It is noteworthy that worker displacement does not occur instantly: rather, displace-

ment 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 and time-consuming to fire them. Further, these gradual changes could in part also result from a time delay in the effective implementation of automation technologies and work process changes 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 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 (see Davis and Von Wachter (2011) for an overview).

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 ($\tau = 0$), 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 5.3 days annually after 5 years, with a total cumulative increase in non-employment of around 18 days compared to the control group. By comparison, in the event year, matched control group incumbents spend around 11 days in non-employment on average, suggesting automation produces an average increase of 9% in non-employment days in the automation year itself. The cumulative five-year impact corresponds to a 12% average increase relative to the five-year cumulative non-employment duration of 145 days experienced by control group incumbents.

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.³⁶ 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.³⁷

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 344 euros on average. Given that the average annual wage income loss cumulates to about 3,700 euros after 5 years (see above), this implies that less than 10% 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.³⁸ Consistent with high

³⁶As shown in Appendix E.4, we do not find statistically significant impacts on log hourly wages for the subperiod where these are observed, suggesting our finding is not driven by mismeasurement.

³⁷See Deelen et al. (2018); Mooi-Reci and Ganzeboom (2015) who find evidence of substantial wage scarring after mass lay-offs in the Dutch context, using the same administrative data as we do here.

³⁸From 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.

unemployment benefit eligibility, we do not see any increase in welfare payments. Lastly, the uptake of disability benefits is actually slightly decreasing for treated versus control group workers over time.

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 0.7 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 $\tau = 4$ is around 1.73%. As such, the treatment effect represents a 40% increase in the incidence of early retirement. Besides early retirement, the figure also examines the possibility that displaced workers enter self-employment. We do not find any effects there: estimates are very close to zero and statistically insignificant, highlighting that self-employment is not a compensating income source.

5.3 Effect heterogeneity

First, we consider effect heterogeneity by incumbent worker characteristics: firm size, worker age, and education level.³⁹ We find substantial heterogeneity in the effects of firm-level automation for incumbent workers employed in firms of different sizes, as measured by their number of workers in $\tau = -1$.

Annual treatment effects (averaged over the post-treatment period $\tau = 0$ through $\tau = 4$) are reported for six firm size classes in panel A of Table 4. Several results emerge. First, while the average incumbent worker experiences losses in earnings across firm size classes, these losses are generally larger for smaller firms: in firms with up to 50 employees, workers lose around 3.2% earnings annually over the four years following the automation event, whereas workers lose only 0.8 to 2.4% annually in larger firms. These differences are only partly driven by differences in firm separation and non-employment: automation leads to declining log daily wages (conditional on employment) for workers in smaller but not larger firms. Overall, this is consistent with the firm-level analyses re-

³⁹See Appendix E.5 for details.

ported in Section 4.2, where we found stronger wage growth and no employment losses for larger firms following automation events.⁴⁰ Since losses are higher for workers at smaller firms, we would probably find somewhat higher average wage losses from automation if our data were more representative in terms of firm size. Theory highlights that effects on incumbent workers are less negative in firms where productivity increases from automation are larger. Viewed through the lens of this model, our results are consistent with larger firms more effectively substituting capital for labor. This could for example arise if larger firms are better able to relocate workers across jobs within the firm.

Further, we find that workers aged 50 and older are most negatively affected by automation events, as shown in panel B of Table 4: differences with younger age groups are statistically significant, implying all other groups experience smaller or even negligible income losses. This is the result of older workers separating from the automating firm at higher rates, and experiencing larger increases in non-employment duration. Unsurprisingly (and not reported here), the early retirement effects we found are entirely driven by the oldest age group. Taken together, older workers appear to face substantially higher adjustment costs from automation than do younger workers.

Wage losses from automation also differ by workers' education level. We only observe education for a subset of workers, reducing statistical power. However, the estimates show losses are highest for middle-educated workers, and lowest for high-educated workers. This is consistent with a long literature which has documented that jobs that can at present be automated are predominantly concentrated in the middle of the wage distribution (see Acemoglu and Autor (2011) for an overview), although effects on aggregate skill demand may largely play out between rather than within firms as automating firms gain employment share (Doms et al., 1997; Dunne and Troske, 2005).⁴¹

⁴⁰Appendix E.5 shows that effects for incumbent workers employed at firms which import automation technology are very similar to the effects for larger firms.

⁴¹Appendix E.5 reports further effect heterogeneity results. We find a similar pattern as for workers' education level when using within-firm age-specific wage quartiles as our skill measure.

We find that effects are quite pervasive across sectors: this suggests that displacement effects are not limited to robotics or process innovations in manufacturing. However, as we do not observe the specific technology being adopted or work process being changed, our estimates—even within sectors—may capture an average across different automation technologies which individually may have more positive or negative effects on workers.

Lastly, 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.⁴² Unlike for incumbent workers, we find no income losses from automation for the average recent hire. This could be 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 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.

5.4 Robustness tests

Our estimates can be interpreted as causal effects only under the identifying assumption of parallel trends in the absence of automation events and no anticipation of the automation event. A simple falsification test for parallel trends is to see whether $\delta_\tau = 0$ for $\tau < 0$, 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. 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

⁴²See Appendix E.6 for details.

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 remove firms where there was a new worker among the firm’s top-decile annual wage income earners⁴³ 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.

For all these robustness checks, estimates are very similar, 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.⁴⁴

As a further robustness check, we construct events for firm-level investments which we expect to be neutral from the perspective of incumbent workers: spikes in “other material fixed assets”. This is a residual category comprising items like furniture, shelving, silos, containers, and pellets. It excludes buildings, land, means of transportation, machines, communication equipment, and computers; as well as immaterial assets such as software and licenses. We construct spikes in other material fixed assets analogously to automation costs— these expenditures are also lumpy, allowing for the construction of spikes.⁴⁵ We estimate worker-level difference-in-differences models for the overlapping sample of firms where we observe both automation events and at least three years of data on other fixed material assets. In contrast to automation events, events based on other fixed material assets do not have any impact on firms’ incumbent workers.⁴⁶ This shows that the worker impacts we find for automation events are not a mechanical outcome of

⁴³Conditional on this worker earning at least 150 euros a day, i.e. 40,000 euros a year.

⁴⁴Appendix Figure E.4a in Appendix E.7.2 summarizes our estimates for relative annual wage income for all four robustness checks pertaining to firm-level pre-trends and events.

⁴⁵See Appendix E.7.1 for details.

⁴⁶Results are shown in Appendix Figure E.4b.

constructing firm-level events the way we do.⁴⁷

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 conducts a separate and partially overlapping firm survey on investments, including computer investments.⁴⁸ This item is called ‘computers’ or ‘computers and other hardware’ 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, 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.

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 spikes in computer investment. Second, if automation cost expenditures and computer investments are correlated at the firm level, we can remove firms which have computer investment

⁴⁷Appendices E.7.3, E.7.4, and E.7.5 report results from three further robustness tests: 1) using alternative definitions of automation events, including different spike thresholds, 2) changing the model specification, and 3) performing a randomization test. Also here, we find that our results are robust.

⁴⁸We only consider computer investments because investments in software and communication equipment are only observed from 2012 onward. In 2012, software investments are of a similar magnitude as computer investments.

spikes within our estimation window to rule out that our automation event is partially capturing investment in computers. Conversely, we also estimate the effects of computer investments in isolation, that is, excluding any events where automation events occur within the estimation window.⁴⁹

Automation costs are higher than computer investments across the distribution, both in total and per worker⁵⁰, though of course 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 worker.⁵¹ We use the same threshold, assigning firms a computer investment spike if their computer investment per worker exceeds three times their usual level.⁵² Compared to automation spikes, computer investment spikes are more frequent. However, similar to automation events, there is 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.

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 and computerization events in isolation, identifying treated and control group workers for one type of event while ignoring the

⁴⁹Throughout our analyses, 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 25,118 instead of 35,580 firms, and is more skewed towards larger firms as these are most likely to be sampled in both surveys.

⁵⁰As shown in Appendix F, along with further details of computer investments by sector and firm size.

⁵¹We 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.

⁵²The resulting distribution of computer investment spikes is reported in Appendix F.

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: firms that have recently had one type of event are more likely to also experience the other sometime soon, 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: 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 5.

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. Our findings are robust to considering automation events without concurrent computerization and vice versa. 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. This could be due to various factors, which we cannot disentangle here. First, computer technology could be less worker-displacing than automation overall because it more strongly complements worker capabilities. Second, firms' investments in computer technology may be less displacing from workers' perspective if these investments replace older vintages of computers (and computerization events reflect replacement costs), as this technology has already reached higher adoption levels. In that case, computerization effects reflect capital deepening, which, in contrast to automation, does not lead to worker displacement (Acemoglu and Restrepo, 2018).

7 Conclusion

We estimate 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 difference-in-differences event-study design exploiting spikes in automation costs, 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 disruptive 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 9 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 substantially 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, workers with medium levels of education, 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 (at present) a more labor-displacing force.

References

- Abbring, J. H. and Van Den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517.
- Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4:1043–1171.
- Acemoglu, D., Lelarge, C., and Restrepo, P. (2020). Competing with Robots: Firm-

- Level Evidence from France. *American Economic Association Papers and Proceedings*, 110:383–388.
- Acemoglu, D. and Restrepo, P. (2018). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment. *American Economic Review*, 108(6):1488–1542.
- Acemoglu, D. and Restrepo, P. (2019). Artificial Intelligence, Automation and Work. In Agrawal, A. K., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence*. University of Chicago Press.
- Acemoglu, D. and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, 128(6):2188–2244.
- Acemoglu, D. and Restrepo, P. (2021). Demographics and Automation. *The Review of Economic Studies*, 89(1):1–44.
- Aghion, P., Antonin, C., and Bunel, S. (2019). Artificial 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. Working paper.
- Azoulay, P., Graff Zivin, J. S., and Wang, J. (2010). Superstar Extinction. *The Quarterly Journal of Economics*, 125(2):549–589.
- Baker, A., Larcker, D. F., and Wang, C. C. Y. (2021). How much should we trust staggered difference-in-differences estimates? Technical report.
- Bessen, J., Goos, M., Salomons, A., and van den Berge, W. (2020). Firm-Level Automation: Evidence from the Netherlands. *American Economic Association Papers & Proceedings*, 110:389–393.
- Blackwell, M., Iacus, S., King, G., and Porro, G. (2009). CEM: Coarsened Exact Matching in Stata. *Stata Journal*, 9(4):524–546.

- Bonfiglioli, A., Crino, R., Fadinger, H., and Gancia, G. (2021). Robot Imports and Firm-Level Outcomes. Working paper.
- Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. Working paper.
- Brier, G. W. (1950). Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, 78:1–3.
- Callaway, B. and Sant’Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.
- Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.
- Cheng, H., Jia, R., Li, D., and Li, H. (2019). The Rise of Robots in China. *Journal of Economic Perspectives*, 33(2):71–88.
- Clemens, J. and Strain, M. R. (2021). The heterogeneous effects of large and small minimum wage changes: Evidence over the short and medium run using a pre-analysis plan. Working Paper 29264, National Bureau of Economic Research.
- Cortes, G. M. (2016). Where Have the Middle-Wage Workers Gone? A Study of Polarization using Panel Data. *Journal of Labor Economics*, 34(1):63–105.
- 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., Suedekum, J., and Woessner, N. (2021). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association*, 19(6):3104–3153.
- Davis, S. J. and Von Wachter, T. (2011). Recessions and the Costs of Job Loss. *Brookings Papers on Economic Activity*, 42(2):1–71.

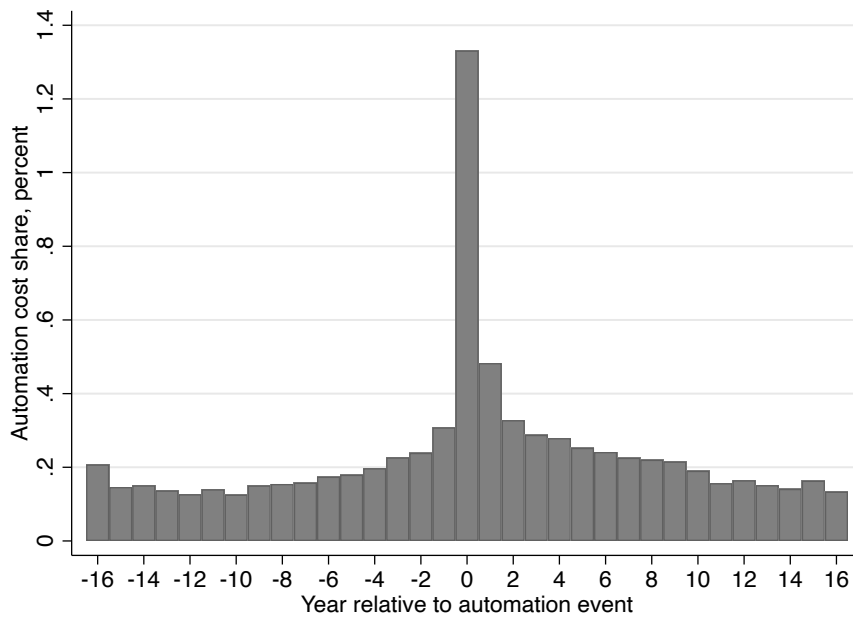
- de Chaisemartin, C. and D’Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.
- Deelen, A., de Graaf-Zijl, M., and van den Berge, W. (2018). Labour Market Effects of Job Displacement for Prime-age and Older Workers. *IZA Journal of Labor Economics*, 7(1):3.
- Deshpande, M. and Li, Y. (2019). Who Is Screened Out? Application Costs and the Targeting of Disability Programs. *American Economic Journal: Economic Policy*, 11(4):213–48.
- 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 3422581, SSRN.
- Doms, M., Dunne, T., and Troske, K. R. (1997). Workers, Wages, and Technology. *The Quarterly Journal of Economics*, 112(1):253–290.
- Doms, M. E. and Dunne, T. (1998). Capital Adjustment Patterns in Manufacturing Plants. *Review of Economic Dynamics*, 1(2):409–429.
- Doraszelski, U. and Jaumandreu, J. (2013). R&D and Productivity: Estimating Endogenous Productivity. *Review of Economic Studies*, 80(4):1338–1383.
- 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.
- Dunne, T. and Troske, K. (2005). Technology adoption and the skill mix of US manufacturing plants. *Scottish Journal of Political Economy*, 52:387–405.
- 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.
- Fadlon, I. and Nielsen, T. H. (2021). Family labor supply responses to severe health shocks: Evidence from Danish administrative records. *American Economic Journal: Applied Economics*, 13(3):1–30.
- Fisher, S. R. A. (1935). *The Design of Experiments*. Macmillan.
- Goldschmidt, D. and Schmieder, J. F. (2017). The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure. *The Quarterly Journal of Economics*, 132(3):1165–1217.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277. Themed Issue: Treatment Effect 1.
- Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–2526.
- Graetz, G. and Michaels, G. (2018). Robots at Work. *Review of Economics and Statistics*, C(5):753–768.
- 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.
- Hirvonen, J., Stenhammar, A., and Tuhkuri, J. (2021). The Employment Consequences of Robots: Firm-Level Evidence. Working Paper.
- Humlum, A. (2021). Robot Adoption and Labor Market Dynamics. Working paper.
- Iacus, S. M., King, G., and Porro, G. (2012). Causal Inference Without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20(1):1–24.

- Illing, H., Schmieder, J. F., and Trenkle, S. (2021). The Gender Gap in Earnings Losses after Job Displacement. NBER Discussion Paper 29251.
- Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings Losses of Displaced Workers. *The American Economic Review*, 83(4):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. (2021). Robots and Firms. *The Economic Journal*, 131(638):2553–2584.
- 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):152–190.
- 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.
- Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–1297.
- Pew (2017). Automation in Everyday Life. Technical report, Pew Research Center.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.
- 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.

Figures

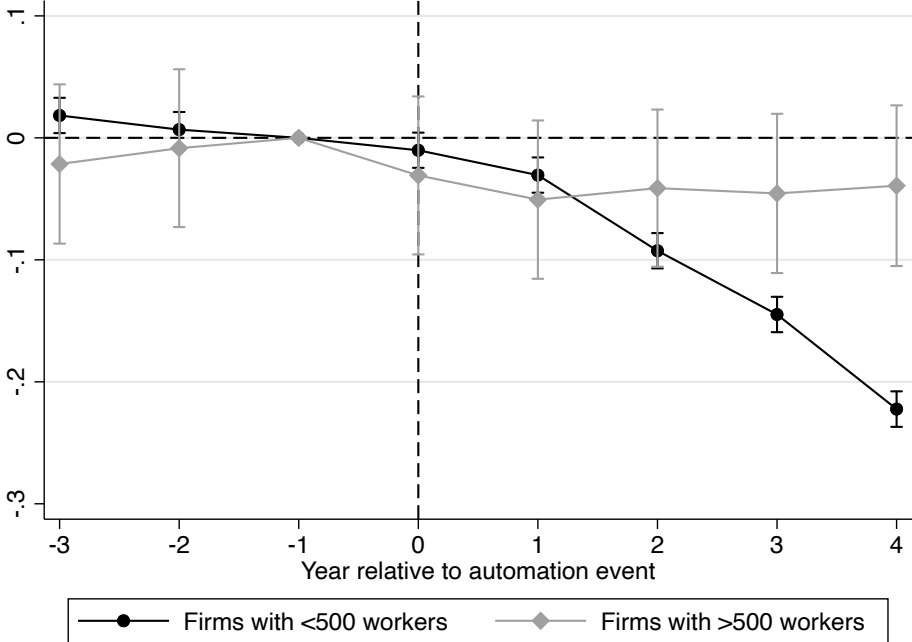
Figure 1. Automation cost shares around automation events



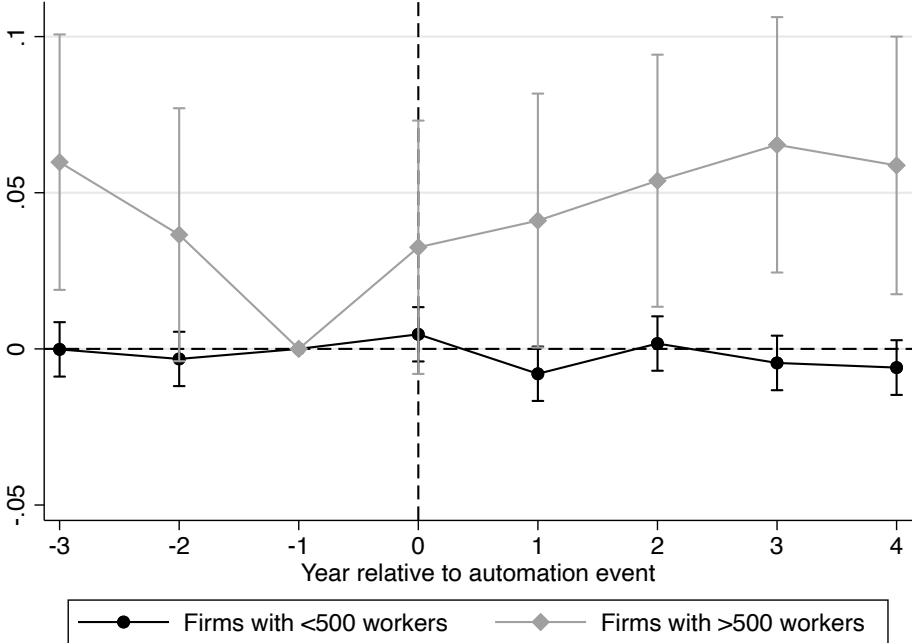
Notes: $N = 10,425$ in $\tau = 0$.

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

(a) Log employment change



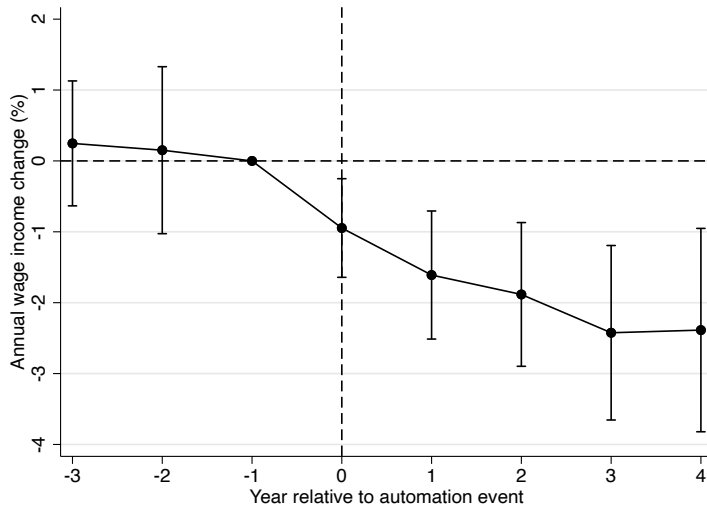
(b) Log daily wage change



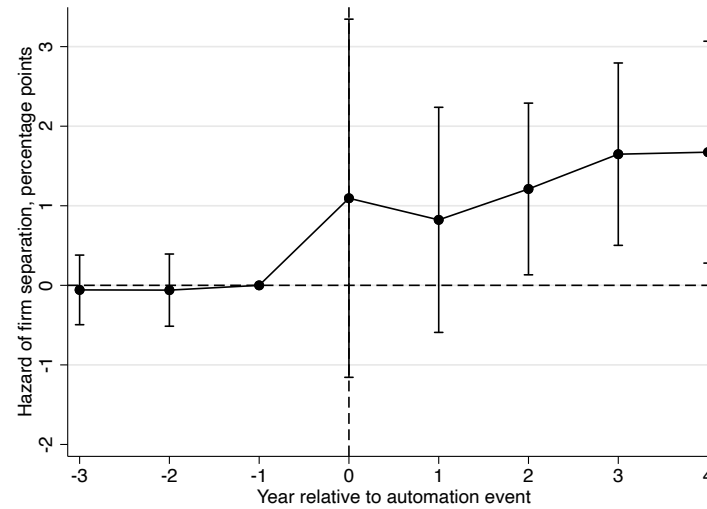
Notes: All automating firms that exist in all years $\tau \in \{-3, \dots, 4\}$. Both models are weighted by firm-level employment size in $\tau = -1$. Whiskers reflect 95% confidence intervals.

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

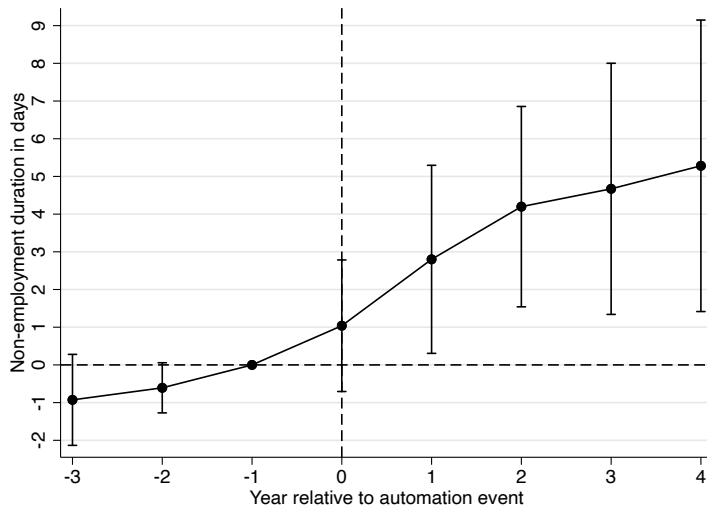
(a) Relative annual wage income



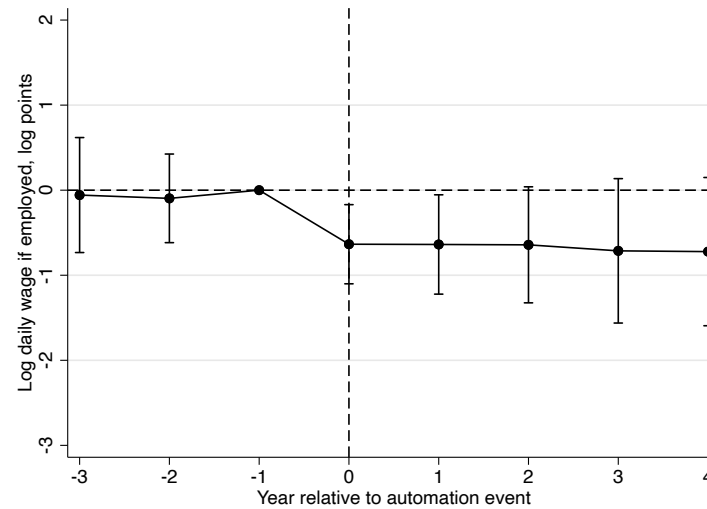
(b) Hazard rate of leaving the firm



(c) Days in non-employment



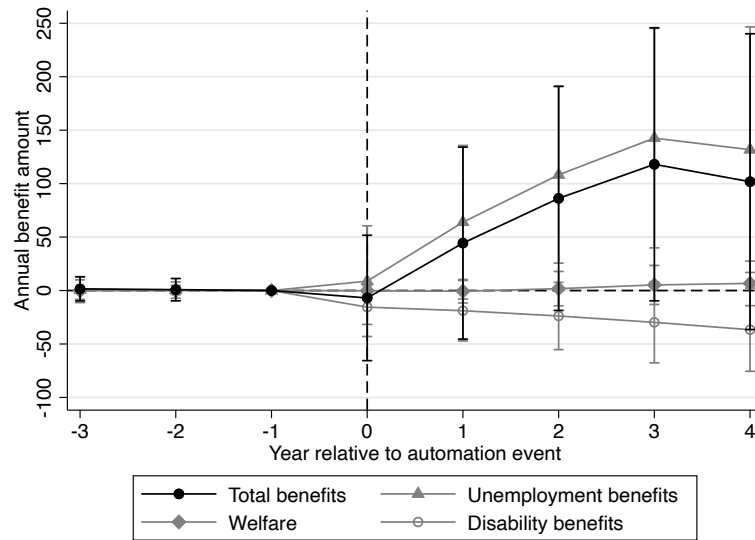
(d) Log daily wage



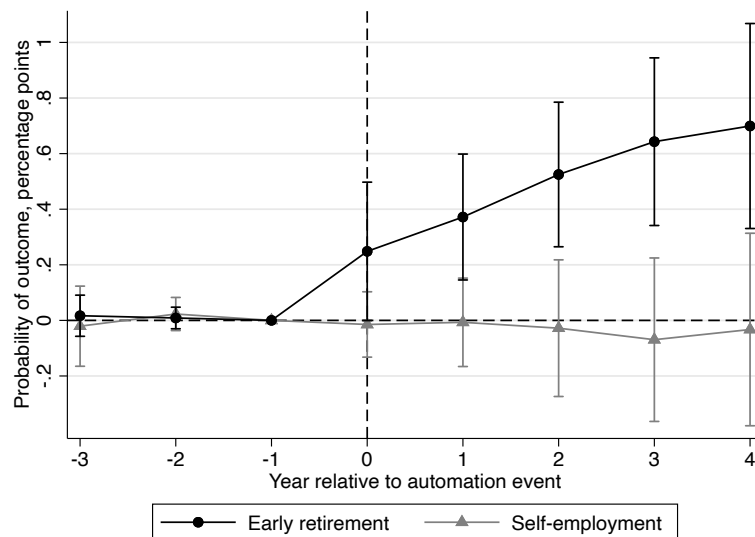
Notes: $N = 8,791,392$ for annual real wage income and days in non-employment, $N = 8,021,759$ for hazard of firm separation, $N = 8,429,129$ for log daily wages. Whiskers represent 95 percent confidence intervals.

Figure 4. Effect of automation on annual benefit income and early retirement

(a) Annual benefit income

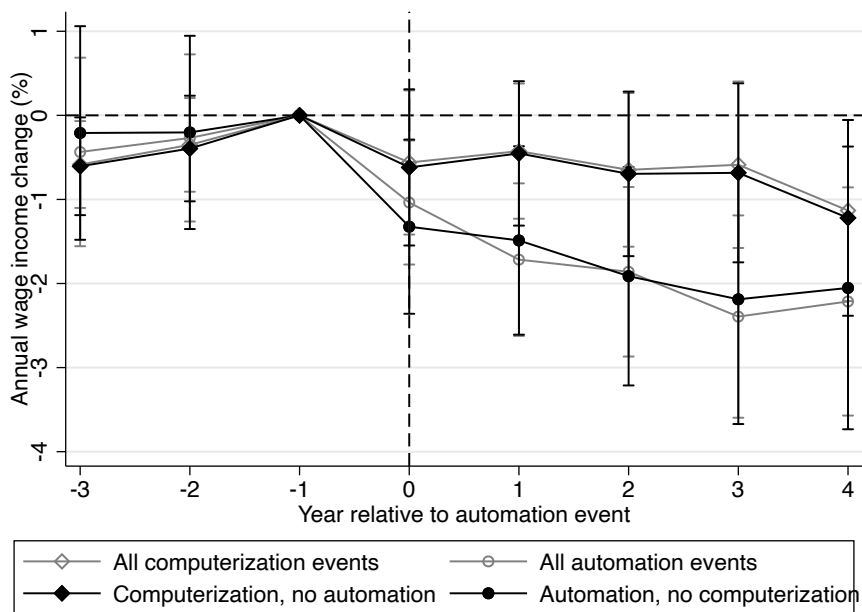


(b) Early retirement and self-employment



Notes: $N = 8,791,392$. Whiskers represent 95 percent confidence intervals.

Figure 5. Relative annual wage income effects of automation and computerization



Notes: All estimates are for the overlapping sample where we observe data on both automation costs and computer investments. $N = 10,022,288$ for all computerization; $N = 8,525,056$ for all computerization excluding automation; $N = 8,499,112$ for all automation; and $N = 5,167,464$ for automation excluding computerization.

Tables

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	0.00	2,211	59	0.04
p10	0	0	0.00	3,987	101	0.06
p25	0	0	0.00	10,487	256	0.14
p50	11,736	283	0.16	30,000	641	0.32
p75	52,824	986	0.47	93,711	1,446	0.68
p90	192,458	2,256	1.06	305,290	2,949	1.37
p95	453,582	3,625	1.69	713,180	4,591	2.13
<i>mean</i>	<i>211,307</i>	<i>1,045</i>	<i>0.44</i>	<i>307,794</i>	<i>1,522</i>	<i>0.64</i>
N firms × years		238,713			163,881	
Pct 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.

Table 2. Automation costs by sector and firm size class

	Total cost	Cost per worker		Cost share (%)		Nr of obs	
	<i>Mean</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Firms</i>	<i>Firms × yrs</i>
A. Sector							
Manufacturing	430,125	1,078	7,065	0.36	0.58	5,513	44,375
Construction	78,225	457	1,463	0.20	0.36	4,429	28,182
Wholesale & retail trade	116,252	1,174	4,176	0.31	0.80	10,918	75,277
Transportation & storage	280,151	909	4,962	0.41	1.07	3,124	21,270
Accommodation & food serving	55,624	244	896	0.30	0.50	1,184	6,545
Information & communication	447,852	1,789	25,852	0.85	2.93	2,627	16,801
Prof'l, scientific & technical activities	149,069	1,283	5,490	1.02	1.75	3,977	23,581
Administrative & support activities	133,887	839	18,825	0.50	1.18	3,808	22,682
B. Firm size class							
1-19 employees	12,269	921	14,568	0.40	1.30	9,499	48,073
20-49 employees	27,689	893	4,547	0.42	1.34	13,424	86,553
50-99 employees	61,601	954	4,345	0.42	0.96	6,192	47,072
100-199 employees	144,891	1,135	5,812	0.44	0.94	3,413	28,666
200-499 employees	406,461	1,573	21,305	0.51	1.11	1,943	17,868
≥500 employees	3,161,867	2,124	14,294	0.76	1.60	1,109	10,481

Notes: Automation costs in 2015 euros, automation cost shares as a percentage of total costs, excluding automation costs. Total N firms is 35,580; Total N firms × years is 238,713.

Table 3. Firm-level outcomes for automating vs non-automating firms

	Δ log employment		Δ log daily wage		Δ log wage bill	
	(1)	(2)	(3)	(4)	(5)	(6)
Automating	0.019*** (0.006)	0.015*** (0.006)	-0.002 (0.003)	-0.002 (0.002)	0.016*** (0.005)	0.013*** (0.005)
Additional controls	No	Yes	No	Yes	No	Yes

Notes: N = 168,091 firm-year observations, where 10,425 out of 35,580 unique firms automate. All models include calendar year fixed effects. 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 average firm-level employment size. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

Table 4. Heterogeneity in incumbent worker impacts

	Annual wage income	Separation hazard	Days non- employed	Log daily wage
<i>A. Firm size</i>				
1–19 employees (reference)	-3.17*** (0.76)	1.83*** (0.37)	4.25*** (1.39)	-2.21*** (0.48)
<i>Deviations from reference group for:</i>				
20–49 employees	0.23 (0.91)	-0.13 (0.46)	-0.19 (1.64)	0.87 (0.57)
50–99 employees	2.42** (0.96)	-0.47 (0.57)	-3.26* (1.78)	1.80*** (0.62)
100–199 employees	1.35 (1.11)	-1.38** (0.69)	-1.93 (2.09)	1.00 (0.69)
200–499 employees	2.21* (1.16)	0.07 (0.90)	-1.21 (2.38)	2.54*** (0.74)
≥500 employees	0.77 (1.51)	-0.17 (2.57)	1.25 (3.19)	1.77* (0.94)
N	8,791,392	8,021,759	8,791,392	8,429,129
<i>B. Worker age</i>				
Age ≥50 (reference)	-3.97*** (1.25)	2.82*** (0.92)	10.14*** (2.28)	-0.45 (0.56)
<i>Deviations from reference group for:</i>				
Age 40–49	2.62* (1.36)	-1.17** (0.58)	-7.56*** (2.64)	-0.15 (0.49)
Age 30–39	2.25* (1.27)	-1.63** (0.74)	-7.39*** (2.68)	-0.55 (0.66)
Age 20–29	3.20* (1.71)	-1.54 (1.07)	-9.15*** (2.70)	0.40 (0.91)
N	8,791,392	8,021,759	8,791,392	8,429,129
<i>C. Worker education level</i>				
Medium education (reference)	-2.54*** (0.77)	2.81 (1.97)	5.25** (2.21)	-1.13*** (0.41)
<i>Deviations from reference group for:</i>				
Low education	0.86 (1.47)	-1.42 (1.33)	-0.78 (2.34)	0.76 (0.90)
High education	1.21* (0.70)	0.30 (1.31)	-2.15 (1.64)	0.78 (0.70)
N	2,511,583	2,177,563	2,511,583	2,413,426

Notes: All coefficients are average annual effects over the post-treatment period ($\tau = 0$ through $\tau = 4$). Coefficients for wage income, separation hazard, and log daily wages have been multiplied by 100 to reflect percentages. Education is classified using ISCED-2011: low is up to and including lower secondary education; medium is upper secondary or post-secondary education excl. tertiary education; and high education is tertiary education. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendices

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

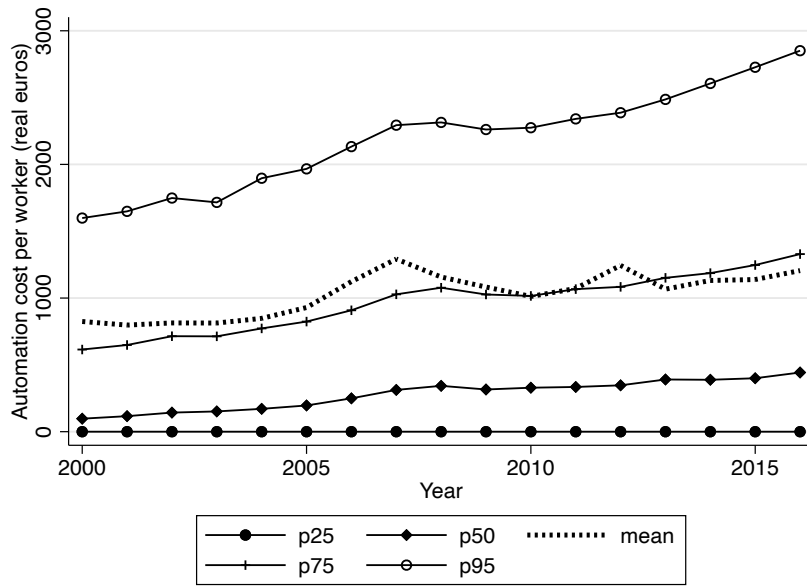
A Data

A.1 Automation costs over time

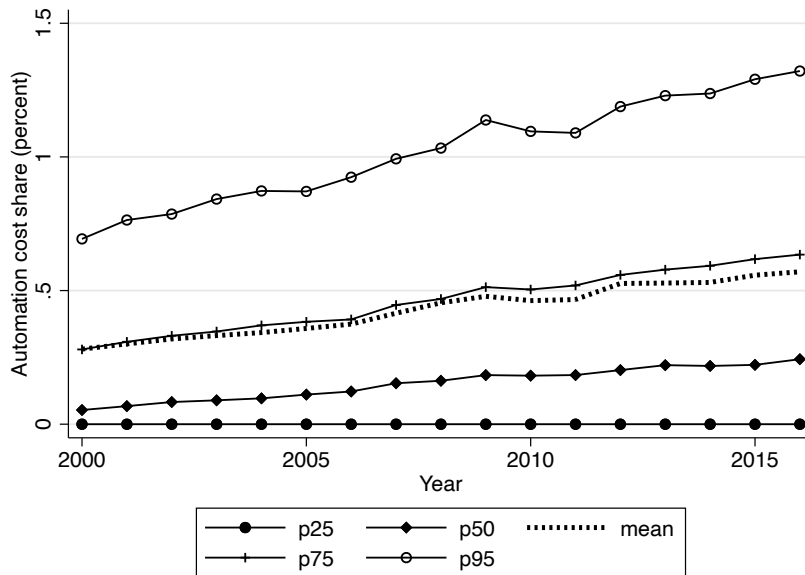
Figure A.1 shows how the distribution of automation costs per worker (top panel) and automation cost shares (bottom panel) have changed over time. Mean automation costs per worker and the mean automation cost share are rising over 2000–2016. Furthermore, besides increases in means, there is a fanning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

Figure A.1. Automation costs over time

(a) Automation costs per worker



(b) Automation costs shares



A.2 Correlations between automation costs and specific technologies

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

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

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

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

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

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

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

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

A.3 Technology usage by sector

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

Table A.3. Technology usage by sector

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

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

A.4 Automation imports

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

Detailed product codes for each of these are as follows:

- Automatically controlled machines:
90321080, 90321000, 90328100, 90320000, 90321020, 90328900, 90328100, 90329000, 90322000
- Automatic transfer machines:
84283100, 84283900, 84573090, 84283300, 84283200, 84283990, 84580000, 84283100, 84283920, 84573000, 84573010
- Automatic welding machines:
85153100, 85153100, 85152100, 85152100
- Numerically controlled machines:
845811000080, 845811200080, 845811410010, 845811410080, 845811490080, 845811800080, 845891000010, 845891000080, 845891200080, 845891800080, 845921000010, 845921000080, 845931000010, 845931000080, 845941000010,

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

- Robots: 84795000

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

Table A.4 compares automation costs and automation imports as a percentage of total operating cost for the overlapping subsample of firms at the sector level⁵³, revealing that average automation expenditures are substantially higher than average automation imports – since few firms are importers–, and observed across a wider range of sectors. Automation imports and automation expenditures are somewhat correlated at the firm-level, as shown in the first two columns of Table A.5 where firm-by-year automation expenditures are regressed onto (net) automation imports while controlling for firms’ total operating cost, and additionally year fixed effects in the second column. However, this correlation disappears when adding firm fixed effects (seen from the last two columns of Table A.5): that is, firms are not more likely to have higher automation costs when they (net) import more automation technology.

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

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

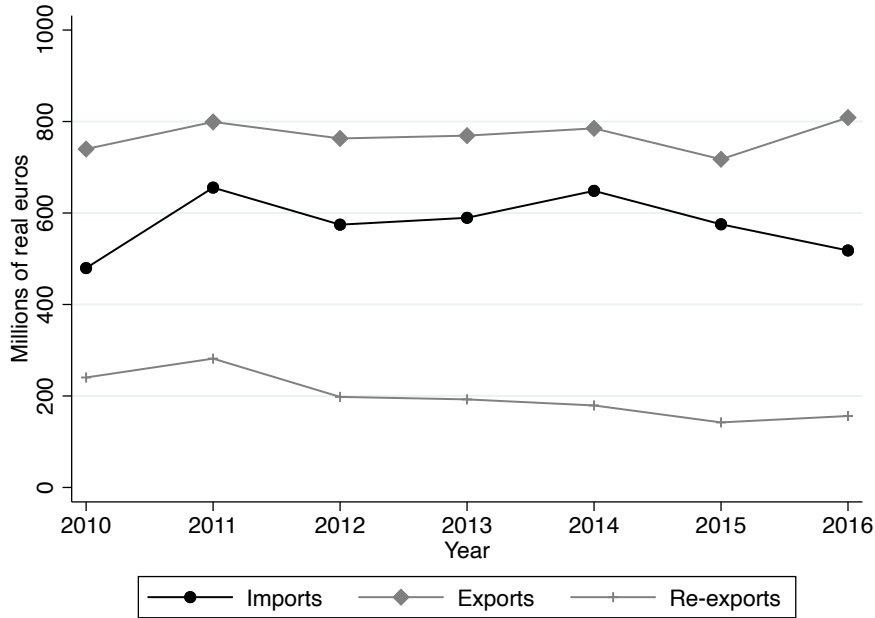


Table A.4. Comparing automation costs to automation imports by sector

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

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

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

<i>Dependent variable: Automation costs (IHS)</i>				
	(1)	(2)	(3)	(4)
Automation imports (IHS)	0.018** (0.007)	0.018** (0.007)	-0.001 (0.004)	-0.002 (0.004)
	(5)	(6)	(7)	(8)
Net automation imports (IHS)	0.016* (0.006)	0.016* (0.006)	-0.003 (0.004)	-0.003 (0.004)
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Log total costs	Yes	Yes	Yes	Yes

Notes: 110,805 firm-year observations for each model. Automation costs, imports, and net imports are transformed using the inverse hyperbolic sine (IHS). Net automation imports are defined as imports minus re-exports. All models control for log total costs at the firm-year level. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

B Theoretical model of firm-level automation spikes

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

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

B.1 Exogenous automation

B.1.1 Consumption

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

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

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

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

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

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

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

$$Y_j = Y \epsilon_j^{\sigma-1} P_j^{-\sigma} \quad (\text{B.3})$$

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

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

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

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

B.1.2 Production

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

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

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

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

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

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

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

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

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

increased by the same proportion.

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

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

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

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

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

B.1.3 Conditional factor demands

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

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

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

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

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

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

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

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

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

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

$$RK_j = I \frac{\sigma-1}{\sigma} P_j Y_j \quad (\text{B.11})$$

and

$$WL_j = [1 - I] \frac{\sigma-1}{\sigma} P_j Y_j \quad (\text{B.12})$$

which gives the firm's conditional factor demands.

B.1.4 Output as a Cobb-Douglas aggregate

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

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

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

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

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

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

where φ_j is defined as:

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

B.1.5 Unconditional labor demand

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

$$\frac{K_j}{I} = \frac{1}{1-I} \frac{W}{R} L_j \quad (\text{B.15})$$

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

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

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

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

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

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

B.1.6 Output

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

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

with

$$P_j = \frac{\sigma}{\sigma-1} \frac{W^{1-I} R^I}{\varphi_j} \quad (\text{B.19})$$

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

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

B.1.7 Profits

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

$$\Pi_j = \frac{P_j Y_j}{\sigma} = \frac{Y \epsilon_j^{\sigma-1}}{\sigma} \left[\frac{\sigma}{\sigma-1} \frac{W^{1-I} R^I}{\varphi_j} \right]^{1-\sigma} \quad (\text{B.20})$$

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

B.2 Endogenous automation

B.2.1 Expected profits

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

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

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

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

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

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

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

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

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

B.2.2 Technological progress

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

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

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

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

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

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

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

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

B.2.3 Spikes in automation expenditures

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

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

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

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

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

B.2.4 Shocks to product demand

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

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

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

B.3 Empirical implications

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

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

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

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

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

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

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

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

Related theories make similar assumptions. Bonfiglioli et al. (2021) present a static

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

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

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

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

at the incumbent worker level.

C Automation events

C.1 Automation spike frequency

Table C.1. Spike frequency

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

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

C.2 Automation events across sectors and firm size classes

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

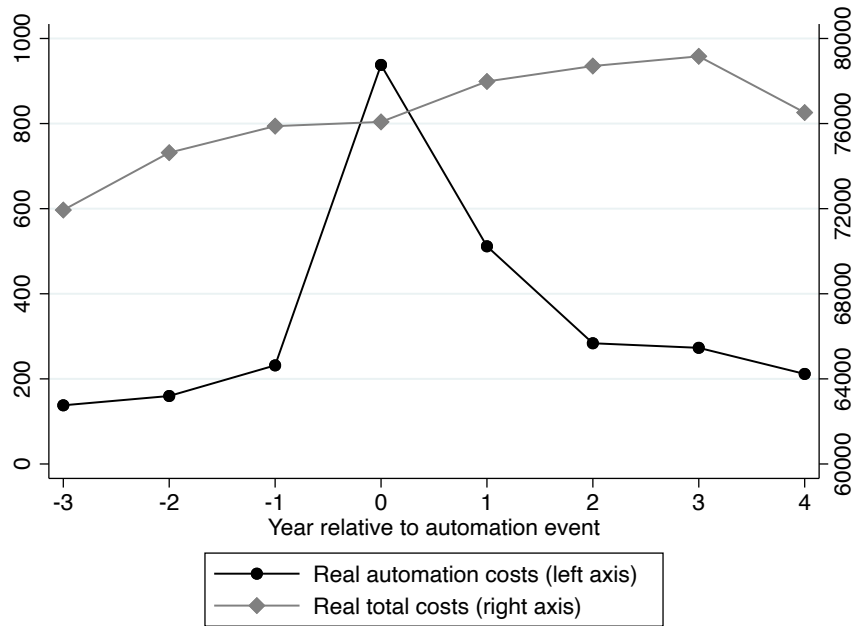
By sector	
Manufacturing	0.29
Construction	0.25
Wholesale & retail trade	0.27
Transportation & storage	0.30
Accommodation & food serving	0.27
Information & communication	0.39
Prof'l, scientific & technical activities	0.33
Administrative & support activities	0.30
By firm size	
1-19 employees	0.26
20-49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

Notes: N=35,580 firms.

C.3 Automation and total costs around automation events

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

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



Notes: Balanced sample of firms.

D Firm-level analyses

D.1 Predicting automation events

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

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

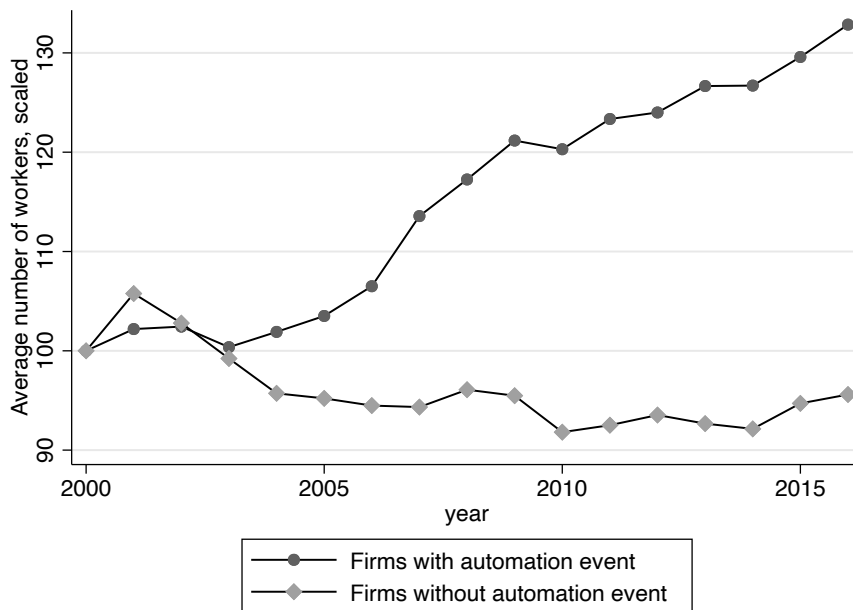
Mean annual wage	0.0009*** (0.0002)	Manufacturing	<i>reference</i>
Share of women	-0.0259** (0.0126)	Construction	-0.0003 (0.0094)
Mean worker age	-0.0034*** (0.0005)	Wholesale & retail trade	0.0161** (0.0080)
Share high educated	0.0368* (0.0197)	Transportation & storage	0.0414*** (0.0102)
1–19 employees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

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

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

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

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



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

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

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

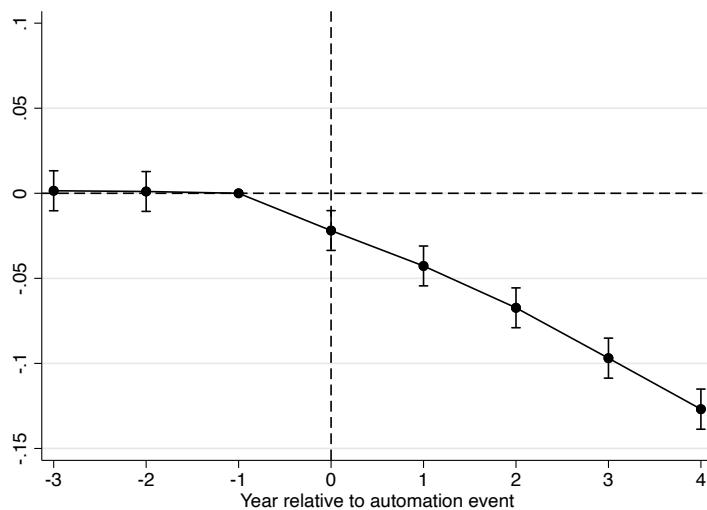
Figure D.2 reports the findings for employment and wages. All estimates are weighted by firm size. The top figures reveal that for employment there are clear divergent pre-trends between automating and non-automating firms (right), which do not seem to exist

when comparing automating with later-automating firms (left). By contrast, for wages the pre-trends appear quite similar, while we find almost no significant impacts on wages using non-automating firms and small positive impacts using later-automating firms as control group.

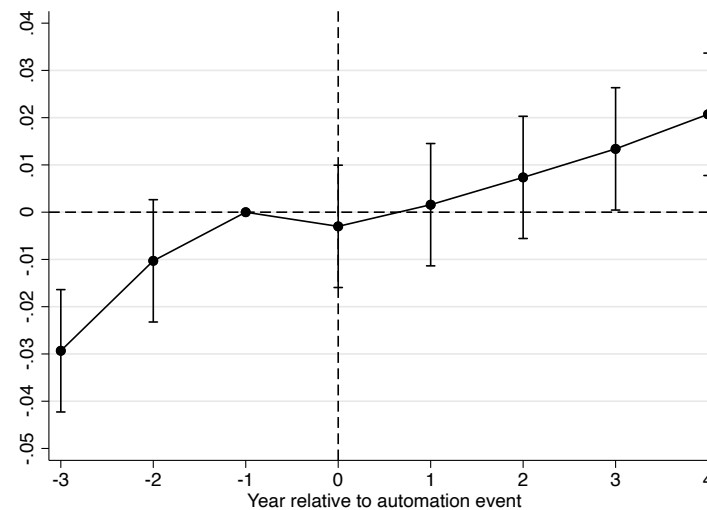
An important limitation of this analysis is that we do not observe automation costs for firms in each year. Hence, we cannot be sure that firms do not automate in a year where we do not observe them in the survey. Restricting the sample to firms that we observe each year would leave us with a very small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

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

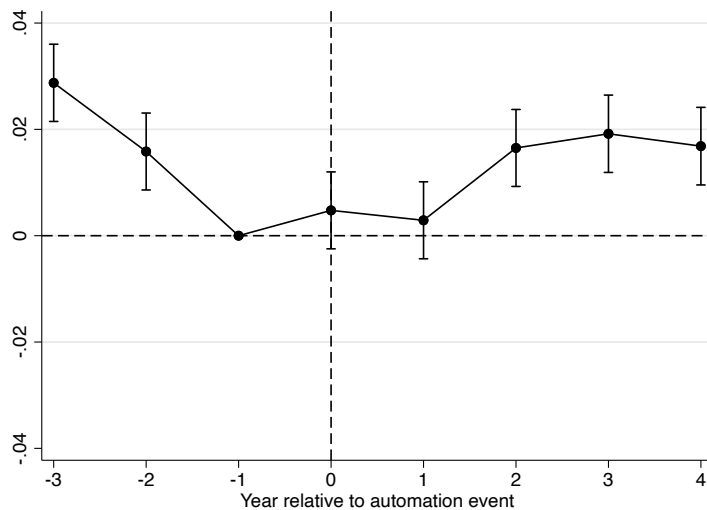
(a) Log employment (baseline)



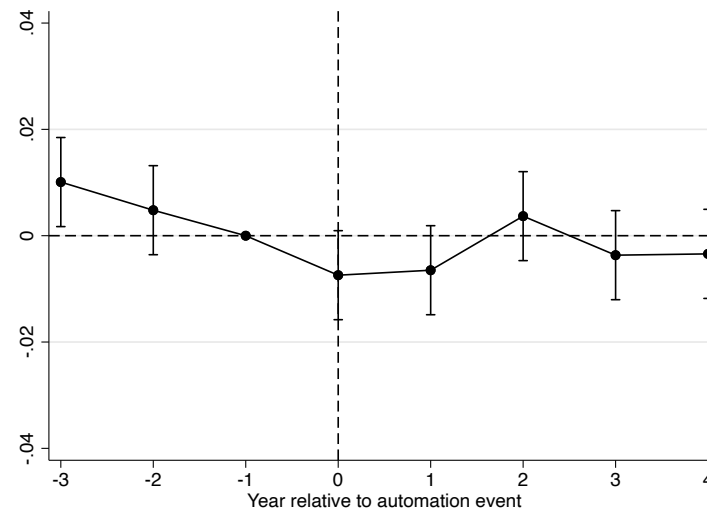
(b) Log employment (using non-automating firms)



(c) Log daily wage (baseline)



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



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

D.4 Comparison to import-based automation measure

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

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

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

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

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

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

<i>Dependent variable: Log firm-level number of employees</i>				
	Automation cost spike		Automation imports	
	(1)	(2)	(3)	(4)
Automating	0.078*** (0.013)	0.085*** (0.013)	0.856*** (0.022)	0.839*** (0.022)
Sector fixed effects	No	Yes	No	Yes

Notes: N = 30,291 firm-level observations. Automation imports measured as non-zero mean automation imports at the firm level. Sector fixed effects are two-digit sector dummies. *p<0.10, **p<0.05, ***p<0.01.

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

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

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

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

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

Notes: N = 30,291 firm observations, where 31% of firms have automation cost spikes, and 8.2% (7.9%) have non-zero (net) imports. Controls are log total costs and sector fixed effects. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

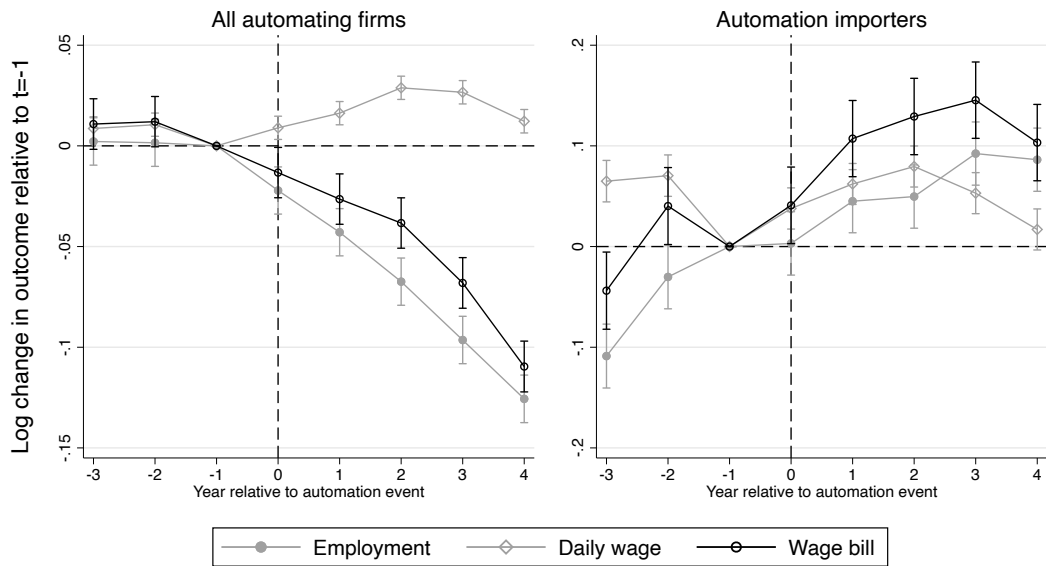
Table D.5. Within-firm correlation between automation events and automation imports

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

Notes: N = 110,805 firm-year observations. Standard errors are clustered at the firm-level. *p<0.10, **p<0.05, ***p<0.01.

and wage growth around automation events.

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



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

E Worker-level analyses

E.1 Sample construction

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

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

We then merge our firm-level data to worker data and keep only events for which we can find at least one incumbent worker who is between 18 and 65 years old at $\tau = -1$. This leaves us with 2,995 potential treatment events merged to 192,755 incumbent workers

and 21,115 potential control events merged to 1,132,190 incumbent workers. We then apply some sample selections to drop outliers. In particular, we drop workers with yearly earnings more than 500,000 euros in any one year or an average daily wage above 2,000 euros. We also exclude students. Finally, we require incumbents to not receive any benefits in the three years before treatment. This leaves us with 997,057 potential control workers and 162,493 potential treated workers.⁶⁶

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

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

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

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

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

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

also serve as treatment event in an earlier year.

E.2 Summary statistics for workers

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

Table E.2. Descriptives for incumbent workers

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

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

E.3 Predicting automation event timing

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

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

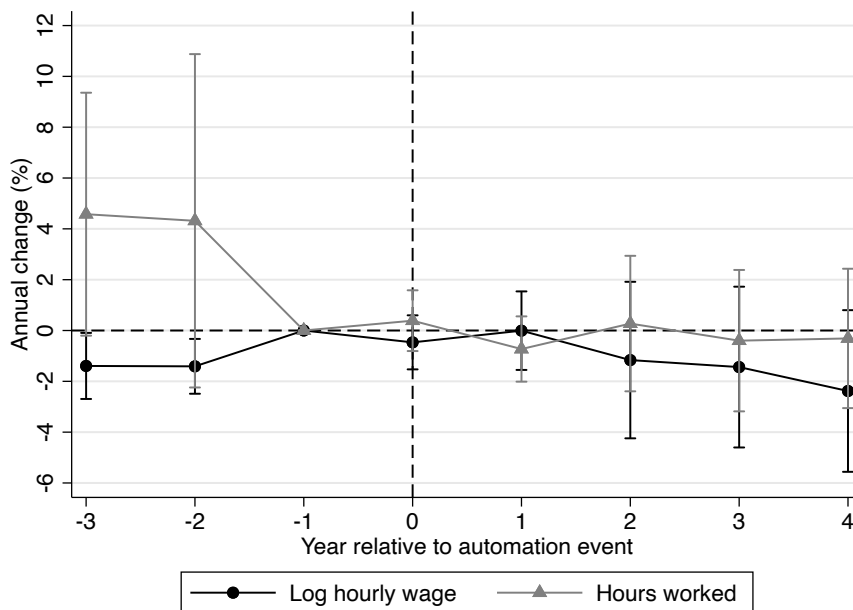
Table E.3. Brier skill scores for predicting automation event timing

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

E.4 Effects on hourly wages

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

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



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

E.5 Effect heterogeneity

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

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

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

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

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

In column 1 of Table E.4, we consider to what extent the impacts of automation differ

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

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

While we also do not find statistically significant differences by workers’ contract type

(column 3), but the estimates suggest losses are larger for workers with flexible contracts as opposed to open-ended contracts. This could reflect differences in employment protection.

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

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

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

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

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

(1) Sector		(3) Contract type	
Manufacturing (reference)	-1.66** (0.83)	Open-ended contract (reference)	-1.74*** (0.44)
<i>Deviations from reference group for:</i>		<i>Deviation from reference group for:</i>	
Construction	0.25 (1.52)	Flexible contract	-2.34 (3.11)
Wholesale & retail trade	-0.58 (1.13)	(4) Overall age-specific wage quartile	
Transportation & storage	1.41 (1.52)	Bottom quartile (reference)	-2.10* (1.24)
Accommodation & food serving	2.89** (1.43)	<i>Deviations from reference group for:</i>	
Information & communication	-1.05 (1.55)	Second quartile	-0.06 (1.20)
Prof'l, scientific, & technical activities	-0.95 (1.53)	Third quartile	0.44 (1.24)
Administrative & support activities	-1.09 (2.46)	Top quartile	0.14
		(5) Within-firm age-specific wage quartile	
		Bottom quartile (reference)	-1.38 (1.78)
		<i>Deviations from reference group for:</i>	
		Second quartile	-0.86 (2.12)
		Third quartile	-1.02 (2.22)
		Top quartile	-0.26 (1.77)
(2) Gender			
Male (reference)	-1.55*** (0.56)		
<i>Deviation from reference group for:</i>			
Female	-0.88 (0.73)		

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

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

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

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

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

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

E.6 Incumbent workers versus recent hires

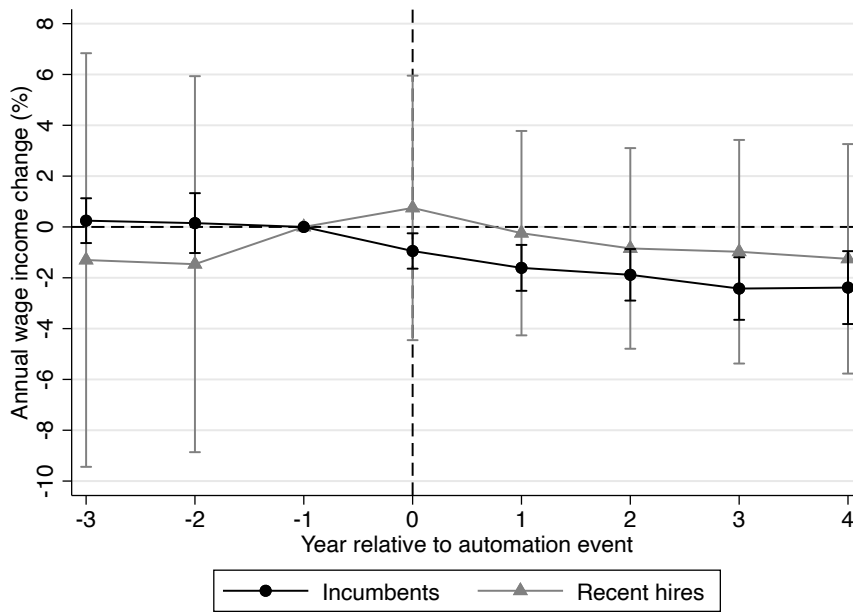
Our identification strategy for the impacts of automation is to consider individual workers who have a pre-existing working relationship with the firm, as evidenced by at least three years of firm tenure. Here we estimate our models for a second group of workers: those with less than three years of firm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

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

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

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

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



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

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

E.7 Robustness tests

E.7.1 Constructing placebo events

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

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

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

	Automation cost		Other material fixed assets	
	<i>level</i>	<i>per worker</i>	<i>level</i>	<i>per worker</i>
p5	0	0	0	0
p10	0	0	0	0
p25	0	0	0	0
p50	18,285	324	1,213	23
p75	75,758	1,043	34,277	456
p90	263,129	2,373	180,821	1,684
p95	620,724	3,839	473,242	3,344
mean	271,888	1,125	181,772	1,067
mean excl. zeros	377,964	1,564	349,980	2,054
N firms \times years	171,875		171,875	
N firms \times years with 0 costs	48,237		82,607	

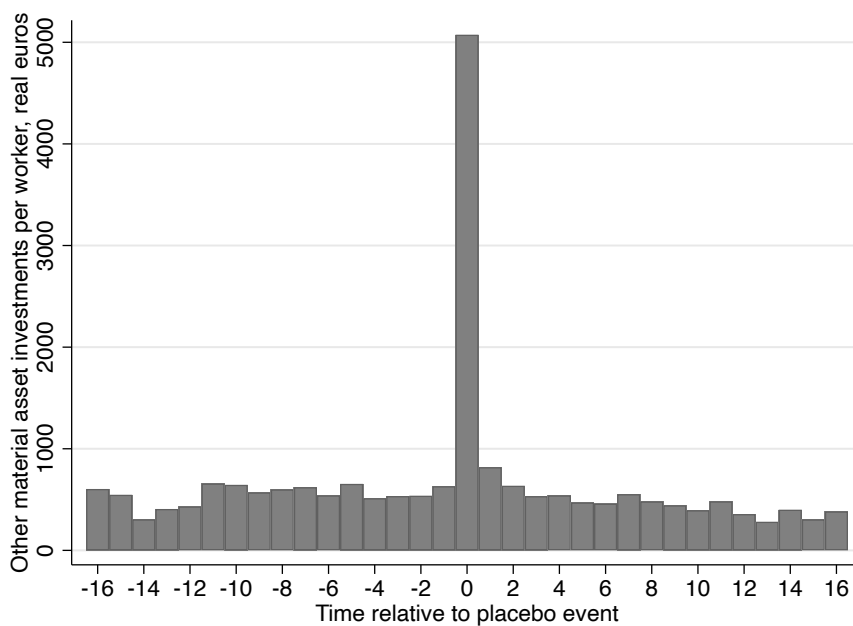
Notes: All numbers are in 2015 euros.

Table E.7. Automation and placebo events

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

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

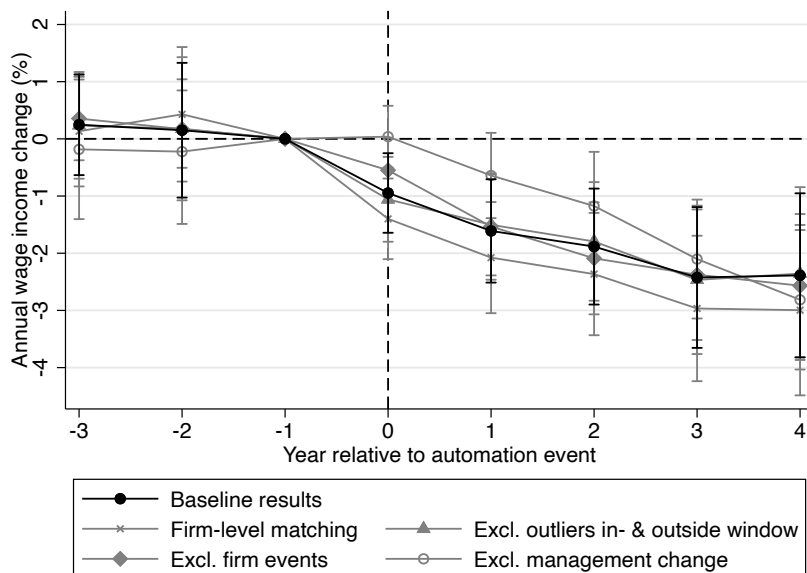
Figure E.3. Placebo events: Spikes in other material fixed assets



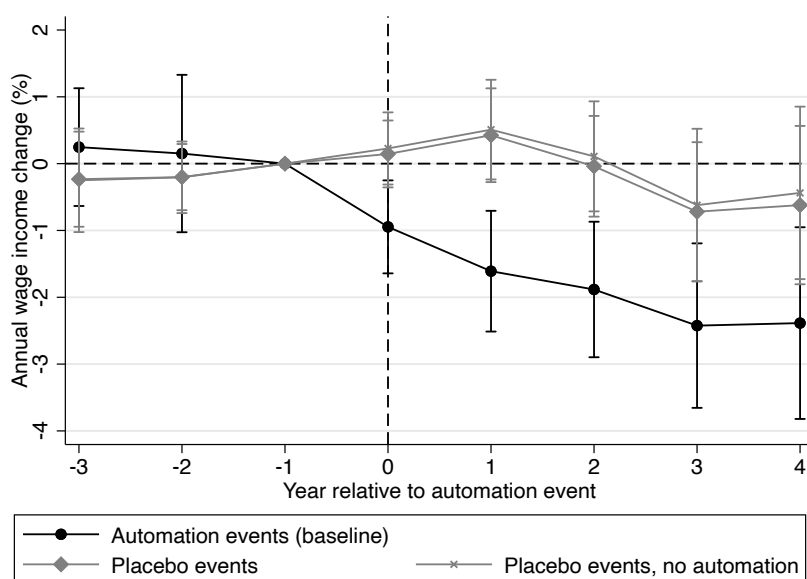
E.7.2 Robustness to other firm events and placebo events

Figure E.4. Robustness tests

(a) Robustness to other firm events



(b) Comparison to placebo events



E.7.3 Alternative definitions of automation events

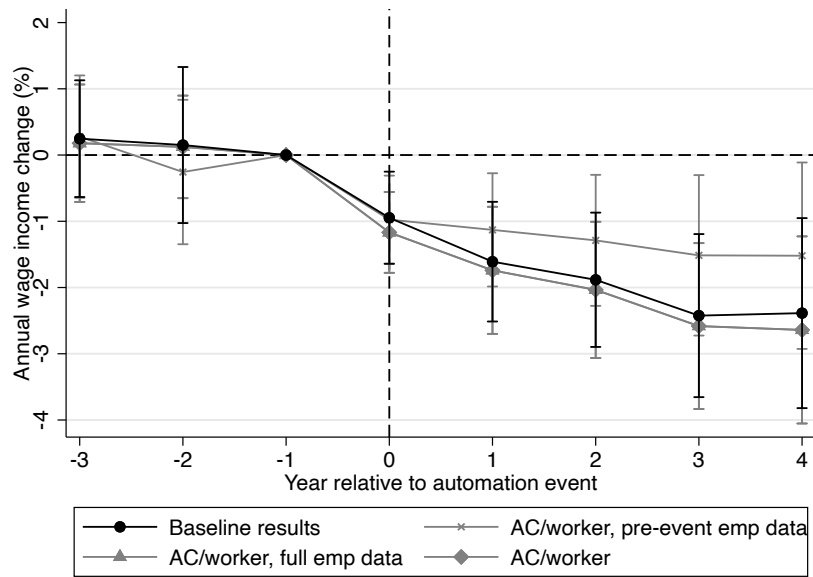
Rather than using automation cost shares (i.e. automation costs in total costs), we can construct automation events from sharp increases in automation outlays per worker.

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

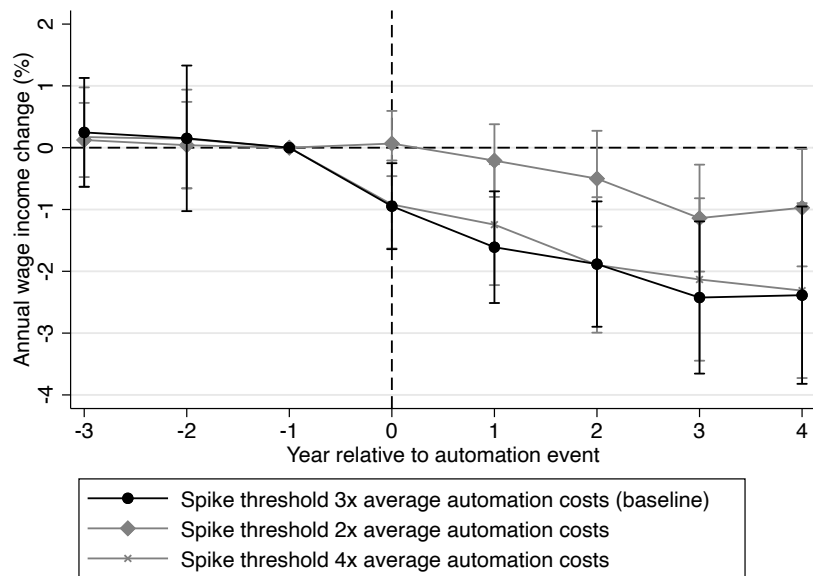
Further, we show that results are robust to varying the spike threshold from two to four times the average automation costs (our baseline is thrice the average automation costs). Panel (b) in Figure E.5 reveals that estimated effect sizes are somewhat larger the higher the threshold, as expected, but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

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

(a) Automation costs per worker



(b) Changes in spike threshold for automation cost shares

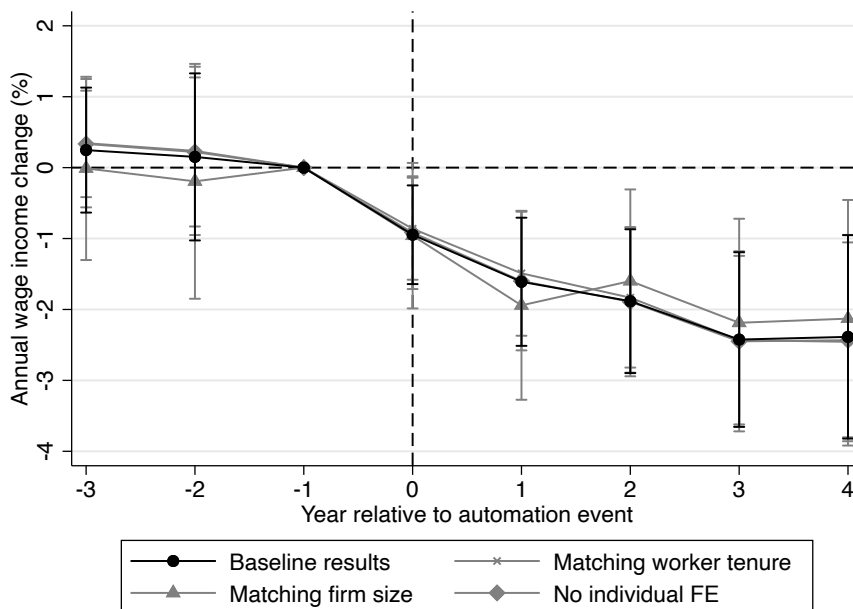


E.7.4 Changes in model specification

Here, we change our model specification in a number of ways. In particular, compared to our baseline estimates, Figure E.6 shows results when additionally matching workers on their firm tenure in years (that is, beyond the three years of firm tenure that all treated and control group workers have); additionally matching workers on firm size; and when removing individual fixed effects from the model (these are then replaced by dummies

for worker gender and nationality, as well as fixed effects for firm size categories, and for firm sector). Although estimates without individual fixed effects are a little less precise, results are extremely robust to these changes in specification.

Figure E.6. Robustness to changes in model specification



E.7.5 Randomization test

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

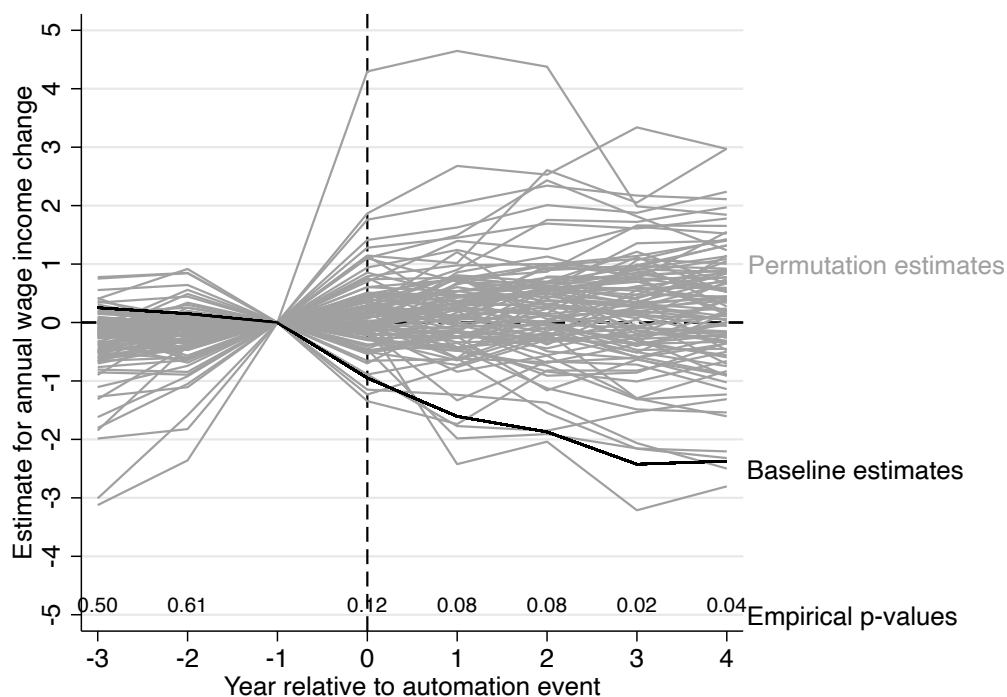
Results are shown in Figure E.7: each gray line presents a set of placebo (dynamic) treatment estimates, whereas the black line presents our actual treatment estimates. The

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

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

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

Figure E.7. Randomization test



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

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

F Computer investments

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

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

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

Table F.1. Automation costs and computer investments distributions

	Automation cost		Computer investment	
	<i>level</i>	<i>per worker</i>	<i>level</i>	<i>per worker</i>
p5	0	0	0	0
p10	0	0	0	0
p25	0	0	0	0
p50	18,285	324	6,046	108
p75	75,758	1,043	33,892	488
p90	263,129	2,373	123,000	1,229
p95	620,724	3,839	273,263	2,039
mean	271,884	1,125	109,390	615
mean excl. zeros	377,959	1,564	170,810	960
N firms × years	171,878		171,878	
N firms × years with 0 costs	48,238		61,804	

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

Table F.2. Automation and computerization events

Nr of spikes	Percentage of firms with event type:	
	Automation	Computerization
0	71.9	47.9
1	22.5	41.9
2	4.8	9.1
3	0.7	1.1
4 or 5	0.1	0.1

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

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

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

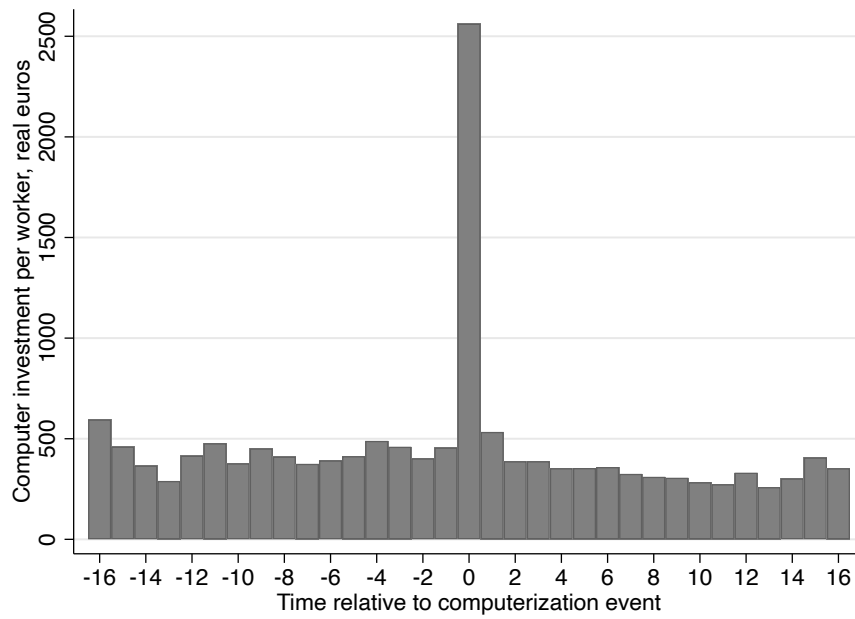
Notes: Overlapping sample, total number of firms is 25,118.

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

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

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

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



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